



ACADEMIA DE FARMACIA DE GALICIA

Discurso de ingreso
como Académico correspondiente

**BIOEQUIVALENCIA Y RIESGO
DEL PACIENTE**

ILMO. SR. DR. MATÍAS A. LLABRÉS MARTÍNEZ



Santiago de Compostela, Abril de 2009

ACADEMIA DE FARMACIA DE GALICIA

Discurso de ingreso como Académico Correspondiente

BIOEQUIVALENCIA Y RIESGO DEL PACIENTE

Ilmo. Sr. Dr. Matías A. Llabrés Martínez

Santiago de Compostela

Abril de 2009

Imprime: EL PRODUCTOR, S.L. *Técnicas Gráficas*

C/ Barrio Nuevo de Ofra, 12- 38320 La Cuesta-La Laguna

Depósito Legal: TF-886-2009

Índice

1	Introducción	1
2	Bioequivalencia de promedios.....	2
2.1	Variabilidad intraindividual y limitaciones de la bioequivalencia de promedios.....	12
3	Bioequivalencia individual y bioequivalencia poblacional	18
3.1	Bioequivalencia individual	22
3.2	Bioequivalencia poblacional.....	23
3.3	Criterios de bioequivalencia	23
4	Relevancia terapéutica de la bioequivalencia individual.....	24

Excmo. Sr. Presidente de la Academia de Farmacia de Galicia,

Excmos. e Ilmos. Sras. y Sres. Académicos,

Sras. y Sres,

Quiero expresar en primer lugar mi gratitud por la distinción que me han otorgado al elegirme Académico Correspondiente de esta Institución; mi satisfacción es hoy doble al verme acogido por compañeros con los que he compartido momentos que determinaron un futuro común y con los que mantengo esa especial amistad que sólo nace de la ilusión y de la complicidad imprescindibles para desarrollar cualquier proyecto.

Mi vinculación con Santiago de Compostela es especial. De niño correteé por sus rúas, aquí cursé la licenciatura de Farmacia, desarrollé gran parte de mi carrera profesional y junto con Queta formamos nuestra familia. Lo que no tenía programado cuando ingresé como alumno de la Facultad de Farmacia son los acontecimientos que me han traído aquí, en no poca medida determinados por el ambiente que encontré en el Laboratorio de Farmacia Galénica, una mezcla poco frecuente de acogimiento, disciplina en el trabajo, ilusión y honestidad intelectual, marcada por la personalidad del profesor D. Rafael Cadórniga al que siempre recordaré con afecto y gratitud. De aquella época recuerdo mi especial relación con el Dr. Isaac Arias que me animó a realizar la tesis doctoral, tarea que comencé armado de un “disponímetro” que requería de ciertas habilidades mecánicas para mantenerlo en funcionamiento; y la confianza, apoyo y asesoramiento que en todo momento tuve del profesor José Luís Vila, sin cuya ayuda difícilmente hubiera podido continuar mi carrera en la Universidad. Tuve también la suerte de contar con dos compañeros excepcionales, los profesores Ramón Martínez-Pacheco y Ángel Concheiro.

Mi traslado a la Universidad de La Laguna supuso emprender una nueva etapa. Pretendí recrear el ambiente que había conocido en Santiago, y allí encontré un grupo de personas, hoy entrañables amigos y compañeros, los profesores Carmen Évora, José Fariña,

Obdulia Munguía, Esther Sánchez, María Isabel Soriano, Araceli Delgado, Alexis Oliva, y Ana Santoveña, cuya colaboración y entusiasmo han sido determinantes en esta segunda etapa de mi vida universitaria.

Aunque sea de forma anónima, también están hoy en mi recuerdo aquellos con los que en un momento u otro trabajé y conté con su colaboración.

Sin duda, el tema de la evaluación de la biodisponibilidad y de la bioequivalencia ha sido una constante a lo largo de mi trayectoria profesional y no es la primera vez que lo elijo para una conferencia. El origen de esta querencia es mi tesis doctoral, cuando tuve que enfrentarme por primera vez al reto de la investigación científica, tarea ardua y las más veces frustrante de transformar observaciones y datos en conocimiento. Apasionante, sí, hasta el punto de hacernos caer en nuestras trampas por introducir de forma inadvertida en nuestros experimentos los ingredientes necesarios para obtener los resultados anhelados. Es cuando nace mi curiosidad por los modelos matemáticos implícitos en el diseño de experimentos y en los métodos de análisis estadísticos, herramientas imprescindibles si deseamos realizar nuestro trabajo de forma objetiva. Pero tampoco nos dejemos engañar por el espejismo de los números; ya nos advirtieron Box y Draper, personalidades reconocidas en la estadística, que *“todos los modelos son falsos, pero algunos son útiles”*¹. Peor aún, podemos acabar utilizando los modelos para ver cumplidas nuestras expectativas o como criterio espurio para eliminar ese dato valioso que por inesperado lo denominamos error.

Es precisamente la necesidad de informar al médico y al paciente de forma objetiva y fiable la que ha conducido la evaluación de la bioequivalencia hacia los terrenos de las estadística biomédica. No se trata de un área en la que se busquen logros científicos o técnicos, que los ha habido, sino de dar respuesta a la legítima preocupación de los ciudadanos por la seguridad y eficacia de los medicamentos.

¹ G. Box y N. Draper. *Evolutionary Operation: A Statistical Method for Process Improvement*. John Wiley & Sons, 1969

Mi intención hoy es analizar la insuficiencia del método clásico y los avances más recientes en el campo de la evaluación de la bioequivalencia y su desarrollo a través de las normas de la European Medicines Agency (EMA) y la Food and Drug Administration (FDA) de los Estados Unidos de América. El trabajo que presento hoy no es sin embargo una simple actualización; el paso del tiempo, si tiene alguna ventaja, es la de facilitarnos una perspectiva más amplia y madura, como individuos y como parte de la comunidad científica. Perspectiva desde la que al mirar atrás siempre provoca la misma pregunta: ¿Cómo no se nos ocurrió antes?

Espero su comprensión y paciencia por el empleo de algunas ecuaciones y fórmulas matemáticas necesarias para expresar de forma coherente y precisa ciertos conceptos, como espero que haya mejorado mi habilidad expositora en los treinta años transcurridos desde que el profesor Fermín Vázquez, al que agradezco una vez más sus palabras de apoyo, me recomendó con su peculiar estilo imperativo: “no lo vuelva a hacer”. Mi pecado había sido aburrir al tribunal que dirimía una plaza de Profesor Asociado de Farmacia Galénica con una lección magistral sobre las funciones de distribución y momentos estadísticos utilizados en el análisis granulométrico. En mi descargo he de decir que me tocó en suerte exponer mi lección magistral a primera hora de la tarde.



1 INTRODUCCIÓN

A pesar de los cuarenta años transcurridos desde las primeras propuestas, el diseño y la evaluación estadística de los ensayos de bioequivalencia no ha dejado de ser origen de controversias. La causa más relevante que alimenta la polémica es sin duda la comercialización de medicamentos genéricos y los conflictos de intereses entre la industria farmacéutica innovadora, fabricantes de genéricos y los sistemas de salud que financian la asistencia médica. Además, esta controversia también ha sido alcanzada por el fenómeno social y económico de globalización. El interés por abrir el mercado internacional a los medicamentos ha llevado a la creación de la *International Conference on Drug Harmonization* (www.ich.org) como cauce para establecer normas comunes de calidad de los medicamentos; sin embargo, hasta la fecha la ICH no ha incluido en su calendario el estudio de una norma común para la evaluación de la bioequivalencia, y las normas de la *Federal Drug Administration* (FDA) y de la *European Agency for Medicinal Products* (EMEA), difieren notablemente. En Europa se mantiene el criterio de *bioequivalencia de promedios* (ABE de sus siglas en inglés, *average bioequivalence*) (1) (2) y que el método estándar de análisis estadístico conocido como doble prueba t-student (3) es suficiente para garantizar la bioequivalencia. Este criterio había sido introducido en los EEUU nueve años antes (4), pero en el año 2001 (5) el FDA adoptó otro basado en los conceptos de *bioequivalencia poblacional* (PBE) y *bioequivalencia individual* (IBE) (6) (5).

Como iremos exponiendo a lo largo de esta conferencia, el criterio de bioequivalencia individual no tiene su origen en la observación de fallos terapéuticos, sino de la opinión de que la bioequivalencia de promedios no es adecuada para limitar el riesgo del paciente, particularmente en el caso de fármacos con un estrecho margen terapéutico. Pero el salto conceptual desde la bioequivalencia de promedios a la bioequivalencia individual no fue inmediato. La bioequivalencia individual se definió por primera vez por Anderson y

Hauck (7) en el año 1990 y no fue hasta el año 2001 cuando se alcanzó el consenso sobre un método de evaluación estadística (5).

Podemos imaginar que a lo largo de todos estos años de discusiones y controversias las opiniones han madurado y muchos criterios desechados; sin embargo hay dos puntos centrales que se han mantenido todo este tiempo:

Primero, *la hipótesis de bioequivalencia se refiere al ensayo de la equivalencia entre productos ensayados, no al ensayo de la diferencia*. Por definición, la equivalencia de dos productos significa que la diferencia entre ellos no es relevante desde el punto de vista terapéutico. Debemos por lo tanto hacer una clara distinción entre lo que entendemos por diferencia terapéuticamente relevante y diferencia estadísticamente significativa; la metodología y criterios estadísticos a utilizar deben orientarse a probar la equivalencia entre los dos productos, no la superioridad de uno de ellos (8).

Segundo, *el riesgo del paciente debe especificarse en el criterio de bioequivalencia*. Se entiende por riesgo del paciente la probabilidad de recibir un producto que sin ser equivalente con el producto de referencia ha sido declarado bioequivalente. Es precisamente la forma en que se especifica el riesgo del paciente donde se sitúan las diferencias entre las bioequivalencias de promedios, poblacional e individual.

En esta conferencia nos centraremos en la evaluación de la bioequivalencia de medicamentos de administración oral; dejaremos al margen problemas íntimamente relacionados como son la bioequivalencia de aerosoles o el de los biogénicos que alargarían la duración de esta intervención más allá de lo que es razonable.

2 BIOEQUIVALENCIA DE PROMEDIOS

El primer documento publicado que reflejaba un consenso básico sobre el procedimiento de evaluación de la bioequivalencia, conocido como el *black book* (9), fue editado en el año 1972 por la American Pharmaceutical Association. Se trata de una corta

monografía donde se exponen los fundamentos de los ensayos, los orígenes de variación esperados y el diseño experimental, y las pruebas estadísticas que se consideraban adecuadas.

Originalmente el ensayo de bioequivalencia estaba orientado a probar lo que hoy día conocemos como *bioequivalencia de promedios* de la formulación evaluada (que denominaremos T) frente a la formulación de referencia (R). Por entonces era común denominar al diseño experimental *cuadrado latino*, denominación que fue desplazada posteriormente por la de *diseño cruzado 2 × 2*. En este diseño participan n individuos, seleccionados al azar entre la población en la que se desea evaluar la bioequivalencia, divididos en dos grupos de tamaño n_1 y n_2 (es deseable pero no imprescindible que $n_1 = n_2$). Los individuos de cada grupo reciben las formulaciones en cada una de las secuencias posibles, RT y TR, separadas por un intervalo de tiempo que asegure que todo el medicamento administrado en la primera ocasión haya sido eliminado. Los parámetros farmacocinéticos utilizados para medir la biodisponibilidad son habitualmente el área bajo la curva de niveles plasmáticos tiempo (ABC) y la concentración plasmática máxima (C_{max}).

Los modelos estadísticos utilizados en la actualidad para evaluar la bioequivalencia se derivan directamente del propuesto por entonces salvo que en la actualidad se considera correcto interpretar el efecto de los individuos como un efecto aleatorio anidado dentro del efecto secuencia. El modelo, denominado logarítmico normal, toma la forma

$$y_{ijk} = \ln(x_{ijk}) = \mu + G_k + P_j + F_{(j,k)} + \xi_{ik} + \varepsilon_{ijk} \quad (1)$$

donde y_{ijk} es el logaritmo natural del parámetro farmacocinético observado en el individuo i ($i = 1 \dots n_k$) perteneciente al grupo que recibe las formulaciones siguiendo la secuencia k durante el período j (en el diseño cruzado 2×2 , $j = 1, 2$ y $k = R, T$).

El modelo incorpora cuatro efectos fijos, es decir, parámetros que se interpretan como constantes a estimar: media general (μ), el efecto fijo “secuencia” (G_k), el efecto debido “período” (P_j), y el

efecto formulación administrada en el período j a los individuos del grupo k ($F_{(j,k)}$, igual a F_R o a F_T). Debido a que estos términos designan categorías es necesario introducir ciertas restricciones siendo habitual las siguientes:

$$\sum_k G_k = 0 \quad \sum_j P_j = 0 \quad F_R + F_T = 0$$

El modelo incluye dos términos aleatorios anidados. ξ_{ik} es el efecto aleatorio debido al individuo i ($i = 1 \dots n_k$) perteneciente al grupo que recibe las formulaciones en la secuencia k . Durante la década de los setenta era habitual asumir que $\xi_{ik} \sim N(0, \sigma_B^2)$ y $\varepsilon_{ijk} \sim N(0, \sigma_W^2)$ y $\text{cov}(\xi, \varepsilon) = 0$.

En los ensayos cruzados 2×2 tenemos 4 promedios correspondientes a otras tantas combinaciones de períodos y secuencias por lo que es posible determinar los 4 parámetros del modelo: la media general, μ , y los tres efectos principales. En la tabla 1 se muestran los coeficientes de las tres combinaciones lineales para la estimación de los efectos formulación, secuencia y período. Obsérvese que el modelo está saturado, es decir, los tres grados de libertad disponibles una vez calculado el promedio general se corresponden con el número de efectos a determinar. Si las secuencias del diseño son RT/TR, la estimada del logaritmo de la biodisponibilidad relativa es por tanto,

$$\bar{y}_T - \bar{y}_R = \frac{1}{2}(\bar{y}_{\cdot 12} + \bar{y}_{\cdot 21}) - \frac{1}{2}(\bar{y}_{\cdot 11} + \bar{y}_{\cdot 22}) \quad (2)$$

y la estimada de la biodisponibilidad relativa:

$$\hat{F} = \exp(\bar{y}_T - \bar{y}_R) \quad (3)$$

Tabla 1. Coeficientes de los promedios por celdas para la estimación de los efectos secuencia, formulación y período, valores esperados y varianzas de las combinaciones lineales resultantes.

efecto	$\bar{y}_{\cdot 11}$	$\bar{y}_{\cdot 12}$	$\bar{y}_{\cdot 21}$	$\bar{y}_{\cdot 22}$	esperanza	varianza
formulación	1/2	-1/2	-1/2	1/2	$F_T - F_R$	$\left(\frac{1}{n_1} + \frac{1}{n_2}\right) \sigma_w^2$
secuencia	-1/2	-1/2	1/2	1/2	$G_2 - G_1$	$\frac{3}{2} \left(\frac{1}{n_1} + \frac{1}{n_2}\right) \sigma_w^2$
período	-1/2	1/2	-1/2	1/2	$P_2 - P_1$	$\left(\frac{1}{n_1} + \frac{1}{n_2}\right) \sigma_w^2$

Criterio de bioequivalencia

El primer criterio de bioequivalencia aceptado por la comunidad científica (9) y por las autoridades regulatorias es la bioequivalencia de promedios; está basado en probar la hipótesis de equivalencia

$$\tilde{\Delta}_L \leq \frac{\tilde{\mu}_T}{\tilde{\mu}_R} \leq \tilde{\Delta}_U$$

o bien, en términos de las variables transformada logarítmicamente (en lo sucesivo utilizaremos los logaritmos naturales)

$$\Delta_L \leq \mu_T - \mu_R \leq \Delta_U$$

Es importante resaltar dos aspectos. Primero, que $\tilde{\mu}_T$ y $\tilde{\mu}_R$ son los valores esperados para los efectos de las formulaciones, es decir, excluyendo la posible influencia de los demás factores del modelo. Segundo, esta hipótesis se basa en el criterio de que la diferencia de los logaritmos de los promedios debe ser suficientemente pequeña como para tener relevancia clínica. Inicialmente se propuso como intervalo de equivalencia [0,80 , 1,20] siendo modificado posteriormente a [0,80 , 1,25] para que fuera simétrico en escala logarítmica de forma que $\Delta_L = -\Delta_U = 0,2331$. También se ha considerado como intervalo de equivalencia para fármacos con elevada variabilidad intraindividual [0,75 , 1,33].

La primera prueba estadística propuesta se conoció como *prueba de la potencia (power rule)*, basada en las siguientes hipótesis nula y alternativa;

$$H_0 : \mu_T - \mu_R = 0$$

$$H_1 = |\mu_T - \mu_R| > \Delta$$

De acuerdo con la prueba de la potencia, la formulación T (ensayada) se declara bioequivalente frente a la formulación R (referencia) si la hipótesis nula se acepta para un nivel de significación $\alpha = 0,05$ y potencia del ensayo $1 - \beta \geq 0,80$. Para ello podemos utilizar la prueba t-student,

$$t = \frac{\bar{y}_T - \bar{y}_R}{s_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad (4)$$

donde s_w es la estimada muestral de la desviación típica intraindividual.

Este método ha sido ampliamente criticado por tres razones. Primero, el éxito de la verificación de la bioequivalencia por este método depende tanto de que μ_T sea próximo a μ_R como del número de individuos utilizados en el ensayo; si el número de individuos es escaso, se corre el riesgo de no alcanzar una potencia suficiente; si el número de individuos es excesivo, el riesgo es rechazar la hipótesis nula es elevado aunque la diferencia de los promedios sea inferior a la máxima diferencia permitida. Segundo, el cálculo de la potencia requiere conocer el parámetro de no centralidad de la distribución t-student no central, función de la varianza de la población, no de la varianza muestral. Como es sabido, la varianza muestral tiene una distribución χ^2 y para los grados de libertad usuales en los ensayos de bioequivalencia es notablemente sesgada por la derecha, por lo que el cálculo de la potencia es poco fiable. Tercero, la prueba de hipótesis nula está orientada a probar la superioridad de un tratamiento sobre otro, no su equivalencia. La importancia de este aserto puede verse comparando los cuadros mostrados en las tablas 2 y 3.

En la tabla 2 se muestra la tabla de decisión cuando se aplica la prueba de hipótesis nula en su verdadero contexto, es decir, cuando deseamos probar que un tratamiento es superior a otro, como es el caso de probar la efectividad de un fármaco comparándolo con un placebo; obsérvese que la eficacia del fármaco se deduce del rechazo de la hipótesis nula y que podemos protegernos de probar falsamente la eficacia del fármaco reduciendo el nivel de significatividad tanto como sea necesario. Es decir, el riesgo del paciente coincide con el nivel de significatividad, que puede fijarse *a priori* y mantenerse bajo control.

Tabla 2. Tabla de decisión en la prueba de efectividad de un fármaco comparándolo con un placebo. La hipótesis ensayada H_0 es de no diferencia entre ambos.

Hipótesis → Resultado de la prueba ↓	H_0 verdadera Fármaco no efectivo	H_0 falsa Fármaco efectivo
Aceptada Fármaco no efectivo	NO ERROR La eficacia del fármaco no ha sido probada	ERROR TIPO II La eficacia del fármaco no ha sido probada; la potencia del experimento es insuficiente y debe aumentarse el número de individuos
Rechazada Fármaco efectivo	ERROR TIPO I La eficacia del fármaco ha sido probada falsamente; debe reducirse el valor de α	NO ERROR La eficacia del fármaco ha sido probada

Si se utiliza la prueba de hipótesis nula para probar la bioequivalencia de dos tratamientos tendremos la situación mostrada en el en la tabla 3. Concluir falsamente lo que deseamos probar es ahora un error tipo II que no podemos controlar directamente. Conclusión, la regla de la potencia no permite controlar directamente el riesgo del paciente y no debe utilizarse para evaluar la bioequivalencia.

A lo largo de la década de los setenta y ochenta se propusieron numerosas alternativas que sin dejar de estar basada en la hipótesis nula paliara los defectos de la prueba, pero no fue hasta el año 1983 cuando Anderson y Hauck (10) pusieron de manifiesto un aspecto que había pasado desapercibido y que sería decisivo en los años siguientes. La prueba de hipótesis nula es adecuada cuando se desea probar la superioridad de un tratamiento frente a otro ya que de esta forma podemos controlar la probabilidad de declarar superior el nuevo tratamiento (rechazo de H_0) cuando ello es falso haciendo tan pequeño que creamos necesario el nivel de significación. En otras palabras, concluimos que lo que deseamos verificar es cierto a partir del rechazo de la hipótesis nula, no de su aceptación. Por lo tanto, si deseamos verificar la bioequivalencia, la hipótesis ensayada deberá ser la de no bioequivalencia. Estos autores proponen como hipótesis nula e hipótesis alternativa las siguientes:

Tabla 3. Tabla de decisión de un ensayo de equivalencia utilizando la prueba de hipótesis nula. La hipótesis nula H_0 asume que la formulación ensayada es equivalente con la de referencia.

Hipótesis → Resultado de la prueba ↓	H_0 verdadera T y R equivalentes	H_0 falsa T y R no equivalentes
Aceptada T y R equivalentes	NO ERROR La equivalencia ha sido aceptada	ERROR TIPO II La equivalencia ha sido aceptada porque la sensibilidad del experimento es insuficiente para detectar una diferencia $\geq \Delta$
Rechazada T y R no equivalentes	ERROR TIPO I La equivalencia ha sido rechazada independientemente de la verdadera diferencia entre tratamientos	NO ERROR La equivalencia ha sido rechazada

Tabla 4. Tabla de decisión para la prueba propuesta por Anderson y Hauck. La hipótesis ensayada es que los tratamiento no son equivalentes.

Hipótesis → Resultado de la prueba ↓	H ₀ verdadera T y R no son bioequivalentes	H ₀ falsa T y R son bioequivalentes
Aceptada T y R no son bioequivalentes	NO ERROR La hipótesis de bioequivalencia no ha sido rechazada	ERROR TIPO II La no bioequivalencia ha sido aceptada debido a la baja potencia del ensayo; debe aumentarse el número de individuos
Rechazada T y R son bioequivalentes	ERROR TIPO I La bioequivalencia ha sido probada falsamente; debe disminuirse α	NO ERROR La hipótesis de no bioequivalencia ha sido rechazada

$$H_0 : \mu_T - \mu_R \leq \Delta_L \quad \text{ó} \quad \mu_T - \mu_R \geq \Delta_U$$

$$H_1 : \Delta_L < \mu_T - \mu_R < \Delta_U$$

El estadístico para ensayar H₀ es:

$$T_{AH} = \frac{\bar{y}_T - \bar{y}_R - (\Delta_L + \Delta_U)/2}{s_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad (5)$$

En la tabla 4 se muestra la tabla de decisión de la prueba de Anderson y Hauck; obsérvese como el riesgo del paciente se protege tomando un valor suficientemente pequeño del nivel de significatividad.

Tal como los autores reconocen, la principal limitación de este método es que el estadístico T tiene una distribución t-student no central y que el parámetro de no centralidad sólo se puede calcular si conocemos la varianza de la población. Este punto determinó que la prueba de Anderson y Hauck no fuera aceptada por la comunidad científica, pero no por ello hemos de dejar de reconocer que centraron el debate sobre la forma en que debía plantearse la hipótesis de equivalencia.

En el año 1987, Shuirmann (11) propuso desdoblar la hipótesis de no bioequivalencia de Anderson y Hauck en la forma siguiente:

$$H_{01} : \mu_T - \mu_R \leq \Delta_L$$

$$H_{11} : \mu_T - \mu_R > \Delta_L$$

y

$$H_{02} : \mu_T - \mu_R \geq \Delta_S$$

$$H_{12} : \mu_T - \mu_R < \Delta_S$$

Para probar H_{01} y H_{02} propone sendas pruebas t-student:

$$t_1 = \frac{(\bar{x}_T - \bar{x}_R) - \Delta_L}{s_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \geq t_{1-\alpha, n_1+n_2-2} \quad (6)$$

$$t_2 = \frac{\Delta_U - (\bar{x}_T - \bar{x}_R)}{s_w \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \geq t_{1-\alpha, n_1+n_2-2} \quad (7)$$

No parece que Shuirmann fuera consciente de que al ser Δ constante las dos pruebas t-student están correlacionadas; de hecho, el argumento estadístico de más peso que aportó es que su método era operacionalmente equivalente con el intervalo de confianza de la diferencia de promedios, es decir, que los resultados numéricos siempre coinciden. En el año 1990 Muller-Cohrs (12) puso de manifiesto que veintidós años antes Owen (13) ya había estudiado un problema similar en el ámbito del control de la calidad, y que la doble prueba t-student sigue la distribución t-student no central bivalente, lo que justificaba desde el punto de vista estadístico la adopción de la prueba de Shuirmann. Esta prueba fue finalmente adoptada en el año 1992 por el FDA (4), y nueve años más tarde por la EMEA (14).

Antes de seguir adelante quiero presentar los resultados de un ensayo de bioequivalencia entre dos formulaciones de granisetron, un antagonista específico de los receptores de la 5-hidroxitriptamina utilizado para prevenir las náuseas y vómitos en pacientes oncológicos. La formulación de referencia es la utilizada en los

ensayos clínicos, y la formulación ensayada una en desarrollo para su comercialización (15). En ensayo se realizó sobre 11 (secuencia 1, RT) más 13 (secuencia 2, TR) individuos. En el cuadro 1 se muestra el resultado del análisis de la varianza (ver modelo en la ecuación 1) obtenido utilizando la aplicación estadística R versión 2.8.1 (16), un programa muy potente, interactivo, totalmente programable y orientado al objeto (la licencia es pública y puede obtenerse en www.R-project.org).

La figura x resume gráficamente los experimentales utilizando el diagrama de caja. La estimada de la biodisponibilidad relativa es $\exp(0,0730) = 1,08$ y los límites de confianza 90%, equivalente a la doble prueba t-student, $[-0.1242, 0,2704]$, equivalente a $[0,883, 1,310]$; obsérvese además que $\hat{\sigma}_B = 1,340^{1/2} = 1,158$, equivalente a un coeficiente de variación interindividual aproximadamente igual 115% y que $\hat{\sigma}_w = 0,1572^{1/2} = 0,397$, un coeficiente de variación intraindividual próximo al 40%. En resumidas cuentas, un fármaco con una elevada variabilidad inter e intraindividual. Más adelante volveremos a analizar estos datos.

Table 5. Análisis de la varianza de un ensayo de bioequivalencia con dos formulaciones de granisetron.

Error: subj						
	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
seq	1	0.7959	0.7959	0.5938	0.4492	
Residuals	22	29.4893	1.3404			
Error: Within						
	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
per	1	0.1794	0.1794	1.1413	0.2970	
form	1	0.0636	0.0636	0.4044	0.5314	
Residuals	22	3.4578	0.1572			

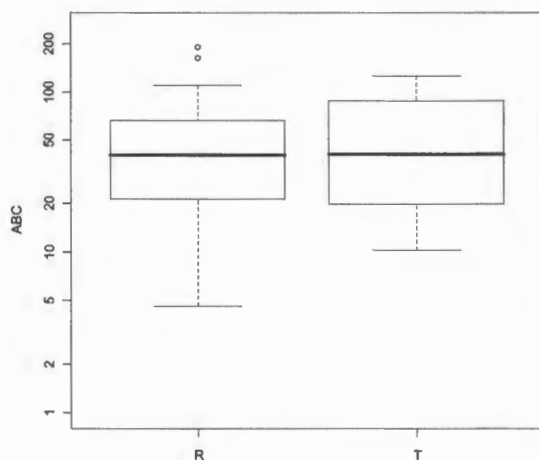


Figura 1. Diagrama de caja para los datos del ensayo de bioequivalencia de dos especialidades de granisetrón.

2.1 Variabilidad intraindividual y limitaciones de la bioequivalencia de promedios

A finales de los años 70 empezó a cuestionarse la suficiencia del criterio de bioequivalencia de promedios. La figura 2 muestra los cocientes individuales de áreas bajo la curva, ABC_T/ABC_R del ejemplo analizado anteriormente; los valores oscilan entre 0,39 y 3,0 y sólo 5 individuos muestran cocientes individuales dentro del intervalo de bioequivalencia, es decir, el riesgo de sub o sobredosificación es elevado. La conclusión es inmediata: para los fármacos con una variabilidad elevada, el nivel de significatividad no es adecuado para medir el riesgo del individuo.

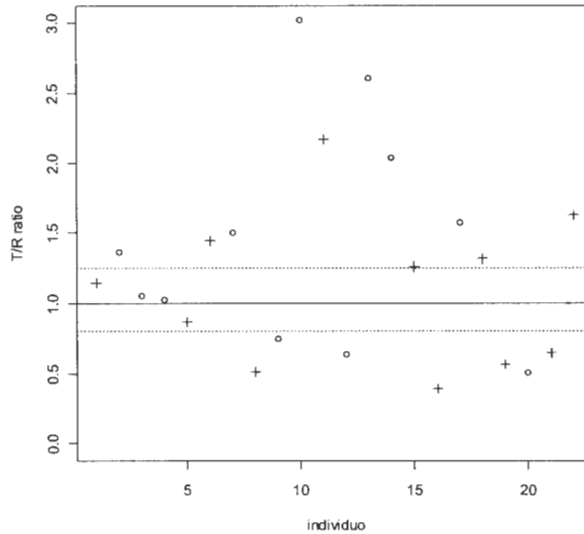


Figura 2. Cocientes individuales de áreas bajo la curva de niveles plasmáticos tiempo para el granisetrón.

Para evitar la declaración de bioequivalentes casos como el del granisetrón, El FDA introdujo una norma adicional a la prueba de bioequivalencia de promedios según la cual el 75% de los individuos que participan en el ensayo deben mostrar una biodisponibilidad relativa individual superior al 75%, requisito conocido como *regla 75/75 (75/75 rule)*. Inicialmente se aplicó a los fármacos antidepresivos (17), ampliándose posteriormente a los inhibidores de la anhidrasa carbónica (18), probenecid (19), fenotiazina (en forma de regla 70/70) (20) y quinidina (21). La regla 75/75 ocupó gran parte del debate sobre los métodos de evaluación. Haynes (22) analizó sus propiedades mediante el método de Monte Carlo concluyendo que eran inadecuadas y proponiendo en su lugar ensayar la prueba de hipótesis de igualdad de las varianzas totales marginales de ambas formulaciones mediante la prueba de Pitman-Morgan (23) (24). FDA continuó defendiendo su postura de mantener la prueba en vigor (25) y los detractores sus críticas (26); la final el FDA se vio obligado a retirarla en el año 1984.

La figura 2 permite presentar de una forma sencilla el problema que presenta la dispersión de las respuestas individuales, pero no permite explorar toda la información contenida en el diseño experimental. La figura 3 muestra el denominado gráfico de interacción individuo \times formulación para ambas secuencias. Hemos de resaltar: la segregación del grupo de la secuencia 2 en dos subgrupos; la elevada variabilidad interindividual, y la elevada interacción individuo \times formulación.

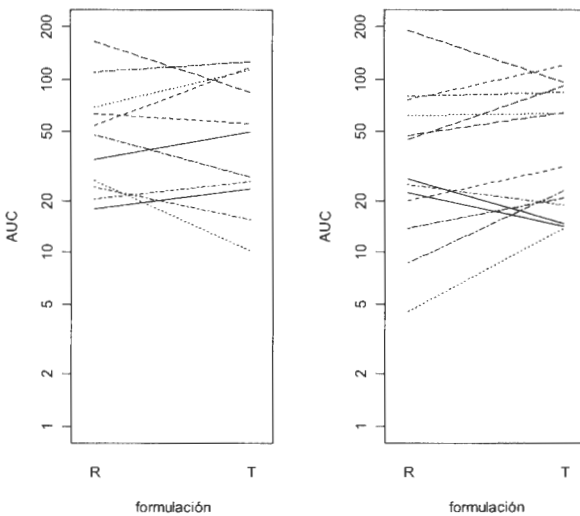


Figura 3.- Gráfico de interacción individuo \times formulación correspondiente al estudio del granistrón.

Las primeras referencias a la interacción individuo \times formulación en la aparecen a finales de los setenta. En una carta dirigida al editor del Journal of Pharmaceutical Sciences en el año 1978, Hwang y colaboradores de los laboratorios Merck Sharp & Dohme (27) manifestaron que la elevada variabilidad de los cocientes individuales puede deberse a la interacción individuo \times formulación. También indicaron que para estimar la varianza debida a esta interacción de forma separada a la variabilidad intraindividual era

necesario utilizar diseños experimentales en los que los individuos recibieran la misma formulación en al menos dos ocasiones diferentes, y que ante la presencia de la interacción individuo \times formulación debemos declarar a la formulación ensayada *no intercambiable* con la formulación de referencia. Es importante resaltar como el criterio de *intercambiabilidad* (*interchangeability*) comienza a ser considerado central en la evaluación de la bioequivalencia. En aquella época la información disponible la interacción individuo \times formulación era escasa y la carta de Hwang y colaboradores tuvo poco eco. Tampoco lo tuvo el estudio sobre el diseño experimental adecuado y evaluación estadística de la interacción individuo \times interacción realizado por el profesor Martínez- Pacheco en el mismo año (28).

En el año 1989, Ekbohm y Melander (29) proponen de nuevo considerar la interacción individuo \times formulación en el modelo estadístico y la prueba F-Snedecor para ensayar la hipótesis $H_0: \sigma_D^2 = 0$ tal como había propuesto Martínez-Pacheco. Estos autores re-analizaron un ensayo de bioequivalencia de dos formulaciones de furosemida publicado previamente por Grahnén y colaboradores (30). El ensayo incluía 8 individuos divididos en dos grupos de 4 que recibieron las formulaciones siguiendo las secuencias RTRT/TRTR. La figura 3 muestra el gráfico de interacción, donde es apreciable la interacción individuo \times formulación.

Tabla 6. Análisis de la varianza para el ensayo de bioequivalencia de dos formulaciones de furosemida.

Error: ind						
	Df	Sum Sq	Mean Sq			
seq	1	0.08572	0.08572			
ind	6	1.19442	0.19907			
Error: Within						
	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
per	3	0.12767	0.04256	0.7030	0.5658	
trt	1	0.00283	0.00283	0.0467	0.8320	
trt:ind	6	0.79432	0.13239	2.1868	0.1069	
Residuals	14	0.84754	0.06054			

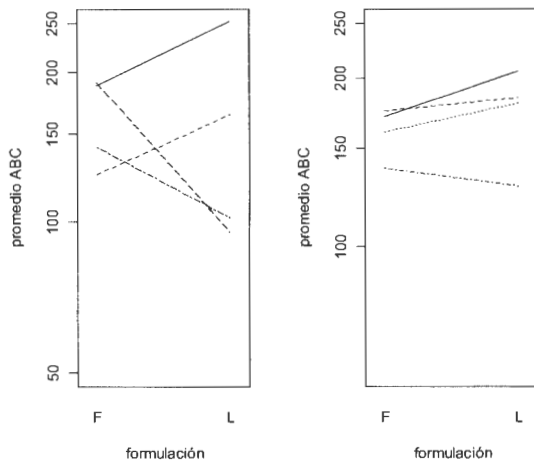


Figura 4. Gráfico de interacción individuo \times formulación para los datos de furosemda.

El trabajo de Ekbohm y Melander tuvo un eco notable y abrió una nueva vía para la evaluación de la bioequivalencia a pesar de que los resultados presentados no eran concluyentes. La tabla 4 muestra el ANOVA, debido destacar los puntos siguientes: primero, hemos de aceptar la hipótesis nula $H_0 : \sigma_D^2 = 0$ ($p = 0,107$) debido a que la potencia del ensayo es baja por el reducido número de individuos utilizados; segundo, la variabilidad de la biodisponibilidad de la furosemda es menor que en el caso anterior; los coeficiente de variación inter e intraindividuales son aproximadamente 45% y 25% respectivamente.

También se han publicado numerosos estudios en los que no está presente la interacción individuo \times formulación. La figura 4 muestra el gráfico de interacción de un estudio de bioequivalencia de 2 formulaciones de verapamilo en el que utilizó un diseño de 4 secuencias y 4 períodos (RTRT/TRTR/RTTR/TRRT) publicado por Chinchilli y Esinhart (31). Obsérvese que los datos muestran una elevada variabilidad interindividual, que ambas formulaciones

presentan una biodisponibilidad similar y que la interacción individuo \times formulación no parece ser relevante.

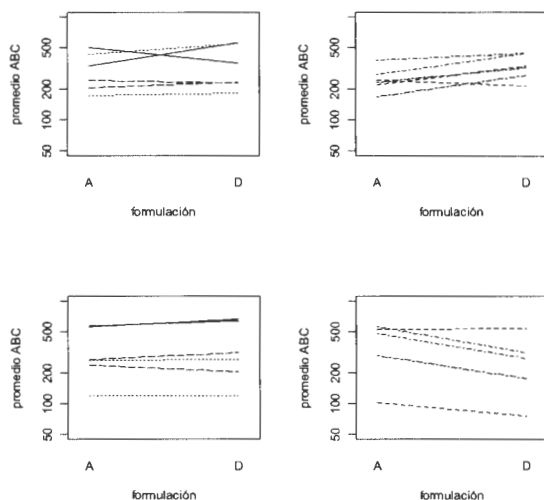


Figura 5. Gráfico de interacción individuo \times formulación del caso 2.

Con la ventaja que siempre da el análisis retrospectivo, podemos hoy concluir que a finales de los años ochenta se había constatado que bioequivalencia de promedios no siempre garantiza la equivalencia para todos los individuos y que el nivel de significatividad no es el riesgo del paciente. Sin embargo, no se había logrado definir la bioequivalencia incorporando el criterio de intercambiabilidad ni establecer una metodología estadística adecuado a tal fin. De hecho, durante la década de los años 90 el FDA mantiene como único criterio de evaluación de la bioequivalencia el método propuesto por Shurman (4).

3 BIOEQUIVALENCIA INDIVIDUAL Y BIOEQUIVALENCIA POBLACIONAL

En el año 1990 Anderson y Hauck introdujeron un nuevo concepto que permitiría replantear la evaluación de los ensayos de bioequivalencia: la bioequivalencia individual y la situación de *intercambiabilidad* (*interchangeability*) o *conmutabilidad* (*switchability*) (7). Posteriormente Sheiner (32) y Schall y Luus (33), siguiendo la línea iniciada por Anderson y Hauck desarrollaron el concepto de *bioequivalencia poblacional* y la situación de *prescribibilidad* (*prescribability*).

La forma más directa de explicar estos conceptos es situándonos en la posición de un médico que tiene que decidir la sustitución de la especialidad que está tomando un paciente por otra especialidad. La preocupación del médico será asegurarse que el cambio de tratamiento no modifique substancialmente los niveles plasmáticos de fármaco y evitar un fallo terapéutico debido a una sobre o sub-dosificación. Si las dos formulaciones que estamos considerando se comportan de forma similar a las de verapamilo podremos confiar en que el cambio de tratamiento no será perjudicial para el paciente; sin embargo, si el comportamiento es similar para las formulaciones de furosemida, lo más prudente sería abstenerse de la sustitución. Otra situación es cuando tenemos que iniciar el tratamiento en un nuevo paciente. La información de que disponemos en ese momento es nuestra experiencia con pacientes anteriores. Si la variabilidad interindividual no es relevante, esperamos que nuestro paciente se comporte como cualquier otro; sin embargo, si la variabilidad interindividual es elevada en relación con el índice terapéutico del fármaco tendremos que tomar precauciones ante la respuesta del paciente y si fuera necesario monitorizar los niveles plasmáticos. Vemos por tanto que para fármacos con índice terapéutico bajo en relación con la variabilidad farmacocinética, la información suministrada por los ensayos de bioequivalencia de promedios no es relevante.

Nuevamente Anderson y Hauck desbrozaron el camino al definir la bioequivalencia individual en los términos siguientes (7): dos formulaciones son bioequivalentes individuales si la biodispo-

nibilidad de la nueva formulación es suficientemente próxima a la de referencia en la mayoría de los individuos. La expresión formal de esta definición es:

$$\text{pr} \left[\tilde{\theta}_L \leq \frac{\mu_{Ti}}{\mu_{Ri}} \leq \tilde{\theta}_U \right] \geq \text{MINP} \quad (8)$$

MINP es la mínima proporción de individuos de la población que deben satisfacer el criterio de bioequivalencia y $[\tilde{\theta}_L, \tilde{\theta}_U]$ el intervalo de bioequivalencia individual. Este criterio recuerda la regla 75/75, pero las diferencias entre ambos métodos son fundamentales. Primero, MINP se refiere a la proporción de la población, no a una proporción de la muestra como en la regla 75/75. Segundo, el cociente μ_{Ti}/μ_{Ri} es el cociente de los valores esperados para los efectos de las formulaciones T y R para un individuo de la población, mientras que el cociente y_{ijT}/y_{ijR} de la regla 75/75 es un valor observado condicionado por los efectos secuencia y período. Anderson y Hauck propusieron un método para evaluar la biodisponibilidad individual basado en la distribución binomial denominado TIER (*test on individual equivalence ratios*) que fue rápidamente desechado por su baja potencia. La ecuación anterior puede aplicarse igualmente a la bioequivalencia poblacional con tan solo referirnos a dos individuos distintos.

A pesar del fracaso del método TIER, el impacto del artículo de Anderson y Hauck fue importante porque recondujo la discusión sobre la importancia de la variabilidad en la evaluación de la bioequivalencia, y dio origen a diversos métodos para la evaluación de la bioequivalencia individual y poblacional. Se han propuesto tres grupos de métodos: métodos basados en los intervalos de tolerancia, propuestos inicialmente por Esinhart y Chinchilli (34) utilizando el modelo estadístico clásico del análisis de la varianza para 2, 3 y 4 períodos, y por Brown y colaboradores (35) utilizando el modelo de efectos mixtos; métodos basados en la probabilidad de similitud de las funciones de distribución estadística, propuesto inicialmente por Sheiner (32), y métodos basados en los momentos propuesto por Schall y Luus (33). En el año 1997 el FDA recomendó el empleo del

método de los momentos escalado (36) descrito por Schall y Williams (37). En el año 2001 el FDA adopta definitivamente una variante de este método (5). Esta variante, propuesta por Hyslop, Husuan y Holder (38), permite simplificar notablemente los cálculos implicados; los componentes de la varianza se estiman utilizando contrastes intraindividuales (31) en lugar de utilizar el método iterativo de máxima verisimilitud restringida, y el límite de confianza superior del criterio de bioequivalencia se calcula mediante el método descrito entre otros por Ting y colaboradores (39) en sustitución de la técnica bootstrap.

Modelo de efectos mixtos

Los métodos de evaluación de la bioequivalencia individual, y en particular el método utilizado por el FDA, requieren la estimación de siete parámetros: los promedios poblacionales para ambas formulaciones, las varianzas intra e inter individuales y la varianza de la interacción individuo \times formulación, $\{\mu_R, \mu_T, \sigma_{BR}^2, \sigma_{BT}^2, \sigma_{WR}^2, \sigma_{WT}^2, \sigma_D^2\}$. Algunos autores han sugerido utilizar las medias cuadráticas esperadas en el análisis de la varianza (40), pero las estimadas resultantes no son independientes y no es descartable estimadas negativas de la varianza interindividual y de la de interacción individuo \times formulación. La mayoría de los autores se decantan por utilizar el modelo de efectos mixtos; en lo sucesivo describiremos el propuesto por Chinchilli y Esinhart (31).

El modelo parte de un ensayo cruzado $s \times p$ (s , número de secuencias; p , número de períodos); en ausencia de efecto residual, toma la forma:

$$Y_{ijkl} = \mu_k + \gamma_{ikl} + \delta_{ijk} + \epsilon_{ijkl} \quad (9)$$

donde $i = 1, \dots, s$ indica la secuencia; $j = 1, \dots, n_i$ el individuo anidado dentro de las secuencias; $l = 1 \dots p_{ik}$ el replicado de la formulación k en la secuencia i ; $k = R, T$ la formulación administrada. y_{ijkl} es el logaritmo de la observación (ABC, Cmax) realizada en el individuo j de la secuencia i en el replicado l del tratamiento k . γ_{ikl} es el efecto fijo debido al replicado l del tratamiento k en la secuencia i ; δ_{ijk} es el

efecto aleatorio debido al individuo j de la secuencia i cuando recibe la formulación k ; ϵ_{ijkl} es el efecto del individuo asociado a la observación. En el modelo se asume además que los términos ϵ_{ijkl} son mutuamente independientes con distribución

$$\epsilon_{ijkl} \sim N(0, \sigma_{wk}^2) \quad (10)$$

Los efectos aleatorios del individuo j de la secuencia i viene dado por el vector

$$\boldsymbol{\mu}_{ij} = \begin{pmatrix} \mu_{ijR} \\ \mu_{ijT} \end{pmatrix} = \begin{pmatrix} \mu_R + \delta_{ijR} \\ \mu_T + \delta_{ijT} \end{pmatrix} \quad (11)$$

La distribución de los valores de μ_{ij} se asumen igualmente independientes y con una distribución normal bivalente:

$$\boldsymbol{\mu}_{ij} \sim N \left(\begin{pmatrix} \mu_R \\ \mu_T \end{pmatrix}, \begin{pmatrix} \sigma_{BR}^2 & \rho \sigma_{BR} \sigma_{BT} \\ \rho \sigma_{BR} \sigma_T & \sigma_{BT}^2 \end{pmatrix} \right) \quad (12)$$

Para evitar la sobre-parametrización del modelo se aplican las siguientes restricciones:

$$\sum_{i=1}^s \sum_{k=1}^{p_k} \gamma_{ikl} = 0 \quad (13)$$

La interacción individuo \times formulación mide la dispersión interindividual de la diferencia entre las respuestas a ambas formulaciones:

$$\sigma_D^2 = \text{var} [\mu_{T,i} - \mu_{R,i}] = \sigma_{BT}^2 + \sigma_{BR}^2 - 2 \rho \sigma_{BT} \sigma_{BR} \quad (14)$$

o bien

$$\sigma_D^2 = (\sigma_{BT} - \sigma_{BR})^2 + 2(1-\rho)\sigma_{BT} \sigma_{BR} \quad (15)$$

Consecuentemente, la interacción individuo \times formulación tiene de acuerdo con este modelo dos orígenes: una diferencia entre las varianzas inter-individuales de las formulaciones T y R, y que las observaciones dentro de un individuo no estén perfectamente correlacionadas.

3.1 Bioequivalencia individual

Sean Y_T e Y_R las variables aleatorias que describen el valor esperado cuando un individuo recibe la formulación T y R respectivamente. De acuerdo con el modelo anterior el valor esperado del cuadrado de esta diferencia es igual:

$$E[(Y_T - Y_R)^2] = (\mu_T - \mu_R)^2 + \sigma_D^2 + \sigma_{WT}^2 + \sigma_{WR}^2 \quad (16)$$

Por otro lado, el valor esperado cuando el individuo recibe la formulación de referencia en dos ocasiones diferentes es:

$$E[(Y_R - Y'_R)^2] = 2\sigma_{WR}^2 \quad (17)$$

y la diferencia de esperanzas:

$$M_U = E[(Y_T - T_R)^2] - E[(Y_R - Y'_R)^2] = (\mu_T - \mu_R)^2 + \sigma_D^2 + \sigma_{WT}^2 - \sigma_{WR}^2 \quad (18)$$

Esta diferencia fue el primer criterio de bioequivalencia individual propuesto por Schall y Luus (33). Es interesante resaltar la equivalencia de este criterio con el de bioequivalencia cuando $\sigma_{WR}^2 = \sigma_{WT}^2$ y $\sigma_D^2 = 0$, razón por la cual se propuso como criterio de bioequivalencia individual:

$$M_U < 0,2331^2 \quad (19)$$

Schall y Williams (37) pusieron de manifiesto que el empleo de la diferencia de esperanzas como criterio de bioequivalencia supone un criterio excesivamente estricto cuando la formulación de referencia presenta una variabilidad intraindividual baja, por lo que estos autores proponen el escalado del criterio utilizando una estrategia mixta:

$$\eta_I = \frac{(\mu_T - \mu_R)^2 + \sigma_D^2 + \sigma_{WT}^2 - \sigma_{WR}^2}{\sigma_0^2} < \theta_I \quad (20A)$$

$$\sigma_0^2 = \begin{cases} 0,20 & \text{si } \sigma_{WR}^2 < 0,20 \\ \sigma_{WR}^2 & \text{si } \sigma_{WR}^2 \geq 0,20 \end{cases}$$

Este estadístico no tiene una distribución conocida por lo que es necesario calcular su intervalo de confianza mediante la técnica

bootstrap. Hyslop, Hsuan y Holder (38) propusieron el criterio linealizado que figura en el documento del FDA del año 2001:

$$\eta_I = (\mu_T - \mu_R)^2 + \sigma_D^2 + (\sigma_{WT}^2 - \sigma_{WR}^2) - \theta_I \sigma_0^2 < 0 \quad (21)$$

3.2 Bioequivalencia poblacional

El desarrollo del criterio de bioequivalencia poblacional ha sido paralelo al de bioequivalencia poblacional. La diferencia es que ahora Y_T e Y_R son las variables aleatorias que describen el valor esperado cuando un individuo recibe la formulación T y otro individuo la formulación R. El criterio de bioequivalencia linealizado es:

$$\eta_P = (\mu_T - \mu_R)^2 + \sigma_{TT}^2 - \sigma_{TR}^2 - \theta_P \sigma_0^2 < 0 \quad (22)$$

3.3 Criterios de bioequivalencia

Los criterios de bioequivalencia poblacional (ecuación 22) e individual (ecuación 21) introducen un nuevo punto de discusión: los valores de los criterios θ_I y θ_P . Para la bioequivalencia individual,

$$\theta_I = \frac{\ln(1,25) + \varepsilon_I}{\sigma_0^2} \quad (23)$$

Obsérvese que el criterio tiene la misma forma que el criterio η_I ; el valor 1,25 se corresponde con los límites impuestos en la biodisponibilidad de promedios y el coeficiente ε_I con los límites impuestos a los componentes de la varianza; σ_0^2 mantiene su significado. El criterio general recomendado por el FDA se basa en el hecho de que $\sigma_D \approx 0,1356$, aproximadamente el 10% de los individuos mostraran bioequivalencias individuales fuera del intervalo 0,80 – 1,25 aunque $\mu_T - \mu_R = 0$, por lo que el criterio recomendado es:

$$\theta_I = \frac{\ln(1,25)^2 + 0,05}{0,20^2} = 2,4948 \quad (24)$$

No existe sin embargo un acuerdo general para el valor de θ_P ; de hecho, el documento del FDA señala que aquellos que deseen

realizar un ensayo de bioequivalencia poblacional deben ponerse en contacto con la agencia.

Analizados los datos de las formulaciones de furosemida encontramos:

$$\hat{\sigma}_D^2 = 0,1032 \quad \hat{\sigma}_{WR}^2 = 0,04559 \quad \hat{\sigma}_{WT}^2 = 0,04381$$

Obsérvese que las varianzas intraindividuales son similares, pero que la varianza de la interacción individuo \times formulación, 0,1032, es notablemente alta. El parámetro de bioequivalencia individual (ecuación 21) es $\hat{\eta}_1 = 0,0687 > 0$, por lo que no podemos rechazar la hipótesis de no equivalencia.

Por contrario, las dos formulaciones de verapamilo son bioequivalentes individuales; las varianzas estimadas fueron:

$$\sigma_D^2 = 3,85 \cdot 10^{-4} \quad \hat{\sigma}_{WR}^2 = 0,0592 \quad \sigma_{WT}^2 = 0,0636$$

y el parámetro de bioequivalencia $\hat{\eta}_1 = -0,0505 < 0$, por lo que hemos de rechazar la hipótesis de no bioequivalencia individual, quedando por lo tanto probada la equivalencia. Obsérvese que en este segundo caso las varianzas intraindividuales estimadas son similares a las observadas para la furosemida, pero que la varianza estimada para la interacción es mucho menor.

4 RELEVANCIA TERAPÉUTICA DE LA BIOEQUIVALENCIA INDIVIDUAL

Todo lo expuesto hasta ahora se ha centrado en el riesgo del paciente analizado desde una perspectiva exclusivamente estadística, y las pruebas de que la bioinequivalencia individual sea un problema real que hemos presentado se han limitado al estudio de Ekbhom y Melander.

Ciertamente, la evidencia experimental de que la bioinequivalencia individual sea un problema terapéutico es escasa por varias razones. Desde el punto de vista de la investigación académica, los ensayos de bioequivalencia tienen un coste elevado y

escaso interés científico salvo algún hallazgo casual inesperado. Por otro lado, no se conocen los mecanismos responsables de la interacción individuo \times formulación ni disponemos de un modelo animal que nos permitiera elucidar el origen de la interacción. Podemos suponer que un efecto de primer paso elevado, el polimorfismo metabólico y procesos de absorción y metabolismo saturables juegan un papel relevante, pero no disponemos de la información necesaria para mantener esta opinión.

Por otro lado, los datos que obtiene la industria, aunque conocidos por las agencias reguladoras, son de carácter reservado. No obstante. El FDA hizo pública en el año 1997 una base de datos con los resultados de 80 ensayos de bioequivalencia (www.fda.gov/cder). Diez ensayos se realizaron con administración replicada de la formulación de referencia o la ensayada y fueron analizados por Chen y colaboradores (41); los nombres de los fármacos no son públicos, sólo su categoría terapéutica. En 3 de los 10 ensayos (se corresponde con el 40% de los fármacos ensayados) se observó una interacción individuo \times formulación elevada, es decir $\sigma_D > 0,15$.

Más complicado es analizar la relevancia terapéutica de la bioequivalencia individual. Sólo es de esperar la observación de fallos terapéuticos en fármacos con estrecho margen terapéutico y paciente sometidos a un estricto control médico, pero los datos publicados son escasos y el origen del fallo terapéutico no suele estar debidamente documentado ni excluidos otros motivos como pueden ser las interacciones medicamentosas, olvido de dosis, etc. Pasqualotto y Denning (42) recogen tres casos de fallos terapéuticos por la sustitución de una marca registrada de itraconazol (Sponanox™) por un genérico. Van Paesschem y colaboradores (43) advierten del riesgo de sustitución de antiepilépticos en niños, pero no aportan evidencias de fallos terapéuticos. Sí los aportan Nielsen y colaboradores (44) al relacionar fracasos terapéuticos con los niveles plasmáticos de lomotrigina. En sentido contrario tenemos el trabajo de Lee y colaboradores (45) quienes observan la bioequivalencia y equivalencia terapéutica de un genérico respecto a un producto de marca de warfarina sódica.

Lo cierto es que estamos ante un problema que ha generado desconfianza frente a la sustitución de un medicamento por otro bioequivalente. Reiffel y Kowey (46) publicaron una encuesta realizada entre cardiólogos y concluyen que aquellos consideran que los procedimientos para evaluar la bioequivalencia en los EEUU no son adecuados; la mayoría de los médicos consultados se decantan contra la sustitución. Haskins y colaboradores (47) estudiaron la opinión que merecía a médicos y pacientes la sustitución de anticonvulsivantes mediante una encuesta realizada a 974 pacientes y 435 neurólogos de 5 países, Canadá, Reino Unido, Francia, Alemania y España. La opinión mayoritaria es de rechazo a la sustitución por el temor a la aparición de convulsiones.

A día de hoy la discusión sobre la relevancia de la interacción individuo \times formulación en el riesgo del paciente está abierta y tendremos que esperar a obtener información adicional para tomar una decisión definitiva. En ocasiones se ha argumentado que el criterio de bioequivalencia individual es una barrera para impedir la entrada de genéricos en el mercado; mi opinión sin embargo es la contraria; probar la bioequivalencia individual es sin duda la mejor carta de presentación de un genérico en el cada vez más competitivo mercado farmacéutico.

Muchas gracias por la atención que me han prestado.

Bibliografía

1. *Note for guidance on the investigation of bioavailability and bioequivalence. The European Agency for the Evaluation of Medicinal products.* London : s.n., 1981. CPMP/EWP/QWP/1401/98.
2. *Guideline on the Investigation of Bioequivalence (draft). European Medicines Agency.* Londres : s.n., 2008. CPMP/EWP/QWP/1401/98 rev. 1.
3. *A comparison of the two one sided test procedure and the power approach for assessing the equivalence of average bioavailability.* **Schirmann, D. J.** 1987, Journal of Pharmacokinetics and Biopharmaceutics, Vol. 15, págs. 657 - 680.
4. **FDA.** *Guidance on Statistical Procedures for Bioequivalence Studies Using a Standard Two-Treatment Crossover Design.* Rockville, Maryland : Center for Drug Evaluation and Research, Food and Drug Administration, 1992.
5. —. *Guidance for Industry on Statistical Approaches to Establishing Bioequivalence.* Rockville, Maryland : Center for Drug Evaluation and Research, Food and Drug Administration, 2001.
6. —. *Guidance for Industry: Bioavailability and Bioequivalence Studies for Orally Administered Drug Products - General Considerations.* Rockville, Maryland : Center for Drug Evaluation and Research, Food and Drug Administration, 2000.
7. *Consideration of individual bioequivalence.* **Anderson, S. y Hauck, W. W.** 1990, Journal of Pharmacokinetics and Biopharmaceutics, Vol. 18, págs. 259 - 773.
8. *Bioavailability - A problem in equivalence.* **Metzler, C. M.** 1974, Biometrics, Vol. 30, págs. 309 - 317.
9. **American Pharmaceutical Association.** *Guidelines form Biopharmaceutical Studies in Man.* Washington D. C. : Academy of Pharmaceutical Sciences, 1972.
10. *A new procedure for testing equivalence in comparative bioavailability and other clinical trials.* **Anderson, S. y Hauck, W.**

-
- W. 1983, Communications in Statistics - Theoretical Methods, Vol. 12, págs. 2663 - 2692.
11. *A comparison of the two one-sided test procedure and the power approach for assessing the equivalence of average bioavailability.* **Schirmann, D. J.** 1987, Journal of Pharmacokinetics and Biopharmaceutics, Vol. 15, págs. 657 - 680.
12. *The power of the Anderson-Hauck test and the double t-test.* **Müllers-Cohrs, J.** 1990, Biometrical Journal, Vol. 32, págs. 259 - 266.
13. *A special case of a bivariate non-central t-distribution.* **Owen, D. B.** 1965, Biometrika, Vol. 52, págs. 437 - 446.
14. **Committee for Proprietary Medicinal Products.** *Note for Guidance on the Investigation of Bioavailability and Bioequivalence.* The European Agency for the Evaluation of Medicinal Products. London : s.n., 2001. CPMP/EWP/QWP/1401/98.
15. *Evaluation of bioequivalence of tablet and capsule formulacion of granisetron in patients undergoing cytotoxic chemotherapy for malignant disease.* **Cupissol, D., y otros.** 1993, Journal of Pharmaceutical Sciences, Vol. 82, págs. 1281 - 1284.
16. **R Development Core Team.** *R: A language and enviroment for statistical computing.* Vienna : R Foundation for Statistical Computing, 2008.
17. **FDA.** 1978, Federal Register, Vol. 43, pág. 6968.
18. **FDA.** s.l. : Federal Register, 1980, Vol. 45, pág. 11853.
19. **FDA.** 1980, Federal Register, Vol. 45, pág. 48160.
20. **FDA.** 1980, Federal Register, Vol. 45, pág. 56832.
21. **FDA.** 1980, Federal Register, Vol. 45, pág. 72200.
22. *Statistical simulation study of new proposed uniformity requirement for bioequivalency studies.* **Haynes, J. D.** 1981, Journal of Pharmaceutical Sciences, Vol. 70, págs. 673 - 675.

23. *A note on normal correlation.* **Pitman, E. J. G.** 1939, *Biometrika*, Vol. 31, págs. 9 - 12.
24. *A test for the significance of the difference between the two variances in a sample from a bivariate population.* **Morgan, W. A.** 1939, *Biometrika*, Vol. 31, págs. 13 - 19.
25. *Assessment of 75/75 rule: FDA viewpoint.* **Cabana, B. E.** 1983, *Journal of Pharmaceutical Sciences*, Vol. 72, págs. 98 - 99.
26. *FDA 75/75 rule: a response.* **Haynes, J. D.** 1983, *Journal of Pharmaceutical Sciences*, Vol. 72, págs. 99 - 100.
27. *Bioequivalence and interchangeability.* **Hwang, S., y otros.** 1978, *Journal of Pharmaceutical Sciences*, Vol. 67, pág. IV.
28. **Martçínez-Pacheco, R.** *Diseño y Evaluación Biofarmacéutica de Comprimidos de Cesión Sostenida.* Santiago de Compostela : s.n., 1978.
29. *The subject-by-formulation interaction as a criterion of interchangeability of drugs.* **Ekbohm, G. y Melander, H.** 1989, *Biometrics*, Vol. 45, págs. 1249 - 1254.
30. *Implications of intraindividual variability in bioavailability studies of furosemide.* **Grahnén, A., Hammrlund, M. y Lundqvist, T.** 1984, *European Journal of Clinical Pharmacology*, Vol. 27, págs. 595 - 602.
31. *Design and analysis of intra-subject variability in cross-over experiments.* **Chinchilli, V. y Esinhart, J.** 1996, *Statistics in Medicine*, Vol. 15, págs. 1619 - 1634.
32. *Bioequivalence revisited.* **Sheiner, L. B.** 1992, *Statistics in Medicine*, Vol. 11, págs. 1777 - 1788.
33. *On population and individual bioequivalence.* **Schall, R. y Luus, H. G.** 1993, *Statistics in Medicine*, Vol. 12, págs. 1109 - 1124.
34. *Extension to the use of tolerance intervals for the assessment of individual bioequivalence.* **Esinhart, J. D. y Chinchilli, V. M.** 1994, *Journal of Biopharmaceutical Statistics*, Vol. 4, págs. 39 - 52.

-
35. *Tolerance intervals for assessing individual bioequivalence.* **Brown, E. B., Iyer, H. K. y Wang, C. M.** 1997, *Statistics in Medicine*, Vol. 16, págs. 803 - 820.
36. **FDA.** *Guidance for Industry: In vivo equivalence studies based on population and individual bioequivalence approaches.* Center for Drug Evaluation and Research (CDER), Food and Drug Administration. US Department of health and Human Services. 1997.
37. *Towards a practical strategy for assessing individual bioequivalence.* **Schall, R. y Williams, R. L.** 1996, *Journal of Pharmacokinetics and Biopharmaceutics*, Vol. 24, págs. 133 - 149.
38. *A small sample confidence interval approach to assess individual bioequivalence.* **Hyslop, T., Hsuan, F. y Holder, D. J.** 2000, *Statistics in Medicine*, Vol. 19, págs. 2885 - 2897.
39. *Confidence intervals on linear combinations of variance components taht are unrestricted in sign.* **Ting, N., y otros.** 1990, *Journal of Statistical Computing and Simulation*, Vol. 35, págs. 135 - 143.
40. *Definition of individual bioequivalence: occasion-to-occasion versus mean switchability.* **Kimanani, E. K.** 2000, *Statistics in Medicine*, Vol. 19, págs. 2729 - 2810.
41. *An individual bioequivalence criterion: regulatory considerations.* **Chen, M.-L., y otros.** 2000, *Statistics in Medicine*, Vol. 19, págs. 2821 - 2842.
42. *Generic substitution of itraconazol resulting in sub-therapeutic levels and resistance.* **Pasqualotto, A. C. y Denning, D. W.** 2007, *International Jouranal of Antimicrobial Agents*, Vol. 30, págs. 93 - 94.
43. *The use of generic medication in epilepsy: a review of potential issues and challenges.* **Van Paesschen, W., Hauman, H. y Lagae, L.** 2009, *Official Journal of the european Paediatric Neurology society*, Vol. 13, págs. 87 - 92.

44. *Comparative daily profiles with different preparations of lamotrigine: a pilot investigation.* **Nielsen, K. A., y otros.** 2008, *Epilepsy & Behavior*, Vol. 13, págs. 127 - 130.
45. *Efficacy and tolerability of the swithc from branded to generic warfarin sodium product; an observer-blinded, randomized cross-over study.* **Lee, H.-L., Kan, C.-D. y Yang, Y.-J.** 2005, *Clinical Therapeutics*, Vol. 27, págs. 309 - 315.
46. *Generic antiarrhythmics are not therapeutically equivalent for the treatment of tachyarrhythmias.* **Reiffel, J. A. y Kowey, P. R. K.** 2000, *The American Journal of Cardiology*, Vol. 85, págs. 1151 - 1153.
47. *Patient and physician reactions to generics antiepileptic substitution in the treatment of epilepsy.* **Haskins, L. S., Tomaszewski, K. J. y Crawford, P.** 2005, *Epilepsy & Behaviour*, Vol. 7, págs. 98 - 105.
48. *A new statistical procedure for testing equivalence in two-groups comparative bioavailability trials.* **Hauck, W. W. y Anderson, S.** 1984, *Journal of Pharmacokinetics and Biopharmaceutics*, Vol. 12, págs. 83 - 91.
