



BIOEQUIVALENCIA PARA DISEÑOS REPLICADOS

María Teresa CASPARRI, María José BIANCO

Centro de Investigación en Métodos Cuantitativos Aplicados a la Economía y la Gestión (CMA), Instituto de Investigaciones en Administración, Contabilidad y Métodos Cuantitativos para la Gestión (IADCOM), Universidad de Buenos Aires, Av. Córdoba 2122 - 1120AAQ - Ciudad Autónoma de Buenos Aires, República Argentina.

mariajose.bianco@economicas.uba.ar

Resumen

Recibido: 12/2016

Aceptado: 03/2017

Palabras clave

Bioequivalencia en

promedio.

Bioequivalencia individual.

Diseños replicados.

La bioequivalencia es un estudio que consiste en comparar una formulación Test (nuevo medicamento) y una formulación Referencia (versión original) utilizando un diseño crossover 2x2. A cada sujeto se le administra una dosis de T y R y se les toman muestras de sangre con el fin de determinar la concentración de droga en el plasma. Si los promedios están “suficientemente cerca” las dos formulaciones se consideran *bioequivalentes en promedio*.

El diseño estándar de crossover de dos períodos dos secuencias no es útil en presencia de efecto carryover. Además, no provee de estimadores independientes para las varianzas dentro de los sujetos. Para salvar estas desventajas, en la práctica, se considera el uso de diseños de orden mayor a dos.

Si bien dos drogas pueden ser similares en media sobre la población de potenciales pacientes, podrían producir efectos diferentes cuando un paciente cambia de la formulación T a la R o viceversa. En otras palabras, habría una interacción significativa entre sujeto-formulación. Para demostrar que esto no sucede habría que probar que ambas formulaciones cumplen la *bioequivalencia individual*.

Los diseños de crossover replicados son necesarios para aproximar IBE ya que permiten la estimación de la varianza en el sujeto de T y R y la componente de la varianza de la interacción sujeto formulación.

En el presente trabajo se presentan distintos métodos para el cálculo de bioequivalencia en promedio y bioequivalencia individual utilizando diseños de crossover replicados basándose en los propuestos en las guías draft del 2001 de la FDA.

Copyright: Facultad de Ciencias Económicas, Universidad de Buenos Aires.

ISSN: 2250-687X - ISSN (En línea): 2250-6861

REPLICATE DESIGNS FOR BIOEQUIVALENCE STUDIES

María Teresa CASPARRI, María José BIANCO

Centro de Investigación en Métodos Cuantitativos Aplicados a la Economía y la Gestión (CMA), Instituto de Investigaciones en Administración, Contabilidad y Métodos Cuantitativos para la Gestión (LADCOM), Universidad de Buenos Aires, Av. Córdoba 2122 - 1120AAQ - Ciudad Autónoma de Buenos Aires, República Argentina.

mariajose.bianco@economicas.uba.ar

Abstract

KEYWORDS

Average bioequivalence.
Individual bioequivalence.
Replicate designs.

The bioequivalence summary is a study which is to compare a formulation (new drug) Test and a reference (original version) formulation using a 2 x 2 crossover design. Each subject was given a dose of T and R and blood samples are taken to determine the concentration of drug in the plasma. If the averages are "close enough" two formulations are Bioequivalent on average (ABE).

Standard two period crossover design two sequences is not useful in the presence of carryover effect. In addition, not provides of estimators independent for the variances within the subject. To save these disadvantages, in practice, is considered the use of order greater than two designs. Although two drugs can be similar in half on the population of potential patients, could produce effects different when a patient changes of the formulation T to the R or vice versa. In other words, would have a meaningful interaction between subject-formulation. To demonstrate that this does not happen, it would have to prove that both formulations to meet individual bioequivalence.

Them designs of crossover replicated are necessary for approach IBE since allow the estimation of the variance in the subject of T and R and the component of the variance of the interaction subject formulation. This paper presents various methods for the calculation of bioequivalence on average and Individual bioequivalence using replicated crossover designs based on the proposed draft of the 2001de guidelines the FDA.

Copyright: Facultad de Ciencias Económicas, Universidad de Buenos Aires.

ISSN: 2250-687X - ISSN (En línea): 2250-6861

INTRODUCCIÓN

El desarrollo de un medicamento se divide en dos etapas. La etapa preclínica, en la que se identifica una droga o molécula relacionada con cierta patología, se desarrolla y se verifica su interacción con el sitio de interés donde se busca que actúe. La etapa clínica se divide en fases: Fase I (seguridad), Fase II (eficacia), Fase III (eficacia comparativa) y Fase IV (farmacovigilancia).

En Argentina, si bien se realizan estudios para las cuatro fases de análisis clínicos, la mayoría de los recursos de investigación y desarrollo se destinan a la Fase III, que se focaliza en ensayos sobre pacientes sobre la seguridad y eficacia de los medicamentos¹.

En esta etapa la regulación actual acepta para la evaluación de un medicamento el criterio de bioequivalencia en promedio (ABE)² el cual sólo prueba la bioequivalencia en medias. Pero esa métrica no controla las varianzas de ambas formulaciones y no tiene en cuenta una posible interacción sujeto–formulación. Por lo tanto, la FDA (Food and Drug Administration) propone el uso de diseños replicados.

En la primera sección de este trabajo se realizará una descripción de los modelos de diseños replicados para, en la segunda sección, aplicar esta metodología en la evaluación de bioequivalencia en promedio. En la tercera sección se presentarán los procedimientos para el cálculo de bioequivalencia individual (IBE)³, los cuales se aplicarán a datos de la bibliografía.

Finalmente se realizará una comparación de ambos métodos y se concluirá en término de los resultados obtenidos.

1. ¿QUÉ ES UN DISEÑO REPLICADO?

Un diseño de crossover de orden mayor a dos se define como un diseño donde el número de períodos es mayor a dos o bien el número de secuencias es mayor que el número de tratamientos que se comparan.

El diseño estándar de crossover de dos períodos dos secuencias no es útil en presencia de efecto carryover. Además, no provee de estimadores independientes para las varianzas dentro de los sujetos (*within subject variance*). Para salvar estas desventajas, en la práctica, se considera el uso de diseños de orden mayor a dos.

¹ Para profundizar sobre la evolución de la industria farmacéutica y la producción de medicamentos en nuestro país ver Bianco (2015).

² ABE: *average bioequivalence*

³ IBE: *individual bioequivalence*

Un rasgo de este tipo de diseño es que no es necesario asumir que los efectos del sujeto son variables aleatorias con el fin de testear la diferencia entre los efectos carryover. Aunque al principio se puede cubrir la información directa del tratamiento desde la variación entre sujetos, es sumamente difícil probar que vale la pena, debido a las diferencias entre sujetos típicas de las pruebas crossover.

Consecuentemente, con el fin de comparar diseños se asume que los efectos de los sujetos son fijos y que los errores dentro de los sujetos (*within subject errors*) son independientes con media cero y varianza σ^2 .

Otro punto importante es la forma en que los efectos pueden ser estimados mediante los contrastes dentro de los sujetos. Con el fin de hacer esto primero se debe determinar cuántos grados de libertad están disponibles y luego determinar los contrastes que están asociados a esos grados de libertad.

Si un diseño tiene s secuencias y p períodos, entonces hay $s \cdot p$ medias de períodos. Por lo tanto, hay $(s \cdot p - 1)$ grados de libertad entre estas medias, los cuales pueden dividirse en $(s - 1)$ entre grupos, $(p - 1)$ entre períodos y $(s - 1) \cdot (p - 1)$ para los efectos de la interacción grupo por período. Este último grupo contiene los grados de libertad asociados a los efectos de mayor interés. Aunque este conjunto de grados de libertad puede ser particionado de distintas formas, siempre tendremos en cuenta estos tres conjuntos básicos: (a) el efecto directo de los tratamientos, (b) la interacción directa del período y los efectos carryover, (c) otros efectos que no son de interés directo asociados con la interacción grupos y períodos.

En algunos diseños los efectos asociados con los grados de libertad en (b) pueden ser divididos en otros efectos, y por lo tanto se deberá elegir qué términos deberán incluirse en el modelo.

En resumen, se deben identificar los grados de libertad que están asociados a los efectos de interés y luego formular un modelo que particione estos grados de libertad.

El diseño de crossover replicado puede ser usado independientemente de la aproximación seleccionada para establecer bioequivalencia. Estos diseños son necesarios para aproximar IBE ya que permiten la estimación de la varianza en el sujeto de T y R y la componente de la varianza de la interacción sujeto formulación.

En el Cuadro 1 se presenta el diseño de cuatro períodos dos secuencias dos formulaciones recomendado para estudios de bioequivalencia replicados.

Cuadro 1

		Período			
		1	2	3	4
Secuencia	1	T	R	T	R
	2	R	T	R	T

Fuente: FDA Guidance (2001).

Para este diseño el mismo conjunto de T y R debe ser usado para la administración replicada. Cada período debe ser separado por un adecuado período de lavado.

Los efectos fijos que usualmente se incluyen en un análisis estadístico son secuencia, período y tratamiento (formulación). El número de grados de libertad atribuible a cada efecto fijo es generalmente igual al número de niveles del efecto menos uno. Por eso, en el caso del diseño de dos secuencias cuatro períodos habrá un grado de libertad para la secuencia (2-1), tres grados de libertad para el período (4-1) y un grado de libertad para el tratamiento (2-1) dando un total de cinco grados de libertad para el total de los efectos fijos (1+3+1). Estos 5 grados de libertad difieren de los 7 grados de libertad que tiene el modelo, por lo tanto se dice que es *no saturado*. Para lograr un modelo *saturado* se podría agregar a los tres efectos fijos principales (secuencia, período y tratamiento), por ejemplo, el efecto de la interacción secuencia-tratamiento y el efecto de la interacción secuencia-período.

Si el diseño de crossover replicado sólo tiene dos secuencias, el uso de los tres efectos principales en el modelo de efectos fijos o el uso del modelo saturado, prácticamente no tienen diferencia.

En el Cuadro 2 se presenta el modelo de cuatro períodos dos formulaciones cuatro secuencias.

Cuadro 2

		Período			
		1	2	3	4
Secuencia	1	T	R	R	T
	2	R	T	T	R
	3	T	T	R	R
	4	R	R	T	T

Fuente: FDA Guidance (2001).

Si el diseño de crossover replicado contiene más de dos secuencias, el modelo de efectos fijos generalmente produce una estimación diferente de $\mu_T - \mu_R$ con respecto al modelo saturado, a menos que el número de sujetos en cada secuencia sea el mismo. Además, la estimación de la varianza por el método de los momentos es eficiente sólo para el modelo saturado mientras que para el modelo de efectos fijos el estimador más eficiente debería incluir en el modelo alguna componente entre secuencias, complicando el modelo.

Por eso el uso de diseños con sólo dos secuencias minimiza o evita ciertas ambigüedades con respecto al método de estimación de la varianza elegido o al modelo de efectos fijos que se considere.

La única razón para usar un modelo replicado de cuatro secuencias cuatro períodos es que se piense que es el óptimo si se incluye el efecto carryover.

Se considera el siguiente modelo estadístico, el cual corresponde a un diseño de 4 períodos con igual número de replicaciones de T y R en cada secuencia s . Se asume que los efectos carryover están ausentes (la respuesta a una formulación administrada en un período particular del diseño no está afectada por la formulación administrada en períodos anteriores) o bien son iguales para cada formulación (se consideran dentro del efecto del período).

$$Y_{ijkl} = \mu_k + \gamma_{ikl} + \delta_{ijk} + \varepsilon_{ijkl}$$

$$i = 1, \dots, s \quad \text{secuencia}$$

$$j = 1, \dots, n_i \quad \text{sujeto de la secuencia } i$$

$$k = T, R \quad \text{formulación}$$

$$l = 1, 2 \quad \text{replicación de la formulación } k \text{ en la secuencia } i$$

Y_{ijkl} = respuesta en la secuencia i del sujeto j con la formulación k en la replicación l .

μ_k = media de la formulación k

γ_{ikl} = efecto fijo en la secuencia i de la formulación k en la replicación l (*efecto fijo de la formulación*)

δ_{ijk} = efecto aleatorio en la secuencia i del sujeto j con la formulación k (*random subject effect*)

ε_{ijkl} = error aleatorio en la secuencia i del sujeto j con la formulación k en la replicación l (*within subject random error*)

δ_{ijk} y ε_{ijkl} son mutuamente independientes.

ε_{ijkl} son independientes e idénticamente distribuidas

$$\varepsilon_{ijkl} \sim N(0, \sigma_{wk}^2)$$

σ_{wk}^2 = varianza dentro del sujeto para la formulación k (*within subject variance*)

$m_{ij} = \begin{pmatrix} \mu_R + \delta_{ijR} \\ \mu_T + \delta_{ijT} \end{pmatrix}$ son independientes y tienen distribución

$$m_{ij} \sim N \left[\begin{pmatrix} \mu_R \\ \mu_T \end{pmatrix}; \begin{pmatrix} \sigma_{BR}^2 & \rho \sigma_{BT} \sigma_{BR} \\ \rho \sigma_{BT} \sigma_{BR} & \sigma_{BT}^2 \end{pmatrix} \right]$$

Es decir,

$$Var(\delta_{ijR}) = \sigma_{BR}^2 \quad Var(\delta_{ijT}) = \sigma_{BT}^2 \quad Cov(\delta_{ijR}, \delta_{ijT}) = \rho \sigma_{BT} \sigma_{BR} = \sigma_{BTR} = w_{RT}$$

σ_{Bk}^2 = varianza entre sujetos para la formulación k (*between subject variance*)

Bajo este modelo, $Var(\delta_{ijT} - \delta_{ijR}) = \sigma_D^2$ donde σ_D^2 es la varianza interacción sujeto formulación (*subject by formulation interaction variance*)

Para evitar la sobreparametrización se aplica la siguiente restricción:

$$\sum_{i=1}^s \sum_{l=1}^2 \gamma_{ikl} = 0$$

Este modelo se puede resolver utilizando el Modelo Lineal Generalizado (correspondiente a la aproximación por el método de los momentos), máxima verosimilitud o máxima verosimilitud restringida.

2. BIOEQUIVALENCIA EN PROMEDIO EN DISEÑOS REPLICADOS

2.1 Procedimientos inferenciales para ABE en diseños replicados

La regla de decisión para determinar bioequivalencia en promedio para diseños replicados es igual a la de un diseño crossover de dos períodos.

Sean μ_T y μ_R las medias poblacionales de concentración plasmática máxima (C_{max}) o bien del área bajo la curva de concentración plasmática (AUC) cuando los sujetos son tratados con T y R respectivamente.

El clásico intervalo de $(1 - 2\alpha) \times 100\%$ de confianza para $\mu_T - \mu_R$ se obtiene de la siguiente forma

$$L_1 = \hat{\mu}_T - \hat{\mu}_R - t_{\alpha; gl} \sqrt{\hat{Var}(\hat{\mu}_T - \hat{\mu}_R)} \quad (1)$$

$U_1 = \hat{\mu}_T - \hat{\mu}_R + t_{\alpha; gl} \sqrt{\hat{Var}(\hat{\mu}_T - \hat{\mu}_R)}$ donde gl son los grados de libertad de una variable aleatoria con distribución t de Student (observemos que si $T \sim t_{gl}$ entonces $P(T > t_{\alpha; gl}) = \alpha$).

Este intervalo puede ser convertido en un intervalo de $(1 - 2\alpha) \times 100\%$ de confianza para $\frac{\mu_T}{\mu_R}$

dividiendo por $\hat{\mu}_R$ (considerando a $\hat{\mu}_R$ como la verdadera media de la referencia). Por lo tanto:

$$L_2 = \left(\frac{L_1}{\hat{\mu}_R} + 1 \right) \times 100\% \quad (2)$$

$$U_2 = \left(\frac{U_1}{\hat{\mu}_R} + 1 \right) \times 100\%$$

Luego, se concluye que existe bioequivalencia en promedio si se satisface ($\theta_L = -0,2\mu_R$; $\theta_U = 0,2\mu_R$; $\delta_L = 80\%$; $\delta_U = 120\%$):

$$(L_1, U_1) \subset (\theta_L, \theta_U) \quad \text{o bien} \quad (L_2, U_2) \subset (\delta_L, \delta_U) \quad (3)$$

Schuirman (1981, 1987) introdujo el uso de test de hipótesis para establecer bioequivalencia, siendo las hipótesis a testear:

$$H_0 : \mu_T - \mu_R \leq \theta_L \quad \text{o} \quad \mu_T - \mu_R \geq \theta_U \quad H_1 : \theta_L < \mu_T - \mu_R < \theta_U$$

donde $\theta_L = -0,2\mu_R$ y $\theta_U = 0,2\mu_R$

Este test puede ser descompuesto en dos tests de una cola (TOST: *two one – sided tests*):

$$H_{0_1} : \mu_T - \mu_R \leq \theta_L \quad H_{1_1} : \theta_L < \mu_T - \mu_R \quad (4)$$

$$H_{0_2} : \mu_T - \mu_R \geq \theta_U \quad H_{1_2} : \theta_U > \mu_T - \mu_R$$

Bajo la hipótesis de normalidad de las variables, se rechazan las hipótesis nulas si:

$$T_L = \frac{\hat{\mu}_T - \hat{\mu}_R - \theta_L}{\sqrt{\hat{Var}(\hat{\mu}_T - \hat{\mu}_R)}} > t_{\alpha; gl} \quad \text{y} \quad T_U = \frac{\hat{\mu}_T - \hat{\mu}_R - \theta_U}{\sqrt{\hat{Var}(\hat{\mu}_T - \hat{\mu}_R)}} < -t_{\alpha; gl} \quad (5)$$

Se establece bioequivalencia en promedio sí y sólo si se rechazan ambas hipótesis nulas con un nivel de significación α .

Si los datos están transformados, es decir, si μ_T y μ_R son las medias poblacionales del $\ln(AUC)$ (o bien, $\ln(C \max)$) cuando los sujetos son tratados con T y R respectivamente, ABE se demuestra si el intervalo del $(1 - 2\alpha) \times 100\%$ de confianza para $\mu_T - \mu_R$ dado en (1) se encuentra dentro de los límites de aceptación de $-\ln(1,25) = -0,2231$ y $\ln(1,25) = 0,2231$ (FDA Guidance, 1992, 2001, 2003), conocida como la regla del $\pm 20\%$.

Al aplicar la función exponencial, es decir, en la escala natural, ABE se demuestra si hay suficiente evidencia para asegurar:

$$0,80 \leq \exp(\mu_T - \mu_R) \leq 1,25 \quad (6)$$

En efecto, si se consideran a $\tilde{\mu}_T$ y $\tilde{\mu}_R$ como las medias poblacionales de los datos sin transformar, se obtiene:

$$-\ln(1,25) < \mu_T - \mu_R < \ln(1,25)$$

$$-\ln(1,25) < \ln \tilde{\mu}_T - \ln \tilde{\mu}_R < \ln(1,25)$$

$$-\ln(1,25) < \ln \left(\frac{\tilde{\mu}_T}{\tilde{\mu}_R} \right) < \ln(1,25)$$

$$\exp[-\ln(1,25)] < \exp \left[\ln \left(\frac{\tilde{\mu}_T}{\tilde{\mu}_R} \right) \right] < \exp[\ln(1,25)]$$

$$0,80 \leq \frac{\tilde{\mu}_T}{\tilde{\mu}_R} \leq 1,25 \quad (7)$$

Se observa que la simetría del intervalo de confianza es en la escala logarítmica y no en la escala natural.

Cuando se utilizan datos transformados por logaritmo, las hipótesis a testear para TOST son:

$$\begin{aligned} H_{0_1} : \mu_T - \mu_R &\leq -\ln 1,25 & H_{1_1} : -\ln 1,25 < \mu_T - \mu_R \\ H_{0_2} : \mu_T - \mu_R &\geq \ln 1,25 & H_{1_2} : \ln 1,25 > \mu_T - \mu_R \end{aligned} \quad (8)$$

2. 2 Métodos de estimación de ABE para diseños replicados

En esta sección se presentarán distintos procedimientos de estimación. Primero se desarrollará el método que Chow y Liu proponen en su libro *Design and Analysis of Bioavailability and Bioequivalence Studies* (1992) en el que se asume que la varianza sujeto–formulación es nula. Luego, se presentará el *método de los momentos* el cual provee de estimadores insesgados e independientes de la diferencia de medias ($\delta = \mu_T - \mu_R$) y de la varianza ($\sigma_I^2 = \sigma_D^2 + \frac{\sigma_{WT}^2 + \sigma_{WR}^2}{2}$). Finalmente, se discutirá el método de *aproximación asintótica* basado en los estimadores de máxima verosimilitud restringida (REML)⁴. Este método permite encontrar estimadores de cada uno de los componentes de la varianza.

A continuación se enuncian las notaciones y definiciones que se usarán en las siguientes secciones para el cálculo de los estimadores y sus varianzas.

Recordar que s representa la cantidad de secuencias, n_i el número de sujetos en la secuencia i y

$$n - s = \left(\sum_{i=1}^s n_i \right) - s$$

$$Y_{ijT\bullet} = \frac{1}{2} (Y_{ijT1} + Y_{ijT2})$$

$$Y_{ijR\bullet} = \frac{1}{2} (Y_{ijR1} + Y_{ijR2})$$

$$\bar{Y}_{i\bullet k\bullet} = \frac{1}{n_i} \sum_{j=1}^{n_i} \left[\frac{1}{2} \sum_{l=1}^2 Y_{ijkl} \right] = \frac{1}{n_i} \sum_{j=1}^{n_i} Y_{ijk\bullet}$$

$$I_{ij} = Y_{ijT\bullet} - Y_{ijR\bullet}$$

$$\bar{I}_{i\bullet} = \frac{1}{n_i} \sum_{j=1}^{n_i} I_{ij}$$

$$T_{ij} = Y_{ijT1} - Y_{ijT2}$$

$$\bar{T}_{i\bullet} = \frac{1}{n_i} \sum_{j=1}^{n_i} T_{ij}$$

⁴ REML: *Restricted maximum likelihood*

$$R_{ij} = Y_{ijR1} - Y_{ijR2} \qquad \overline{R_{i\bullet}} = \frac{1}{n_i} \sum_{j=1}^{n_i} R_{ij}$$

$$\hat{\mu}_k = \frac{1}{s} \sum_{i=1}^s \overline{Y_{i\bullet k}} \qquad k = T, R \qquad (9)$$

$$\hat{\delta} = \hat{\mu}_T - \hat{\mu}_R = \frac{1}{s} \sum_{i=1}^s \left[\frac{1}{n_i} \sum_{j=1}^{n_i} I_{ij} \right] \qquad (10)$$

$$M_I = \hat{\sigma}_I^2 = \frac{1}{n-s} \sum_{i=1}^s \sum_{j=1}^{n_i} (I_{ij} - \overline{I_{i\bullet}})^2 \qquad (11)$$

$$M_T = \hat{\sigma}_{WT}^2 = \frac{1}{2(n-s)} \sum_{i=1}^s \sum_{j=1}^{n_i} (T_{ij} - \overline{T_{i\bullet}})^2 \qquad (12)$$

$$M_R = \hat{\sigma}_{WR}^2 = \frac{1}{2(n-s)} \sum_{i=1}^s \sum_{j=1}^{n_i} (R_{ij} - \overline{R_{i\bullet}})^2 \qquad (13)$$

Se demuestra que los estimadores $\hat{\delta}$, M_I , M_T y M_R de $\delta = \mu_T - \mu_R$, $\sigma_I^2 = \sigma_D^2 + \frac{\sigma_{WT}^2 + \sigma_{WR}^2}{2}$, σ_{WT}^2 y σ_{WR}^2 , respectivamente, son insesgados e independientes

(Patterson and Jones, 2002). En efecto, $\hat{\delta}$ y M_I son independientes basado en resultados de Fisher (1925) y Student (1908). Por otro lado, como M_T y M_R provienen de observaciones normales multivariadas independientes, bajo el modelo propuesto también son independientes. Además (Chinchilli and Esinhart, 1996):

$$\hat{\delta} = \hat{\mu}_T - \hat{\mu}_R \sim N \left(\delta; \sqrt{\frac{\sigma_I^2}{s^2} \sum_{i=1}^s \frac{1}{n_i}} \right)$$

$$M_I = \hat{\sigma}_I^2 \sim \frac{\sigma_I^2 \chi_{n-s}^2}{n-s}$$

$$M_T = \hat{\sigma}_{WT}^2 \sim \frac{\sigma_{WT}^2 \chi_{n-s}^2}{n-s}$$

$$M_R = \hat{\sigma}_{WR}^2 \sim \frac{\sigma_{WR}^2 \chi_{n-s}^2}{n-s}$$

Consecuentemente $\hat{\delta}$ es un estimador insesgado de $\delta = \mu_T - \mu_R$ (Vonesh and Chinchilli, 1997). Si definimos $\sigma^2 = \sigma_{WT}^2 + \sigma_{WR}^2$ y $r^2 = \frac{1}{s^2} \sum_{i=1}^s \frac{1}{n_i}$, entonces:

$$E(\hat{\delta}) = \mu_T - \mu_R$$

$$\text{Var}(\hat{\delta}) = r^2 \sigma_i^2 = r^2 \left(\sigma_D^2 + \frac{\sigma^2}{2} \right) \quad (14)$$

Si el diseño es de dos secuencias y cuatro períodos entonces $r^2 = \frac{1}{4} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)$ y si el diseño es

de cuatro secuencias y cuatro períodos entonces $r^2 = \frac{1}{16} \left(\frac{1}{n_1} + \frac{1}{n_2} + \frac{1}{n_3} + \frac{1}{n_4} \right)$

Se observa que replicar las mediciones en cada sujeto reduce el error en un factor igual al número de repeticiones. Por ejemplo, en un diseño de crossover de dos períodos la varianza de la media de cada una de las formulaciones es $\sigma_{Bi}^2 + \sigma_{Wi}^2$ ($i = T, R$) donde σ_{Bi}^2 es la varianza entre sujetos y σ_{Wi}^2 es la varianza dentro del sujeto (error). En un diseño replicado, la varianza de la media de cada una de las formulaciones es $\sigma_{Bi}^2 + \frac{\sigma_{Wi}^2}{2}$. Por esa razón, cuando se está frente a una gran variabilidad dentro de los sujetos, el diseño replicado provee una estimación más precisa de la verdadera respuesta individual. Para un producto de baja variabilidad, replicar no influye en la precisión; pero si el producto es altamente variable, un diseño con repeticiones restringe el rango donde la respuesta media individual varía.

i) Método de Chow y Liu

Shein Chung Chow y Jen Pei Liu (1992) proponen métodos para testear ABE en diseños replicados considerando que la varianza sujeto–formulación es nula ($\sigma_D^2 = 0$).

Si un sujeto no recibió alguna de las dosis R o T en alguno de los períodos (*incomplete data*) se excluye del estudio.

Se considerará el modelo planteado en la sección anterior donde los efectos carryover están ausentes.

Si se asume igualdad de varianzas ($\sigma_{WT}^2 = \sigma_{WR}^2 = \sigma_e^2$), se considera $\hat{\sigma}_e^2$ como el error cuadrático medio al realizar el análisis de la varianza del modelo planteado.

Para obtener el estimador de la varianza, en (14) se reemplaza σ^2 por $\hat{\sigma}^2 = 2\hat{\sigma}_e^2$. Este

estimador tiene $3 \left(\sum_{i=1}^s n_i \right) - 4$ grados de libertad.

Si el diseño es de dos secuencias y cuatro períodos,

$$\hat{Var}(\hat{\delta}) = r^2 \hat{\sigma}^2 = \frac{1}{4} \left(\frac{1}{n_1} + \frac{1}{n_2} \right) \frac{2\hat{\sigma}_e^2}{2} = \frac{1}{4} \left(\frac{1}{n_1} + \frac{1}{n_2} \right) \hat{\sigma}_e^2 \quad (15)$$

con $3(n_1 + n_2) - 4$ grados de libertad asociados

Si el diseño es de cuatro secuencias y cuatro períodos,

$$\hat{Var}(\hat{\delta}) = r^2 \hat{\sigma}^2 = \frac{1}{16} \left(\frac{1}{n_1} + \frac{1}{n_2} + \frac{1}{n_3} + \frac{1}{n_4} \right) \frac{2\hat{\sigma}_e^2}{2} = \frac{1}{16} \left(\frac{1}{n_1} + \frac{1}{n_2} + \frac{1}{n_3} + \frac{1}{n_4} \right) \hat{\sigma}_e^2 \quad (16)$$

con $3(n_1 + n_2 + n_3 + n_4) - 4$ grados de libertad asociados.

Reemplazando (10) y (15) (o (16) de acuerdo al diseño que se utilice) en (1), se concluye que los límites del intervalo de $(1 - 2\alpha) \times 100\%$ de confianza para $\mu_T - \mu_R$ están dados por:

$$L_1 = \hat{\mu}_T - \hat{\mu}_R - t_{\alpha; gl} \sqrt{r^2 \hat{\sigma}_e^2} \quad (17)$$

$$U_1 = \hat{\mu}_T - \hat{\mu}_R + t_{\alpha; gl} \sqrt{r^2 \hat{\sigma}_e^2}$$

Si la hipótesis de igualdad de varianzas se rechaza, entonces para obtener un estimador de la varianza, la expresión de σ^2 en (14) debe ser reemplazada (Chow y Liu, 1992; Wang y Hwang, 2000) por la suma de M_T y M_R , definidos en (10) y (13) respectivamente, que son los estimadores de σ_{WT}^2 y σ_{WR}^2 :

$$\hat{\sigma}^2 = M_T + M_R = \frac{1}{2(n-s)} \sum_{i=1}^s \sum_{j=1}^{n_i} \left[(T_{ij} - \bar{T}_{i\cdot})^2 + (R_{ij} - \bar{R}_{i\cdot})^2 \right] \quad (18)$$

En este caso los grados de libertad asociados son $2 \left[\left(\sum_{i=1}^s n_i \right) - s \right]$

Reemplazando (10) y (18) en (1), se concluye que los límites del intervalo de $(1 - 2\alpha) \times 100\%$ de confianza para $\mu_T - \mu_R$ están dados por:

$$L_1 = \hat{\mu}_T - \hat{\mu}_R - t_{\alpha; gl} \sqrt{r^2 \left(\frac{M_T + M_R}{2} \right)} \quad (19)$$

$$U_1 = \hat{\mu}_T - \hat{\mu}_R + t_{\alpha; gl} \sqrt{r^2 \left(\frac{M_T + M_R}{2} \right)}$$

ii) Método de los momentos

Este método permite obtener estimadores insesgados de las componentes del modelo: diferencia de medias y la suma de varianzas de la varianza de la diferencia de medias. Además, tiene en cuenta la varianza de la interacción sujeto formulación. (Patterson and Jones, 2002b).

Si un sujeto no recibió alguna de las dosis R o T en alguno de los períodos (*incomplete data*) no se incluye en el estudio.

Se puede demostrar (Vonesh and Chinchilli, Chapter 4, 1997):

$$\text{Var}(\hat{\mu}_T - \hat{\mu}_R) = r^2 \left[\sigma_{BT}^2 + \sigma_{BR}^2 - 2\rho\sigma_{BT}\sigma_{BR} + \frac{\sigma_{WT}^2 + \sigma_{WR}^2}{2} \right] = r^2 \left[\sigma_D^2 + \frac{\sigma_{WT}^2 + \sigma_{WR}^2}{2} \right]$$

Por lo tanto, para obtener un estimador de la varianza, se reemplaza en la expresión anterior

$$\sigma_I^2 = \sigma_D^2 + \frac{\sigma_{WT}^2 + \sigma_{WR}^2}{2} \text{ por su estimador } M_I, \text{ definido en (11).}$$

Se concluye que los límites del intervalo de $(1 - 2\alpha) \times 100\%$ de confianza para $\mu_T - \mu_R$ están

$$\text{dados por: } L_1 = \hat{\mu}_T - \hat{\mu}_R - t_{\alpha; n_I} \sqrt{r^2 M_I} \quad (20)$$

$$U_1 = \hat{\mu}_T - \hat{\mu}_R + t_{\alpha; n_I} \sqrt{r^2 M_I}$$

iii) Aproximación asintótica

Este método permite calcular una aproximación asintótica de la media y la varianza. Se presentan los resultados basados en la teoría de una aproximación normal utilizando máxima verosimilitud restringida (Patterson and Jones, 2002b; Jones and Kenwards, 2003).

Las estimaciones por el método de máxima verosimilitud restringida pueden ser calculadas por diversos paquetes de software. Si el conjunto de datos es balanceado y completo, los estimadores de las componentes de la varianza y el estimador de los efectos fijos son insesgados.

El mayor interés en la práctica de este método es obtener estimaciones en conjuntos de datos incompletos. Los estimadores obtenidos al utilizar REML en conjuntos de datos incompletos, son independientes y asintóticamente insesgados para los parámetros de interés (Milliken and Johnson, 1992).

Se obtienen los siguientes estimadores: $\hat{\sigma}_{BT}^2$, $\hat{\sigma}_{BR}^2$, $\hat{\sigma}_{WT}^2$, $\hat{\sigma}_{WR}^2$ y $\hat{\omega}_{RT}$ los cuales están normalmente distribuidos en el límite con matriz de varianza-covarianza apropiada al modelo planteado.

Se concluye que los límites del intervalo de $(1 - 2\alpha) \times 100\%$ de confianza asintótico para $\mu_T - \mu_R$ están dados por:

$$U_1 = \hat{\mu}_T - \hat{\mu}_R + z_\alpha \sqrt{\hat{Var}(\hat{\mu}_T - \hat{\mu}_R)}$$

$$L_1 = \hat{\mu}_T - \hat{\mu}_R - z_\alpha \sqrt{\hat{Var}(\hat{\mu}_T - \hat{\mu}_R)}$$

Se debe notar que la estimación de la varianza no está bien caracterizada en ejemplos de pocos individuos. En muchas situaciones de bioequivalencia en promedio, sin embargo, el tamaño de la muestra es menor a 30, con lo cual el procedimiento asintótico se vuelve cuestionable. En Patterson and Jones (2002) sugieren (basándose en Kenward and Jones (1997)) que en situaciones donde el tamaño de la muestra sea pequeño, se utilice el mismo intervalo de confianza pero con una aproximación con *t* de Student.

2.3 Aplicación a datos experimentales

Ejemplo 1:

En este ejemplo se comparan dos tabletas genéricas de furosemide: Lasix (test) y Furix (referencia) con un diseño de dos secuencias cuatro períodos (TRTR / RTRT).

El intervalo del 90% de confianza que se considera en Cuadro 3 corresponde a $\frac{\mu_T}{\mu_R}$ en la escala natural.

Cuadro 3

Método	$\hat{\delta} = \hat{\mu}_T - \hat{\mu}_R$	$Var(\hat{\delta})$	Grados de libertad	IC 90%
Chow – Liu $\sigma_{WT}^2 = \sigma_{WR}^2$	-0,019	0,1025	20	$(0,82; 1,17) \subset (0,80; 1,25)$
Chow – Liu $\sigma_{WT}^2 \neq \sigma_{WR}^2$	-0,019	0,0072	12	$(0,84; 1,14) \subset (0,80; 1,25)$
Método de los Momentos	-0,019	0,0165	6	$(0,76; 1,26) \not\subset (0,80; 1,25)$

Fuente: Elaboración propia basado en datos de la literatura (transformados mediante logaritmo natural) correspondientes a Ekbohm, G. y Melander, H. (1989) “The subject by formulation interaction as a criterion of interchangeability of drugs”. *Biometrics*, 45, 1249 – 1254.

El intervalo del 90% de confianza para $\frac{\mu_T}{\mu_R}$ en la escala natural utilizando aproximación asintótica es $(0,76; 1,26)$. Como no está contenido dentro de los límites 0,8 y 1,25 no se puede considerar que T sea bioequivalente en promedio a R.

Ejemplo 2:

El estudio compara las tabletas (test) y las cápsulas (referencia) de una prodroga de olsalazine con un diseño de dos secuencias y cuatro períodos (TRRT / RTTR). Se presenta un dato faltante en el sujeto 5.

El intervalo del 90% de confianza que se considera en el Cuadro 4 corresponde a $\frac{\mu_T}{\mu_R}$ en la escala natural.

Cuadro 4

Método	$\hat{\delta} = \hat{\mu}_T - \hat{\mu}_R$	$Var(\hat{\delta})$	Grados de libertad	IC 90%
Chow – Liu $\sigma_{WT}^2 = \sigma_{WR}^2$	0,169	0,0106	23	(0,99; 1,41) $\not\subset$ (0,80; 1,25)
Chow – Liu $\sigma_{WT}^2 \neq \sigma_{WR}^2$	0,169	0,0137	14	(0,96; 1,45) $\not\subset$ (0,80; 1,25)
Método de los Momentos	0,169	0,0042	7	(1,05; 1,34) $\not\subset$ (0,80; 1,25)

Fuente: Elaboración propia basada en datos de la literatura (transformados mediante logaritmo natural) correspondientes a Ryde, M.; Huitfeldt, B.; Pettersson, R. (1991) Relative bioavailability of Olsalazine from tables and capsules: a drug targeted for local effect in the colon. *Biopharm & Drug Disposition*, 12, 233-246.

El intervalo del 90% de confianza para $\frac{\mu_T}{\mu_R}$ en la escala natural utilizando aproximación asintótica sin incluir el sujeto 5 es (1,05; 1,34). Como no está contenido dentro de los límites 0,8 y 1,25 no se puede considerar que T sea bioequivalente en promedio a R.

El intervalo del 90% de confianza para $\frac{\mu_T}{\mu_R}$ en la escala natural utilizando aproximación asintótica incluyendo en el estudio al sujeto 5 es (1,04; 1,32). Como no está contenido dentro de los límites 0,8 y 1,25 no se puede considerar que T sea bioequivalente en promedio a R.

Ejemplo 3:

El estudio compara una formulación test con la formulación de referencia de Verapamil con un diseño de cuatro secuencias y cuatro períodos (TRTR / RTRT / TRRT / RTTR).

El intervalo del 90% de confianza que se considera en el Cuadro 5 corresponde a $\frac{\mu_T}{\mu_R}$ en la escala natural.

Cuadro 5

Método	$\hat{\delta} = \hat{\mu}_T - \hat{\mu}_R$	$Var(\hat{\delta})$	Grados de libertad	IC 90%
Chow – Liu $\sigma_{WT}^2 = \sigma_{WR}^2$	-0,02	0,0043	65	(0,88; 1,09) \subset (0,80; 1,25)
Chow – Liu $\sigma_{WT}^2 \neq \sigma_{WR}^2$	-0,02	0,0044	38	(0,88; 1,1) \subset (0,80; 1,25)
Método de los Momentos	-0,02	0,0026	19	(0,90; 1,07) \subset (0,80; 1,25)

Fuente: Elaboración propia basado en datos de la literatura (transformados mediante logaritmo natural) correspondientes a Esinhart, J.D.; Chinchilli, V.M. (1994) Extension to the use of tolerance intervals for the assessment of individual bioequivalence. *Journal of Biopharmaceutical Statistics*, 4, 39-52.

El intervalo del 90% de confianza para $\frac{\mu_T}{\mu_R}$ en la escala natural utilizando aproximación asintótica es (0,90; 1,07). Como está contenido dentro de los límites 0,8 y 1,25 se puede considerar que T es bioequivalente en promedio a R.

3. BIOEQUIVALENCIA INDIVIDUAL

3.1 ¿Qué es la bioequivalencia individual?

Según Anderson y Hauck (1990) se declara *bioequivalencia individual* entre dos formulaciones test y referencia si la biodisponibilidad de la formulación test está “suficientemente cerca” de la biodisponibilidad de la formulación referencia en la mayoría de los individuos.

Hacia principios de los 90, los métodos para evaluar bioequivalencia individual estaban desarrollados desde dos aproximaciones. Una primera aproximación estaba referida a los *métodos de intervalos de tolerancia* que son los criterios basados en probabilidades. Estos métodos trabajaban con cocientes (o diferencias) individuales. Una aproximación totalmente diferente para evaluar bioequivalencia individual se propuso en presencia de diseños de crossover replicados. Estos métodos se basaban en un concepto sumamente simple: la diferencia entre las formulaciones test y referencia debían estar cercanas a las diferencias entre dos respuestas de la formulación referencia. Anderson (1993) denominó esta aproximación *estadísticos derivados (derived statistics)* ya que son funciones de distribución subyacente de los parámetros y no trabaja directamente con los cocientes individuales de las medidas de biodisponibilidad. Tienen además la ventaja de tratar directamente con el interés de la FDA de encontrar métodos que tengan explícitamente en cuenta la varianza dentro del sujeto con la formulación referencia. Con estos modelos es difícil demostrar bioequivalencia si la formulación referencia es muy variable.

3.2 Métrica y procedimientos inferenciales para IBE

La bioequivalencia individual (IBE) es un criterio para decidir si un paciente que está siendo tratado con la droga R puede cambiar a la droga T. Este criterio tiene en cuenta la variabilidad del sujeto (*within subject variability*) y la interacción sujeto–formulación (*subject by formulation interaction*) (FDA Guidance, 1997, 1999b, 2000, 2001).

Se necesitan las siguientes notaciones adicionales:

σ_{WT}^2 = varianza dentro del sujeto (*within subject variance*) con la formulación T

σ_{WR}^2 = varianza dentro del sujeto (*within subject variance*) con la formulación R

σ_{BT}^2 = varianza entre sujetos (*between subject variance*) para la droga T

σ_{BR}^2 = varianza entre sujetos (*between subject variance*) para la droga R

σ_D^2 = varianza interacción sujeto–formulación (*subject by formulation interaction variance*)

ρ = correlación entre sujetos de T y R

$\sigma_D^2 = \sigma_{BT}^2 + \sigma_{BR}^2 - 2\rho\sigma_{BT}\sigma_{BR}$

μ_T = media de T

μ_R = media de R

$\delta = \mu_T - \mu_R$

A partir de ahora se considerarán los datos transformados por logaritmo natural (es decir, $\ln(AUC)$)

Para evaluar IBE se utiliza la siguiente métrica agregada (FDA Guidance, 1997)

$$\eta = \frac{(\mu_T - \mu_R)^2 + \sigma_{WT}^2 - \sigma_{WR}^2 + \sigma_D^2}{\max\{\sigma_{WR}^2; 0,04\}} \quad (21)$$

Esta métrica permite testear:

$$H_0) \quad \eta \geq \theta_I \qquad H_1) \quad \eta < \theta_I \quad (22)$$

- Si $\sigma_{WR}^2 > 0,04$, se observa que:

$$\eta = \frac{(\mu_T - \mu_R)^2 + \sigma_{WT}^2 - \sigma_{WR}^2 + \sigma_D^2}{\max\{\sigma_{WR}^2; 0,04\}} = \frac{\delta^2 + \sigma_{WT}^2 - \sigma_{WR}^2 + \sigma_D^2}{\sigma_{WR}^2}$$

Luego:

$$\frac{\delta^2 + \sigma_{WT}^2 - \sigma_{WR}^2 + \sigma_D^2}{\sigma_{WR}^2} \geq \theta_I$$

$$\delta^2 + \sigma_{WT}^2 - \sigma_{WR}^2 + \sigma_D^2 \geq \theta_I \sigma_{WR}^2$$

$$\delta^2 + \sigma_{WT}^2 + \sigma_D^2 - (1 + \theta_I) \sigma_{WR}^2 \geq 0$$

Definiendo:

$$v_{IBE} = \delta^2 + \sigma_{WT}^2 + \sigma_D^2 - (1 + \theta_I) \sigma_{WR}^2 \quad (23)$$

las hipótesis (22) quedan:

$$H_0) \quad v_{IBE} \geq 0 \qquad H_1) \quad v_{IBE} < 0 \quad (24)$$

- Si $\sigma_{WR}^2 < 0,04$, se observa que

$$\eta = \frac{(\mu_T - \mu_R)^2 + \sigma_{WT}^2 - \sigma_{WR}^2 + \sigma_D^2}{\max\{\sigma_{WR}^2; 0,04\}} = \frac{\delta^2 + \sigma_{WT}^2 - \sigma_{WR}^2 + \sigma_D^2}{0,04} \quad (25)$$

Luego:

$$\frac{\delta^2 + \sigma_{WT}^2 - \sigma_{WR}^2 + \sigma_D^2}{0,04} \geq \theta_I$$

$$\delta^2 + \sigma_{WT}^2 - \sigma_{WR}^2 + \sigma_D^2 \geq \theta_I 0,04$$

$$\delta^2 + \sigma_{WT}^2 - \sigma_{WR}^2 + \sigma_D^2 - \theta_I 0,04 \geq 0$$

Definiendo:

$$v_{CIBE} = \delta^2 + \sigma_{WT}^2 - \sigma_{WR}^2 + \sigma_D^2 - \theta_I 0,04 \quad (26)$$

las hipótesis (22) quedan:

$$H_0) \quad v_{CIBE} \geq 0 \qquad H_1) \quad v_{CIBE} < 0 \quad (27)$$

A la métrica dividida por σ_{WR}^2 se la llama *escala de referencia* y a la que está dividida por 0,04, *escala constante*.

El valor de θ_I en H_0 es un valor igual a 2,49 (FDA Guidance, 2000). En la métrica a escala constante, si se asume que la diferencia de medias es $\ln(1,25)$, la interacción sujeto– formulación es 0,03 y la diferencia entre las varianzas dentro del sujeto para las formulaciones T y R es 0,02, se obtiene:

$$\theta_I = \frac{\delta^2 + \sigma_{WT}^2 - \sigma_{WR}^2 + \sigma_D^2}{0,04} = \frac{\ln(1,25) + 0,02 + 0,03}{0,04} = 2,49 \quad (28)$$

Para demostrar IBE la cota superior del intervalo de 90% de confianza para la métrica agregada (21) debe ser menor a 2,49.

Con el fin de establecer si se usa escala de referencia o escala constante, previamente se realiza el siguiente test:

$$H_0) \quad \sigma_{WR}^2 \geq \sigma_0^2 = 0,04 \qquad H_1) \quad \sigma_{WR}^2 < \sigma_0^2 = 0,04$$

$$\text{Rechazo } H_0 \text{ si } \frac{M_R(n-s)}{\chi_{1-\alpha; n-s}} < 0,04$$

En general, Patterson y Jones (2002b) proponen usar escala de referencia si $\hat{\sigma}_{WR}^2 > 0,04$ y escala constante en caso contrario.

3.2 Métodos de estimación de IBE

i) Método de los momentos

Este método está propuesto por la FDA en la guía del 2001 y 2003 y se basa en Chinchilli (1996) y Chinchilli and Esinhart (1996), permite obtener estimadores insesgados de las componentes del modelo: diferencia de medias, suma de varianzas del numerador y la varianza dentro del sujeto con la formulación R. La varianza entre sujetos no se estima. El problema de este método es que es potencialmente sensible al sesgo producido por datos faltantes (Patterson and Jones, 2002).

Hyslop, Hsuan y Holder (2000) consideraron la siguiente descomposición de las métricas (23) y (26):

$$V_{IBE} = \delta^2 + \sigma_I^2 + 0,5\sigma_{WT}^2 - 1,5\sigma_{WR}^2 - \theta_I \sigma_{WR}^2 \quad (29)$$

$$V_{CIBE} = \delta^2 + \sigma_I^2 + 0,5\sigma_{WT}^2 - 1,5\sigma_{WR}^2 - 0,04\theta_I \quad (30)$$

$$\sigma_I^2 = \sigma_D^2 + \frac{\sigma_{WT}^2 + \sigma_{WR}^2}{2}$$

La razón de esta descomposición es que se pueden obtener estimadores insesgados e independientes $\hat{\delta}$, M_I , M_T y M_R de $\delta = \mu_T - \mu_R$, $\sigma_I^2 = \sigma_D^2 + \frac{\sigma_{WT}^2 + \sigma_{WR}^2}{2}$, σ_{WT}^2 y σ_{WR}^2 respectivamente (Patterson, 2001a, 2001b).

A continuación se enuncian las notaciones y definiciones que se usarán en esta sección para el cálculo de los estimadores y sus varianzas.

Recordar que s representa la cantidad de secuencias, n_i el número de sujetos en la secuencia i y

$$n-s = \left(\sum_{i=1}^s n_i \right) - s$$

$$Y_{ijT\bullet} = \frac{1}{2}(Y_{ijT1} + Y_{ijT2}) \quad Y_{ijR\bullet} = \frac{1}{2}(Y_{ijR1} + Y_{ijR2}) \quad \bar{Y}_{i\bullet k} = \frac{1}{n_i} \sum_{j=1}^{n_i} \left[\frac{1}{2} \sum_{l=1}^2 Y_{ijkl} \right] = \frac{1}{n_i} \sum_{j=1}^{n_i} Y_{ijk}$$

$$I_{ij} = Y_{ijT\bullet} - Y_{ijR\bullet} \quad \bar{I}_{i\bullet} = \frac{1}{n_i} \sum_{j=1}^{n_i} I_{ij}$$

$$\begin{aligned}
 T_{ij} &= Y_{ijT1} - Y_{ijT2} & \bar{T}_{i\bullet} &= \frac{1}{n_i} \sum_{j=1}^{n_i} T_{ij} \\
 R_{ij} &= Y_{ijR1} - Y_{ijR2} & \bar{R}_{i\bullet} &= \frac{1}{n_i} \sum_{j=1}^{n_i} R_{ij} \\
 \hat{\mu}_k &= \frac{1}{s} \sum_{i=1}^s \bar{Y}_{i\bullet k} & k &= T, R
 \end{aligned} \tag{31}$$

$$\hat{\delta} = \hat{\mu}_T - \hat{\mu}_R = \frac{1}{s} \sum_{i=1}^s \left[\frac{1}{n_i} \sum_{j=1}^{n_i} I_{ij} \right] \tag{32}$$

$$M_I = \hat{\sigma}_I^2 = \frac{1}{n-s} \sum_{i=1}^s \sum_{j=1}^{n_i} (I_{ij} - \bar{T}_{i\bullet})^2 \tag{33}$$

$$M_T = \hat{\sigma}_{WT}^2 = \frac{1}{2(n-s)} \sum_{i=1}^s \sum_{j=1}^{n_i} (T_{ij} - \bar{T}_{i\bullet})^2 \tag{34}$$

$$M_R = \hat{\sigma}_{WR}^2 = \frac{1}{2(n-s)} \sum_{i=1}^s \sum_{j=1}^{n_i} (R_{ij} - \bar{R}_{i\bullet})^2 \tag{35}$$

Las métricas linealizadas U_{IBE} y U_{CIBE} dadas en (29) y (30) pueden ser estimadas de la siguiente forma (FDA Guidance, 1999a, 1999b, 2000, 2001):

$$\begin{aligned}
 \hat{U}_{IBE} &= \hat{\delta}^2 + M_I + 0,5M_T - (1,5 + \theta_I)M_R \\
 \hat{U}_{CIBE} &= \hat{\delta}^2 + M_I + 0,5M_T - 1,5M_R - 0,04\theta_I
 \end{aligned}$$

La guía de la FDA del 2001 adopta el método de Hyslop, Hsuan y Holder (2000) para construir la cota superior del intervalo de confianza para demostrar IBE. Este método, basado en los resultados de Howe (1974), Graybill and Wang (1980), y Ting, Burdick, Graybill, Jeyaratnam and Lu (1990), puede ser descripto como sigue.

Sea $\gamma = \gamma_1 + \dots + \gamma_m$. Sean $\hat{\gamma}_i$ los estimadores de γ_i y $\tilde{\gamma}_i$ las cotas superiores de intervalos del $100(1-\alpha)\%$ de confianza de γ_i . Luego, si $\hat{\gamma}_1, \dots, \hat{\gamma}_m$ son independientes, una cota superior aproximada del intervalo de $100(1-\alpha)\%$ de confianza para γ es:

$$\hat{\gamma}_U = \hat{\gamma}_1 + \dots + \hat{\gamma}_m + \sqrt{(\tilde{\gamma}_1 - \hat{\gamma}_1)^2 + \dots + (\tilde{\gamma}_m - \hat{\gamma}_m)^2}$$

La clave para utilizar este método es la independencia de los estimadores.

El Cuadro 6 y el Cuadro 7 dados a continuación (FDA Guidance, 2001) permiten construir la

cota superior del intervalo del $100(1-\alpha)\%$ de confianza donde $n = \sum_{i=1}^s n_i$ y s es el número de

secuencias. Además, si $T \sim t_{n-s}$ entonces $P(T > t_{\alpha;n-s}) = \alpha$ y si $X \sim \chi_{n-s}^2$ entonces $P(X > \chi_{\alpha;n-s}^2) = \alpha$

- Para la escala de referencia:

Cuadro 6

H_q = cota superior del intervalo de confianza	E_q = parámetro estimado	$U_q = (H_q - E_q)^2$
$H_D = \left[\hat{\delta} + t_{\alpha;n-s} \left(\frac{M_I}{s^2} \sum_{i=1}^s \frac{1}{n_i} \right)^{1/2} \right]^2$	$E_D = \hat{\delta}^2$	U_D
$H_I = \frac{(n-s)M_I}{\chi_{1-\alpha;n-s}^2}$	$E_I = M_I$	U_I
$H_T = \frac{0,5(n-s)M_T}{\chi_{1-\alpha;n-s}^2}$	$E_T = 0,5M_T$	U_T
$H_R = \frac{-(1,5 + \theta_I)(n-s)M_R}{\chi_{\alpha;n-s}^2}$	$E_R = -(1,5 + \theta_I)M_R$	U_R

Fuente: FDA Guidance (2001).

Luego, la cota superior del intervalo del $100(1-\alpha)\%$ de confianza para U_{IBE} es:

$$H_{U_{IBE}} = \sum E_q + \left(\sum U_q \right)^{1/2}$$

- Para la escala constante:

Cuadro 7

H_q = cota superior del intervalo de confianza	E_q = parámetro estimado	$U_q = (H_q - E_q)^2$
$H_D = \left[\hat{\delta} + t_{\alpha;n-s} \left(\frac{M_I}{s^2} \sum_{i=1}^s \frac{1}{n_i} \right)^{1/2} \right]^2$	$E_D = \hat{\delta}^2$	U_D
$H_I = \frac{(n-s)M_I}{\chi_{1-\alpha;n-s}^2}$	$E_I = M_I$	U_I
$H_T = \frac{0,5(n-s)M_T}{\chi_{1-\alpha;n-s}^2}$	$E_T = 0,5M_T$	U_T
$H_R = \frac{-1,5(n-s)M_R}{\chi_{\alpha;n-s}^2}$	$E_R = -1,5M_R$	U_R

Fuente: FDA Guidance (2001).

Luego, la cota superior del intervalo del $100(1 - \alpha)\%$ de confianza para ν_{CIBE} es:

$$H_{\nu_{CIBE}} = \sum E_q - 0,04\theta_I + \left(\sum U_q\right)^{1/2}$$

Por lo tanto, si la cota superior $H_{\nu_{IBE}}$ (o bien, $H_{\nu_{CIBE}}$) es negativa o cero se demuestra IBE.

ii) Aproximación asintótica

Este método permite hallar una aproximación asintótica de la media y la varianza de ν_{IBE} y ν_{CIBE} (Patterson, 2001; Patterson and Jones, 2002a,b). El uso de máxima verosimilitud restringida asegura estimadores insesgados aún en conjuntos de datos con datos perdidos o no balanceados. En general, para calcular IBE en crossover replicados se necesitan tamaños de muestra que superen los 20 o 30 sujetos, por lo tanto una aproximación asintótica no es irracional.

Se obtienen los siguientes estimadores: $\hat{\sigma}_{BT}^2$, $\hat{\sigma}_{BR}^2$, $\hat{\sigma}_{WT}^2$, $\hat{\sigma}_{WR}^2$ y $\hat{\omega}_{RT}$ los cuales están normalmente distribuidos en el límite con matriz de varianza-covarianza apropiada al modelo planteado.

Para resolver el modelo por este método se puede usar SAS *proc mixed* con la opción de REML. Los estimadores de la varianza y covarianza obtenida con REML tienen como valores esperados el siguiente vector y la siguiente matriz:

$$\begin{pmatrix} \sigma_{BT}^2 \\ \omega_{RT} \\ \sigma_{BR}^2 \\ \sigma_{WT}^2 \\ \sigma_{WR}^2 \end{pmatrix} \begin{pmatrix} l_{BT} & l_{BT \times \omega} & l_{BT \times BR} & l_{BT \times WT} & l_{BT \times WR} \\ l_{BT \times \omega} & l_{\omega} & l_{BR \times \omega} & l_{\omega \times WT} & l_{\omega \times WR} \\ l_{BT \times BR} & l_{BR \times \omega} & l_{BR} & l_{BR \times WT} & l_{BR \times WR} \\ l_{BT \times WT} & l_{\omega \times WT} & l_{BR \times WT} & l_{WT} & l_{WT \times WR} \\ l_{BT \times WR} & l_{\omega \times WR} & l_{BR \times WR} & l_{WT \times WR} & l_{WR} \end{pmatrix}$$

Si $\hat{\sigma}_{WR}^2 > 0,04$ entonces un estimador de ν_{IBE} para la escala de referencia (FDA Guidance, 2001) es:

$$\hat{\nu}_{IBE} = \hat{\delta}^2 + \hat{\sigma}_{BT}^2 + \hat{\sigma}_{BR}^2 - 2\hat{\omega}_{RT} + \hat{\sigma}_{WT}^2 - (1 + \theta_I) \hat{\sigma}_{WR}^2$$

Este estimador está distribuido en forma normal asintóticamente (Patterson, 2001; Patterson and Jones, 2002b) y es insesgado con esperanza y varianza:

$$E(\hat{\nu}_{IBE}) = \delta^2 + \sigma_D^2 + \sigma_{WT}^2 - (1 + \theta_I) \sigma_{WR}^2$$

$$V(\hat{\nu}_{IBE}) = 4\sigma_{\delta}^2 \delta^2 + l_{BT} + l_{BR} + 4l_{\omega} + l_{WT} + (1 + \theta_I)^2 l_{WR} + 2l_{BT \times BR} - 4l_{BT \times \omega} + 2l_{BT \times WT} - 2(1 + \theta_I)l_{BT \times WR} + 2l_{BR \times WT} - 4l_{BR \times \omega} - 2(1 + \theta_I)l_{BR \times WR} - 4l_{\omega \times WT} + 4(1 + \theta_I)l_{\omega \times WR} - 2(1 + \theta_I)l_{WT \times WR}$$

Análogamente, si $\hat{\sigma}_{WR}^2 \leq 0,04$ entonces un estimador de ν_{CIBE} para la escala constante (FDA Guidance, 2001) es:

$$\hat{\nu}_{CIBE} = \hat{\delta}^2 + \hat{\sigma}_{BT}^2 + \hat{\sigma}_{BR}^2 - 2\hat{\omega}_{RT} + \hat{\sigma}_{WT}^2 - \hat{\sigma}_{WR}^2 - 0,04\theta_I$$

$$E(\hat{\nu}_{CIBE}) = \delta^2 + \sigma_D^2 + \sigma_{WT}^2 - \sigma_{WR}^2 - 0,04\theta_I$$

$$V(\hat{\nu}_{CIBE}) = 4\sigma_\delta^2 \delta^2 + l_{BT} + l_{BR} + 4l_\omega + l_{WT} + l_{WR} + 2l_{BT \times BR} - 4l_{BT \times \omega} + 2l_{BT \times WT} - 2l_{BT \times WR} + 2l_{BR \times WT} - 4l_{BR \times \omega} - 2l_{BR \times WR} - 4l_{\omega \times WT} + 4l_{\omega \times WR} - 2l_{WT \times WR}$$

Por lo tanto, cota superior del intervalo de 90% de confianza asintótico puede ser calculada:

$$\hat{\nu}_{IBE} + 1,645 \sqrt{\widehat{Var}(\hat{\nu}_{IBE})} \quad \text{o bien} \quad \hat{\nu}_{CIBE} + 1,645 \sqrt{\widehat{Var}(\hat{\nu}_{CIBE})}$$

donde los estimadores de la varianza se obtienen reemplazando los valores estimados de las varianzas y covarianzas en las fórmulas de $V(\hat{\nu}_{IBE})$ o $V(\hat{\nu}_{CIBE})$.

Finalmente, si la cota superior es negativa o cero se demuestra IBE.

3.3 Aplicación a datos experimentales

El Cuadro 6 corresponde a la cota superior del intervalo de 90% de confianza para IBE

Cuadro 8

Datos	Método de los Referencia	Momentos Constante	Aproximación Referencia	Asintótica Constante
Ejemplo 1	0,184	0,212	0,184	0,159
Ejemplo 2	-0,205	-0,069	0,049	0,302
Sin sujeto 5				
Ejemplo 2	-----	-----	0,039	0,236
Con sujeto 5				
Ejemplo 3	-0,069	-0,024	-0,009	0,215

Fuente: Elaboración propia.

Los valores en negrita satisfacen la condición de negatividad de la cota superior del intervalo de 90% de confianza para establecer IBE.

El Cuadro 9 corresponde al intervalo de 90% de confianza para los estimadores de la varianza utilizando REML.

Cuadro 9

Datos	Intra Test $\hat{\sigma}_{WT}^2$	Subject Referencia $\hat{\sigma}_{WR}^2$	Inter Test $\hat{\sigma}_{BT}^2$	Subject Referencia $\hat{\sigma}_{BR}^2$
Ejemplo 1	0,0568 (0,027;0,208)	0,0584 (0,028;0,214)	0,1047 (0,042;0,690)	0,0035 (0,002; ∞)
Ejemplo 2 Sin sujeto 5	0,0975 (0,048;0,315)	0,1458 (0,072;0,471)	0,1956 (0,084;1,004)	0,2694 (0,115;1,45)
Ejemplo 2 Con sujeto 5	0,098 (0,049;0,31)	0,1447 (0,073;0,457)	0,1696 (0,074;0,848)	0,2322 (0,1;1,219)
Ejemplo 3	0,1272 (0,08;0,239)	0,0745 (0,047;0,14)	0,2078 (0,115;0,511)	0,2791 (0,166;0,584)

Fuente: Elaboración propia basada en datos de la bibliografía.

4. CONCLUSIONES

En este trabajo se presentaron métodos estadísticos eficientes que permitieran establecer bioequivalencia en promedio en diseños replicados en ausencia de efectos carryover.

La diferencia fundamental en los tres métodos desarrollados radica en la forma de estimar la varianza de la diferencia de medias, razón por la cual se obtienen distintos intervalos de $(1 - 2\alpha) \times 100\%$ de confianza para $\mu_T - \mu_R$.

El primer método presentado fue el propuesto por Chow y Liu (1992) en el que se consideró nula la varianza sujeto–formulación ($\sigma_D^2 = 0$). Se realizó la estimación de la varianza de la diferencia de medias suponiendo primero que las varianzas dentro de los sujetos eran iguales para las dos formulaciones ($\sigma_{WT}^2 = \sigma_{WR}^2$) y luego suponiendo que eran distintas.

Posteriormente se desarrollaron dos métodos propuestos por la Food and Drug Administration en la guía del 2001. Con el método de los momentos se obtuvieron estimadores insesgados e independientes de la diferencia de medias ($\delta = \mu_T - \mu_R$) y de la varianza ($\sigma_I^2 = \sigma_D^2 + \frac{\sigma_{WT}^2 + \sigma_{WR}^2}{2}$). Luego, se presentó el método de *aproximación asintótica* basado en los estimadores de máxima verosimilitud restringida (REML), el cual permitió encontrar estimadores de cada uno de los componentes de la varianza.

Una diferencia importante de estos dos últimos métodos es que en el método de los momentos si un sujeto no recibió alguna de las dosis R o T en alguno de los períodos (*incomplete data*) no se debe incluir en el estudio, mientras que en la aproximación asintótica se puede realizar la estimación aún con datos faltantes. Si bien esta diferencia significa una ventaja para este último

método, ya que es usual en estudios de bioequivalencia que los datos estén incompletos, la desventaja es que en muchas situaciones de bioequivalencia en promedio, sin embargo, el tamaño de la muestra es menor a 30, con lo cual el procedimiento asintótico se vuelve cuestionable.

Estos métodos se aplicaron a los datos experimentales. El primer ejemplo corresponde a un diseño replicado de dos secuencias y cuatro períodos (TRTR / RTRT), el segundo ejemplo corresponde al estudio de dos secuencias y cuatro períodos (TRRT / RTTR) y el tercer ejemplo corresponde a un diseño de cuatro secuencias y cuatro períodos (TRTR / RTRT / TRRT / RTTR).

Se pudo observar que se arribó a conclusiones idénticas en cada uno de los ejemplos al buscar los intervalos del 90% de confianza para $\frac{\mu_T}{\mu_R}$ en la escala natural con cada uno de los métodos

detallados. Más aún, en el ejemplo de olsalazine los datos son incompletos (ya que al sujeto 5 no se le administró ni R ni T en el segundo y tercer período) pero las conclusiones utilizando aproximación asintótica considerándolo en el estudio y excluyéndolo de él fueron iguales.

Luego se describieron métodos estadísticos que permitieran establecer bioequivalencia individual en un modelo con ausencia de efectos carryover.

Se desarrollaron dos métodos propuestos por la FDA. Con el método de los momentos se obtuvieron estimadores insesgados de las componentes del modelo: diferencia de medias, suma de varianzas del numerador y la varianza dentro del sujeto (*within subject variability*) con la formulación R. La varianza entre sujetos (*between subject variability*) no se estima. El problema de este método es que es potencialmente sensible al sesgo producido por datos faltantes (Patterson and Jones, 2002).

La guía de la FDA del 2001 adopta el método de Hyslop, Hsuan y Holder (2000) para construir la cota superior del intervalo de $(1 - \alpha) \times 100\%$ de confianza para demostrar IBE.

El método de aproximación asintótica permitió hallar una aproximación asintótica de la media y la varianza de ν_{IBE} y ν_{CIBE} (Patterson, 2003; Patterson and Jones, 2002,a,b). El uso de máxima verosimilitud restringida asegura estimadores insesgados aún en conjuntos de datos con datos perdidos o no balanceados.

Al aplicar estos métodos a los datos de la bibliografía se obtuvieron cotas superiores de los intervalos del 90% de confianza para IBE distintos. La conclusión a la que se llega es que el tamaño de muestra de los ejemplos no es suficientemente grande como para aplicar la aproximación asintótica, provocando una diferencia en las conclusiones con los dos métodos.

En conclusión se justifica la aplicación de diseños replicados en estudios en los que se desee evaluar una interacción o una varianza intra individual específica, ya que brindan mayor información que otros diseños

La bioequivalencia individual presenta una desventaja económica con respecto a la bioequivalencia en promedio sin sumar beneficios significativos adicionales, ya que constituye una metodología que requiere mayor número de administraciones a voluntarios y un tiempo más prolongado de estudio.

REFERENCIAS

- Anderson, S. (1993) Individual Bioequivalence: a problema of switchability. *Biopharmaceutical Reports*, 2, 1-11.
- Anderson, S.; Hauck, W.W. (1990) Consideration of Individual Bioequivalence. *Journal of Pharmacokinetics and Biopharmaceutics*, 18, 259-273.
- Bianco, M.J. (2015) *Análisis económico y financiero de la implementación de una metodología alternativa en la evaluación de medicamentos genéricos* (Tesis de Doctorado) Facultad de Ciencias Económicas, Universidad de Buenos Aires, Argentina.
- Chinchill, V.M., (1996). The assessment of individual and population bioequivalence. *Journal of Biopharmaceutical Statistics*, 6, 1-14.
- Chinchilli, V.M.; Esinhart, J.D. (1996) Design and analysis of intra-subject variability in cross-over experiments. *Statistics in Medicine*, 15, 1619-1634.
- Chow, S-C; Liu, J-P (1992) *Design and analysis of bioavailability and bioequivalence studies*. Marcel Dekker, NY.
- Ekbohm, G.; Melander, H. (1989) The subject by formulation interaction as a criterion of interchangeability of drugs. *Biometrics*, 45, 1249-1254.
- Esinhart, J.D.; Chinchilli, V.M. (1994) Extension to the use of tolerance intervals for the assessment of individual bioequivalence. *Journal of Biopharmaceutical Statistics*, 4, 39-52.
- FDA Guidance (1992) Statistical procedures for bioequivalence studies using a standard two treatment crossover design.
- FDA Draft Guidance (1997) In vivo bioequivalence studies based on population and individual bioequivalence approaches.

- FDA Draft Guidance (1999a) BA and BE studies for orally administered drug products: General considerations.
- FDA Draft Guidance (1999b) Average, population and individual approaches to establishing bioequivalence.
- FDA Guidance (2000) Bioavailability and bioequivalence studies for orally administered drug products: General considerations.
- FDA Guidance (2001) Statistical approaches to establishing bioequivalence.
- FDA Guidance (2003) Bioavailability and bioequivalence studies for orally administered drug products: General considerations.
- FDA (2003) Statistical Methods for Obtaining Confidence Intervals for Individual and Population Bioequivalence Criteria.
- Fisher, R.A. (1925) Applications of Student's distribution. *Metron*, **5**, 90-104.
- Graybill, F.A.; Wang, C.M. (1980) Confidence intervals on nonnegative linear combinations of variances. *Journal of American Statistics Association*, **75**, 869-873.
- Howe, W.G. (1974) Approximate confidence limits on the mean of $X+Y$ donde X and Y are two tabled independent random variables. *Journal of American Statistics Association*, **69**, 789-794
- Hyslop, T.; Hsuan, F.; Holder, D.J. (2000) A small sample confidence interval approach to assess individual bioequivalence. *Statistics in Medicine*, **19**, 2885-2897.
- Jones, B.; Kenward, M. (2003) *Design and Analysis of Cross-Over Trials*. Chapman and Hall, NY.
- Kenward, M.; Jones, R. (1997) Small sample inference for fixed effects from Restricted Maximum Likelihood. *Biometrics*, **33**, 983-997.
- Milliken, G.A.; Johnson, D.E. (1992) *Analysis of messy data. Vol 1: Designed Experiments*, Chapman and Hall, NY.
- Patterson, S. (2001a) A review of the development of bioestatistical design and analysis techniques for assessing *in vivo* bioequivalence. Part 1, *Indian Journal of Pharmaceutical Sciences*, **63**, 81-100.
- Patterson, S. (2001b) A review of the development of biostatistical design and analysis techniques for assessing *in vivo* bioequivalence. Part 2, *Indian Journal of Pharmaceutical Sciences*, **63**, 81-100.
- Patterson, S.; Jones, B. (2002a) Bioequivalence and the Pharmaceutical Industry. *Pharmaceutical Statistics*, **1**, 83-95.
- Patterson, S.; Jones, B. (2002b) Replicate design and average, individual and population bioequivalence: I. Estimation, inference and retrospective assessment of performance of

novel procedures an the proposed FDA methods for bioequivalence assessment. *Glaxo-SmithKline BDS Technical Report 2002-01*.

Patterson, S.; Jones, B. (2002b) Replicate design and average, individual and population bioequivalence: II. Simulation study of performance of novel procedures and the proposed FDA methods for bioequivalence assessment. *Glaxo-SmithKline BDS Technical Report 2002-02*.

Ryde, M.; Huitfeldt, B.; Pettersson, R. (1991) Relative bioavailability of Olsalazine from tables and capsules: a drug targeted for local effect in the colon. *Biopharm & Drug Disposition*, 12, 233-246.

Schuirmann, D.J. (1981) On hypothesis testing to determine if the mean of a normal distribution is continued in a known interval. *Biometrics*, 37, 617-621

Schuirmann, D.J. (1987) A comparison of the one sided tests procedure and the power approach for assessing the equivalence of average bioavailability. *Journal of Pharmacokinetics and Biopharmaceutics*, 15, 657-680.

Student (1908) On the probable error of the mean, *Biometrika*, 6, 1-25.

Ting, N.; Burdick, R.K.; Graybill, F.A.; Jeyaratnam, S.; Lu, T.F.C. (1990) Confidence interval on linear combinations of variance components that are unrestricted in sign. *Journal of Statistical Computation and Simulation*, 35, 135-143.

Vonesh, E.F.; Chinchilli, V.M. (1997) *Linear and Nonlinear models for the analysis of repeated measurements*. Marcel Dekker, NY.

Wang, W.; Hwang, J.T. (2000) A nearly unbiased test for individual bioequivalence problems using probability criteria. *Journal of Statistical Planning and Inference*, 99, 41-58.