RCAN

Rev Cubana Aliment Nutr 2010;20(1):152-167

Instituto Superior de Ciencias Médicas de La Habana.

PROCEDERES DE REGRESIÓN LINEAL COMO SOLUCIONES AL PROBLEMA DE LA COMPARACIÓN DE MÉTODOS. I. ERRORES ANALÍTICOS CONSTANTES E IGUALES.

Ariel Delgado Ramos[1], *Ramón Ramos Salazar*[1], *Humberto Martínez Canalejo*[2], *Sergio Santana Porbén*[3].

RESUMEN

En este trabajo se discute el desempeño de 4 soluciones del problema de comparación de métodos: P1: Regresión de Mínimos Cuadrados Ordinarios; P2: Regresión de Passing-Bablok; P3: Regresión de Deming con coeficiente λ estimado a partir de las varianzas intrarreplicados; y P4: Regresión de Deming con coeficiente λ estimado a partir de las varianzas de las series; cuando los errores analíticos correspondientes pueden asumirse como iguales y constantes en el rango de concentraciones de interés. Se evaluó también el desempeño de la Regresión de Deming con $\lambda = 1$ (P5). Según el rango de las observaciones, se construyeron 3 escenarios analíticos diferentes: el caso del Sodio, el caso de la Albúmina, y el caso de la Glucosa. El desempeño de cada proceder se documentó con series de 30, 50 y 100 parejas de números generados seudoaleatoriamente. El proceder debía asegurar que el estimado de la pendiente de la recta de comparación de métodos fuera estadísticamente igual a la unidad; el error de estimación de la pendiente fuera el menor posible; y el número de rechazos de la hipótesis nula $H_o : \beta = 1$ sea igual (o menor) del 5%. El proceder P1 se caracterizó por devolver estimados de la pendiente diferentes de la unidad. Se constató que el error de estimación de la pendiente mediante el proceder P1 estaba "inflado" por la presencia de un error sistemático. El uso del proceder P1 resultó en una frecuencia elevada de rechazos de la hipótesis nula $H_o : \beta = 1$. Los otros 3 procederes devolvieron estimados insesgados de la pendiente de la recta de comparación de métodos, pero a costa muchas veces de una frecuencia elevada de rechazos de la hipótesis nula $H_o : \beta = 1$. A pesar de este inconveniente, puede recomendarse su uso como soluciones del problema de comparación de métodos en lugar de la Regresión de Mínimos Cuadrados. En el caso particular del proceder P4, se recomienda su empleo cuando se han obtenido determinaciones únicas con cada método en comparación, y se puede asumir, a partir de los datos de la validación de los métodos, que los errores analíticos son similares. ***Delgado Ramos A, Ramos Salazar R, Martínez Canalejo H, Santana Porbén S***. *Procederes de regresión líneal como soluciones al problema de la comparación de métodos. I. Errores analíticos constantes e iguales. RCAN Rev Cubana Aliment Nutr 2010;20(1):152-167. RNP: 221. ISSN: 1561-2929.*

DeCS: Comparación de métodos / Simulación de métodos / VisualBasic / Mínimos cuadráticos / Passing-Bablok / Deming.

[1] Especialista de Primer Grado en Biostadística. [2] Licenciado en Matemática. [3] Especialista de Segundo Grado en Bioquímica Clínica.

Aparecido en forma abreviada en: Contacto químico (Michoacán) 2007;2(6):21-23.

Sergio Santana Porbén. Hospital Clínico quirúrgico "Hermanos Ameijeiras". La Habana. San Lázaro 701 e/t Marqués González y Belascoaín. Centro Habana. La Habana 10300. Cuba.
Correo electrónico: ssergito@infomed.sld.cu

INTRODUCCIÓN

El problema de la comparación de métodos continúa atrayendo la atención de analistas y matemáticos.[1] Los esfuerzos se dirigen a obtener una solución que sea universalmente aceptable.[2] Con tal fin, se han propuesto tres procederes distintos: la Regresión de Mínimos Cuadrados Ordinarios,[3] la Regresión de Passing-Bablok,[4] y la Regresión de Deming.[5-6] El desempeño de estos procederes se ha caracterizado mediante técnicas de simulación estadístico-matemática.[7-8] A pesar de los resultados documentados en la literatura internacional, todavía se echa de ver una evaluación sistemática y exhaustiva de estos procederes bajo diferentes escenarios analíticos que se correspondan con las situaciones experimentales que el analista puede encontrar en la práctica diaria.

En este trabajo se discute el desempeño de la Regresión de Mínimos Cuadrados Ordinarios (P1), la Regresión de Passing-Bablok (P2), y la Regresión de Deming con coeficiente λ estimado a partir de las varianzas intrarreplicados (P3), cuando los errores analíticos correspondientes pueden asumirse como iguales. También se presenta el desempeño de una modificación de la Regresión de Deming en la que el coeficiente λ se estima a partir de las varianzas de las series (P4). En un trabajo acompañante se discute el desempeño de estos 4 procederes cuando los errores analíticos son estadísticamente diferentes.[9]

MATERIAL Y MÉTODOS

El problema de la comparación de métodos se puede exponer según los lineamientos siguientes. Se desea evaluar la equivalencia de los resultados emitidos por dos métodos analíticos A y B en comparación (que pueden o no diferir en sus fundamentos teóricos y filosóficos). Se asume que la (in)exactitud del método A está exhaustivamente documentada, por lo que se le asigna el papel de METODO DE REFERENCIA, y ocupa el papel de la variable predictora X. El método A puede representar un procedimiento físico, inmunológico, biológico, o químico con un rendimiento analítico conocido. El método B representa aquel cuya (in)exactitud se desea verificar por contrastación con el método A, y por lo tanto, ocupa el papel de la variable a predecir Y. La equivalencia de los resultados emitidos por A y B debe traducirse en sendos estimados de los errores sistemáticos proporcional y constante. Es por ello que la recta $Y = \alpha + \beta X$ puede considerarse como la respuesta intuitiva del problema de la comparación de métodos. En este contexto, α aportaría el estimado del error sistemático constante; mientras que β aportaría el estimado del error sistemático proporcional. Se concluye que ambos métodos son equivalentes si α es estadísticamente igual a cero, y β es estadísticamente igual a 1. La recta asumiría la forma $Y = X$. Este caso particular se conoce como la recta de identidad. Se hace notar que en la práctica analítica, interesa más la segunda demostración, esto es, H_o : $\beta = 1$.

Para dar respuesta al problema de la comparación de métodos, se conduce el diseño experimental siguiente: Se acopian N muestras biológicas frescas; Cada muestra se divide en dos alícuotas; La primera de las alícuotas se ensaya por el método A; La segunda de las alícuotas se ensaya por el método B; y Se cuida de que el error de calibración intermétodo sea el menor posible. Para ello, las alícuotas ensayadas por el método particular se contrastan contra el mismo calibrador. Sean entonces X: los resultados obtenidos con el método A; mientras que Y: los resultados obtenidos con el método B. Las series de valores resultantes se someten a un procedimiento estadístico especificado. Se espera que este procedimiento estadístico devuelva estimados exactos de α y β.

Especificaciones del desempeño del procedimiento estadístico empleado en la solución del problema de comparación de métodos.

Para que un procedimiento estadístico sea una solución aceptable del problema expuesto de la comparación de métodos, debe satisfacer las especificaciones siguientes: (1) Ausencia de sesgo en la estimación de la pendiente β: el procedimiento estadístico debe devolver estimados de β que no sean estadísticamente diferentes de 1 bajo la hipótesis nula H_o : X = Y; y (2) El procedimiento estadístico escogido debe satisfacer la hipótesis nula H_o : $\beta = 1$ si se asegura que X = Y. Dicho con otras palabras, el procedimiento escogido no debe conducir al rechazo de la hipótesis nula H_o : $\beta = 1$ cuando se sabe que X = Y.

Diseño del algoritmo de simulación estadístico-matemática.

El desempeño del procedimiento estadístico particular, esto es, la capacidad para satisfacer las especificaciones apuntadas anteriormente, se evaluó mediante un algoritmo de simulación estadístico-matemática redactado en VisualBasic©® para EXCEL©® de OFFICE©® (Microsoft, Redmont, Virginia, Estados Unidos). Con este algoritmo se generaron N (N = 30, 50, 100) parejas ($\overline{X}$, $\overline{Y}$), donde $\overline{X}$, $\overline{Y}$: media de los duplicados de los números generados (seudo)aleatoriamente bajo el modelo experimental Y = X. Brevemente, el algoritmo funciona de la manera siguiente: Se genera un número x_1 que se distribuye uniformemente en el rango de posibles valores; Se obtiene un número y_1 bajo el modelo experimental escogido; El número x_1 se deforma por adición de un componente de error aleatorio que se distribuye normalmente. El tamaño de este componente aleatorio está determinado por la desviación estándar del método analítico que "supuestamente" genera los valores de x; El número y_1 se deforma por adición de un componente de error aleatorio que se distribuye normalmente. El tamaño de este componente aleatorio está determinado por la desviación estándar del método analítico que "supuestamente" genera los valores de y; Se genera x_2: el duplicado de x_1, por adición de la distancia permisible para el duplicado (Ecuaciones 1-2):

$$\text{Distancia} = \text{Desviación estándar (método A)} * \sqrt{2} \qquad (1)$$

$$x_2 = x_1 + \text{Distancia} \qquad (2)$$

Por su parte, el número x_2 se deforma por adición de un componente de error aleatorio que se distribuye normalmente. El tamaño de este componente aleatorio está determinado por la desviación estándar del método analítico que "supuestamente" genera los valores de x. Se genera entonces y_2: el duplicado de y_1, por adición de la distancia permisible para el duplicado (Ecuaciones 3-4):

$$\text{Distancia} = \text{Desviación estándar (método B)} * \sqrt{2} \qquad (3)$$

$$y_2 = y_1 + \text{Distancia} \qquad (4)$$

El número y_2 se deforma por adición de un componente de error aleatorio que se distribuye normalmente. El tamaño de este componente aleatorio está determinado por la desviación estándar del método analítico que "supuestamente" genera los valores de y.

El número $\overline{X}$ se obtuvo como la media de x_1, x_2 (Ecuación 5):

$$\overline{X} = 0.5 * (x_1 + x_2) \tag{5}$$

Mientras, el número $\overline{Y}$ se obtuvo como la media de y_1, y_2 (Ecuación 6):

$$\overline{Y} = 0.5 * (y_1 + y_2) \tag{6}$$

El desempeño del procedimiento estadístico considerado como solución del problema de comparación de métodos se evaluó en 3 escenarios analíticos, según el rango de valores posibles de los analitos ensayados en la práctica analítica:

Rango	Caso analítico	Li - Ls (Ls/Li)
Estrecho: 1.0 - 1.5	Sodio	125 - 145 (1.16)
Intermedio: 1.6 - 9.9	Albúmina	17.0 - 55.0 (3.23)
Extendido: ≥ 10.0	Glucosa	1.5 - 25 (10.0)

En el caso de la Albúmina, se trató de que el 25% de los valores se encontrara en la mitad inferior del rango, mientras que en el caso de la Glucosa, el 75% de los valores se encontraba en la mitad inferior del rango. En cualquiera de los 3 casos analíticos se asumió que la desviación estándar analítica era constante en el rango de posibles valores, e igual para los métodos en comparación:

Caso analítico	Tamaño de la desviación estándar
Sodio	1.355
Albúmina	1.275
Glucosa	0.7

Los procedimientos analíticos evaluados como soluciones del problema de la comparación de métodos fueron los siguientes: Regresión de Mínimos Cuadrados Ordinarios (P1); Regresión de Passing-Bablok (P2); Regresión de Deming (P3), con el coeficiente λ estimado a partir de las varianzas de los i-replicados (i = 1..N, con N = 30, 50, 100), tal y como se muestra en las ecuaciones (7-9); Regresión de Deming (P4), con el coeficiente λ estimado a partir de las varianzas de las series correspondientes, según las ecuaciones (10-11); y la Regresión de Deming (P5), con el coeficiente λ con los errores analíticos eran iguales: λ = 1.

$$\sigma^2_{xi} = 0.5 * (x_{1i} - x_{2i})^2 \tag{7}$$

$$\sigma^2_{yi} = 0.5 * (y_{1i} - y_{2i})^2 \tag{8}$$

$$\lambda = \frac{\sum_i \sigma^2_{yi}}{\sum_i \sigma^2_{xi}} \tag{9}$$

$$\sigma^2_X = \left. \sum_i \overline{X}^2 - \left(\sum_i \overline{X}\right)^2 \right/ N \tag{10}$$

Siendo $\overline{X}$: media del i-ésimo duplicado de números (seudo)aleatorios generados con el método A.

$$\sigma^2_Y = \left. \sum_i \overline{Y}^2 - \left(\sum_i \overline{Y}\right)^2 \right/ N \tag{11}$$

Siendo $\overline{Y}$: media del i-ésimo duplicado de números (seudo)aleatorios generados con el método B.

$$\lambda = \frac{\sigma_y^2}{\sigma_x^2} \qquad (12)$$

En el caso de que $\lambda < 1$, entonces $\lambda = 1/\lambda$.

El flujograma seguido en el presente trabajo se muestra en la Figura 1. El algoritmo se repitió 500 veces. Al final de cada iteración del algoritmo se estimaron la pendiente y el intercepto de las correspondientes rectas de comparación de métodos, tal y como se ha descrito previamente.[3-4,7,10] También se estimó el error de estimación de la pendiente obtenida mediante los procederes P1, P3, P4 y P5. En el caso del proceder P1, el error de estimación de la pendiente se obtuvo analíticamente, tal y como se ha descrito previamente.[3,7,10]. En el caso de la regresión Deming (P3, P4, P5), el error de estimación de la pendiente se obtuvo mediante técnicas no paramétricas.[11]

Indicadores del desempeño del procedimiento estadístico.

Al final de cada ensayo de simulación se calcularon los siguientes indicadores del desempeño del procedimiento considerado como solución del problema de comparación de métodos:

1) <u>Promedio de las pendientes simuladas</u>:

$$\bar{b} = \frac{1}{W} \sum_i b_i$$

Donde W = 500.

Si el proceder devuelve estimados insesgados de la pendiente, entonces $\bar{b}$ no debe ser estadísticamente diferente de la unidad, bajo la hipótesis nula $H_o: \beta = 1$.

2) <u>Raíz del error cuadrado medio del promedio de las pendientes</u>:

$$\text{RECM} = \sqrt{\frac{\sum_i \left(b_i - \bar{b}\right)^2}{W - 1}}.$$

La raíz del error cuadrado medio del promedio de las pendientes equivale a la desviación estándar de la distribución de las w-ésimas pendientes (W = 1..500) obtenidas al final de cada ensayo de simulación.

3) Error estándar real de la pendiente:

$$\text{ES}(\beta) = \sqrt{\frac{\sum_i \left(b_i - 1\right)^2}{W - 1}}$$

El error estándar real de la pendiente fija la dispersión de la distribución de las w-ésimas pendientes si se asume que $\bar{b} = 1$. Si el proceder en cuestión devuelve estimados insesgados de la pendiente, entonces RCEM y ES(β) deben coincidir.

4) <u>Razón de rechazos f para la hipótesis nula $H_o: \beta = 1$</u>:

$$f = \frac{\text{rechazos observados}}{\text{rechazos esperados}}$$

$$f = \frac{\text{rechazos observados}}{W * \alpha}$$

$$f = \frac{\text{rechazos observados}}{25}$$

Si el proceder devuelve estimados insesgados de la pendiente, entonces la razón $f = 1$.

Figura 1. Diagrama de flujo seguido en la simulación estadístico-matemática.

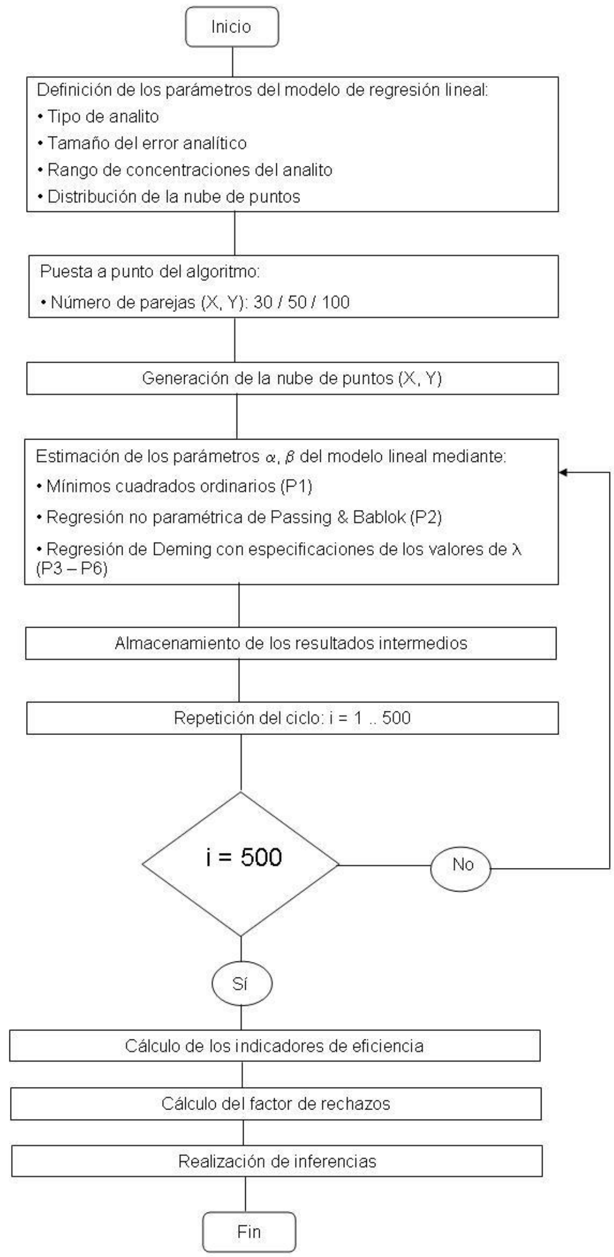
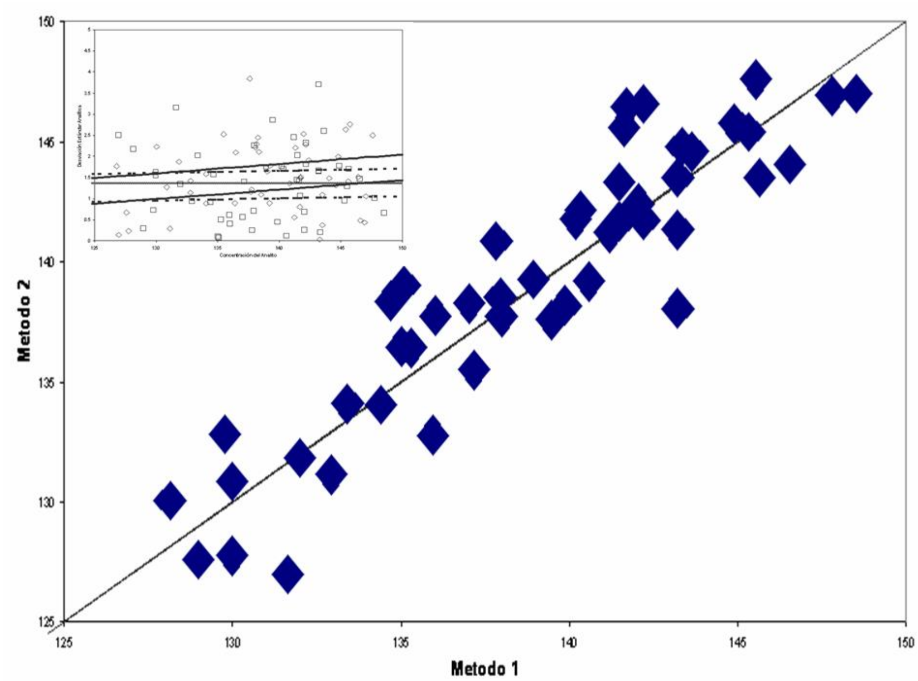


Figura 1. Caso del Sodio: Resultados obtenidos en una iteración del ensayo de simulación. Línea continua (———): Línea de identidad Y = X. Recuadro. Error-en-la-Respuesta vs. Concentración-del-Analito. Línea central: Y = 1.355 (Valor-diana). Línea discontinua (-------): Intervalo de confianza al 99% propio de la recta de regresión Error-en-la-Respuesta vs. Concentración del analito para el método 1 (□). Línea gruesa (———): Intervalo de confianza al 99% propio de la recta de regresión Error-en-la-Respuesta vs. Concentración del analito para el método 2 (◊).



La dócima de los resultados obtenidos al final del ensayo de simulación se realizó mediante las técnicas estadísticas pertinentes.[10] Se fijó un valor crítico del error permisible del 5% en la realización de las pruebas de hipótesis.

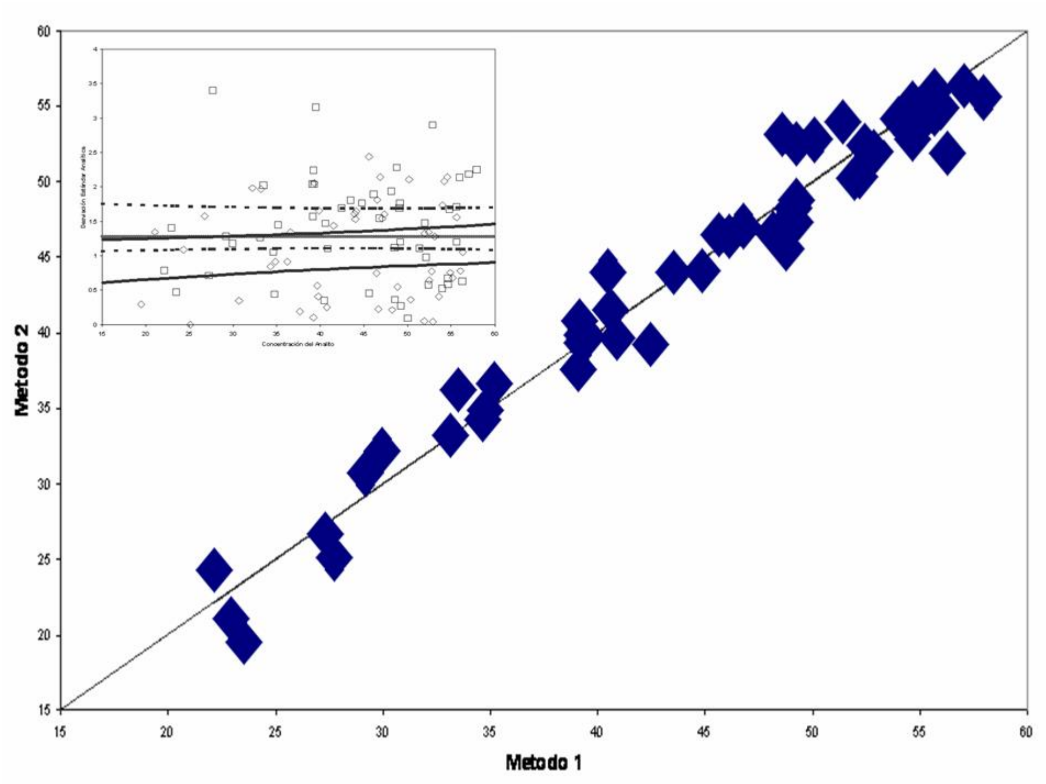
El desempeño del procedimiento estadístico se calificó según los criterios siguientes:

Calificación	Estimado insesgado de la pendiente	Razón de rechazos $f \leq 1$
Adecuado	Sí	Sí
Inadecuado	No	No
Dudoso	Sí	No
	No	Sí

RESULTADOS

En las Figuras 1 – 3 se presentan los diagramas de dispersión de los datos obtenidos para cada caso clínico en una iteración del ensayo de simulación, junto con los correspondientes gráficos del Error-en-la-Respuesta vs. Concentración-del-Analito. En cada caso se pudo satisfacer la constancia del error del método analítico en el rango de valores simulados de la concentración del analito, y la inclusión del valor-diana del error del método analítico dentro del intervalo de confianza al 95% de la recta Error-en-la-Respuesta vs. Concentración-del-Analito.

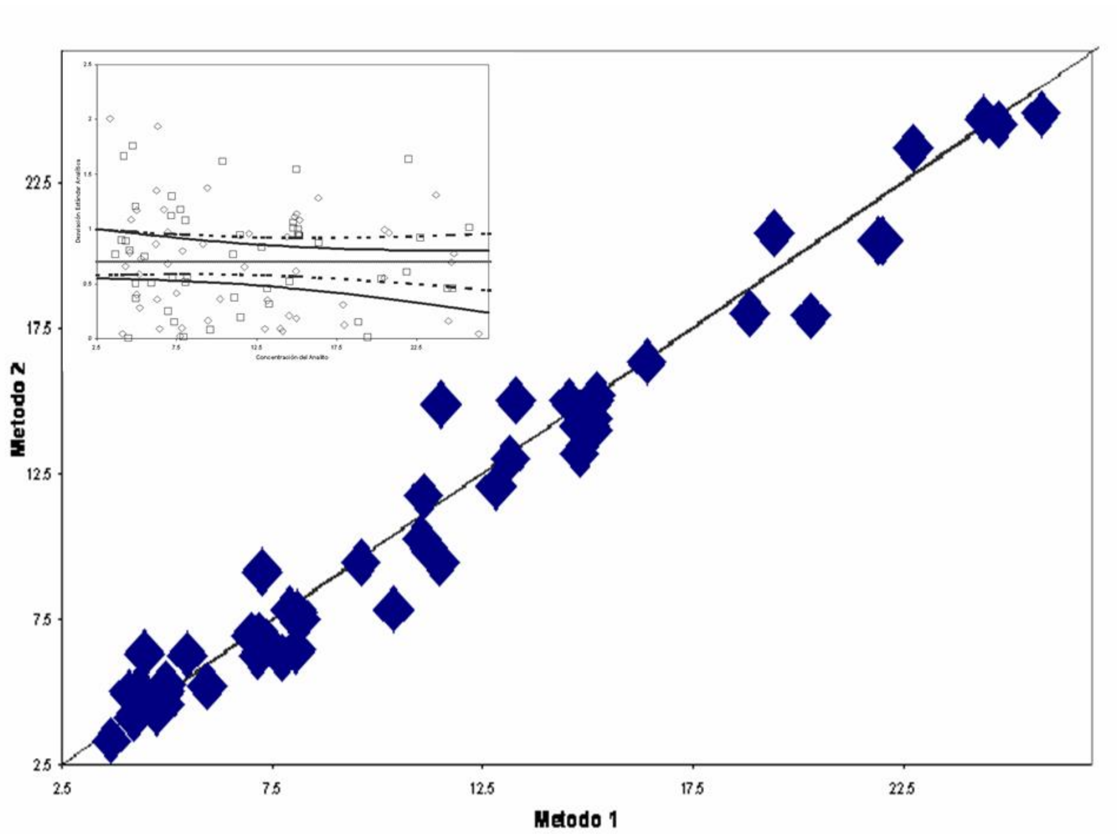
Figura 2. Caso de la Albúmina: Resultados obtenidos en una iteración del ensayo de simulación. Línea continua (———): Línea de identidad Y = X. Recuadro. Error-en-la-Respuesta vs. Concentración-del-Analito. Línea central: Y = 1.275 (Valor-diana). Línea discontinua (-------): Intervalo de confianza al 99% propio de la recta de regresión Error-en-la-Respuesta vs. Concentración del analito para el método 1 (□). Línea gruesa (———): Intervalo de confianza al 99% propio de la recta de regresión Error-en-la-Respuesta vs. Concentración del analito para el método 2 (◊).



En la Tabla 1 se muestra el promedio de los estimados de las pendientes obtenidas con cada uno de los procederes (P1 - P5) después de cada iteración del algoritmo de simulación. Es de notar que el proceder P1 devolvió siempre estimados sesgados de la pendiente, e inferiores a la unidad. El proceder P3 devolvió estimados sesgados de la pendiente en 3 ocasiones, y el proceder P4 en una sola ocasión. El proceder P2 no devolvió estimados sesgados de la pendiente.

En la Tabla 2 se muestran los valores de los errores de estimación de la pendiente devueltos por los procederes P1, P3, P4 y P5. Es de notar que los valores del RECM obtenidos con el proceder P1 fueron siempre superiores en 1.5 – 2.0 veces que los del error estándar real de la pendiente. Para los procederes P3 - P5 los valores del error de estimación de la pendiente fueron virtualmente idénticos, no importa el método empleado en su obtención.

Figura 3. Caso de la Glucosa: Resultados obtenidos en una iteración del ensayo de simulación. Línea continua (———): Línea de identidad Y = X. Recuadro. Error-en-la-Respuesta vs. Concentración-del-Analito. Línea central: Y = 0.7 (Valor-diana). Línea discontinua (-------): Intervalo de confianza al 99% propio de la recta de regresión Error-en-la-Respuesta vs. Concentración del analito para el método 1 (□). Línea gruesa (———): Intervalo de confianza al 99% propio de la recta de regresión Error-en-la-Respuesta vs. Concentración del analito para el método 2 (◊).



Aunque por definición el proceder P2 no proporciona un error de estimación de la pendiente, de todas maneras se presentan los valores del error real de estimación de la pendiente y el RECM, a título de comparación con sus contrapartidas paramétricas.

En la Tabla 3 se muestran los valores del factor *f* de rechazos obtenidos con cada uno de los procederes (P1 - P5). Es de notar que el proceder P1 se caracterizó por una elevada frecuencia de rechazos de la hipótesis nula H_o : $\beta = 1$, bajo el presupuesto Y = X. Sin embargo, los otros procederes no justificaron las expectativas que podrían derivarse de la literatura consultada y del análisis de sus respectivos fundamentos teóricos. Para estos procederes, el factor de rechazos *f* fue ≤ 1 sólo en la mitad aproximadamente de los casos.

Finalmente, en la Tabla 4 se presenta la utilidad de cada proceder validado en este estudio, de acuerdo a los resultados de los ensayos de simulación estadístico-matemático.

Tabla 1. Estimados de las pendientes de los procederes P1 – P5 validados en el presente estudio. Se presenta el promedio de las W = 500 pendientes obtenidas en cada iteración del ensayo de simulación. Para más detalles, consulte la Sección "Material y Métodos" de este trabajo.

Proceder	Tamaño	Analito		
	Muestral	Sodio	Albúmina	Glucosa
P1	30	0.9439*	0.9813*	0.9874*
	50	0.9374*	0.9816*	0.9833*
	100	0.9413*	0.9811*	0.9844*
P2	30	1.0054	1.0014	1.0022
	50	0.9953	1.0010	0.9991
	100	1.0002	0.9997	1.0003
P3	30	0.9970	0.9997	1.0009
	50	0.9911*	0.9996	0.9971*
	100	0.9966*	0.9991	0.9995
P4	30	1.0005	1.0013	1.0022
	50	0.9940*	1.0009	0.9981
	100	0.9986	1.0000	1.0002
P5	30	1.0044	1.0015	1.0024
	50	0.9964	1.0011	0.9986
	100	1.0003	1.0002	1.0004

* Significativamente diferente de 1 ($p < 0.05$)

El proceder P1 fue calificado como Inadecuado en todos los casos, al devolver siempre estimados sesgados de la pendiente, junto con un factor de rechazos intolerablemente alto. Sin embargo, los restantes procederes exhibieron un comportamiento irregular.

De modo general, se puede concluir que la aplicabilidad de cualquiera de los procederes P3 – P5 cuando el rango analítico de trabajo es estrecho (el caso del Sodio) es Dudosa (estimados insesgados de la pendiente unidos a un elevado factor de rechazos), cuando no Inadecuada. En el caso de la Glucosa, la aplicabilidad de los procederes P3 – P5 fue Dudosa. Curiosamente, en el caso de la Albúmina, la aplicabilidad de los procederes P3 – P5 fue Adecuada.

DISCUSIÓN

Con este estudio se pretende evaluar sistemática e integradamente el rendimiento de 3 procederes considerados tradicionalmente como soluciones del problema de comparación de métodos: la Regresión de Mínimos Cuadrados Ordinarios, la Regresión de Passing y Bablok, y la Regresión Deming, ésta última en tres implementaciones diferentes, según la estrategia utilizada en la construcción del estadígrafo λ: a partir de las varianzas intrarreplicadas, a partir de las varianzas de las series generadas de valores, y con valores predeclarados.

Se definieron 3 especificaciones de calidad que debía satisfacer el proceder para ser considerado como una solución adecuada del problema de comparación de métodos: 1) que el estimado de la pendiente de la recta de comparación de métodos sea estadísticamente igual a la unidad; 2) que el error de estimación de la pendiente sea el menor posible, y 3) que el número de rechazos de la hipótesis nula $H_o : \beta = 1$ sea igual (o menor) de un valor nominal del 5%.

El modelo teórico que ha sustentado la simulación de las series establece que los valores "observados" del analito: 1) se muestrean independientemente de una distribución gaussiana del mismo (7), ó 2) se

Tabla 2. Errores de estimación de las pendientes de los procederes (P1 - P5) validados en el presente estudio. Se presenta el promedio de las W = 500 iteraciones del ensayo de simulación. Promedio: Promedio de los valores de los errores de estimación de las pendientes obtenidas en cada iteración. RECM: Desviación estándar de la distribución de las pendientes obtenidas al final del ensayo. ES(β): Error estándar real de la distribución de las pendientes. Para más detalles, consulte la Sección "Material y Métodos" de este trabajo.

Proceder	Tamaño Muestral	Analito								
		Sodio			Albúmina			Glucosa		
		Promedio	RECM	ES(β)	Promedio	RECM	ES(β)	Promedio	RECM	ES(β)
P1	30	0.0637	0.0851	0.0640	0.0372	0.0408	0.0363	0.0322	0.0347	0.0324
	50	0.0484	0.0779	0.0465	0.0281	0.0325	0.0268	0.0302	0.0302	0.0252
	100	0.0340	0.0684	0.0353	0.0195	0.0265	0.0187	0.0238	0.0238	0.0161
P2	30		0.0717	0.0717		0.0402	0.0402		0.0372	0.0371
	50		0.0494	0.0492		0.0299	0.0299		0.0272	0.0272
	100		0.0381	0.0381		0.0204	0.0204		0.0196	0.0196
P3	30	0.0691	0.0708	0.0708	0.0397	0.0378	0.0378	0.0335	0.0347	0.0347
	50	0.0520	0.0516	0.0507	0.0293	0.0278	0.0278	0.0268	0.0268	0.0266
	100	0.0362	0.0378	0.0376	0.0201	0.0195	0.0194	0.0187	0.0187	0.0186
P4	30	0.0694	0.0699	0.0699	0.0398	0.0378	0.0377	0.0335	0.0347	0.0347
	50	0.0522	0.0509	0.0506	0.0293	0.0276	0.0276	0.0266	0.0266	0.0265
	100	0.0363	0.0379	0.0378	0.0201	0.0194	0.0194	0.0187	0.0187	0.0187
P5	30	0.0696	0.0699	0.0698	0.0398	0.0376	0.0376	0.0335	0.0336	0.0335
	50	0.0523	0.0503	0.0502	0.0293	0.0277	0.0277	0.0259	0.0259	0.0258
	100	0.0364	0.0377	0.0377	0.0201	0.0193	0.0193	0.0185	0.0185	0.0185

El rendimiento de estos procederes (y otros) se ha evaluado históricamente mediante ensayos de simulación estadístico-matemática. En tales ensayos, se generan N (habitualmente $N \geq 500$) series de valores bajo la hipótesis nula $H_o : X = Y$ (esto es, de la equivalencia de los métodos en comparación), para después derivar estadígrafos que describan el rendimiento del proceder. Se puede concluir entonces que el proceder es una solución adecuada (o no) si los estadígrafos derivados satisfacen las especificaciones declaradas al inicio del ensayo.

distribuyen uniforme e independientemente en el rango de las concentraciones de interés (12). Sin embargo, el modelo presentado en este trabajo difiere de los presentados anteriormente en que se asume que los resultados 'generados' con cada método analítico son en realidad replicados de una variable que se distribuye uniformemente en el rango de interés de distribución del analito. Esta prescripción del modelo permite asegurar que las raíces cuadradas de las varianzas de los replicados se distribuyan alrededor del valor-diana de la imprecisión del método analítico, y que se cumpla la constancia del error de estimación del método analítico en el rango analítico de trabajo, algo que no se había documentado en los trabajos que sobre este tema se han publicado antes.

Los otros tres procederes se distinguen precisamente porque devuelven (casi) siempre estimados insesgados de la pendiente de la recta de comparación de métodos. En las ocasiones en que alguno de los procederes devolvió un estimado sesgado de la pendiente, la diferencia entre los valores real y estimado podría denotarse como significativa estadísticamente, pero sin ninguna influencia analítica. Hay que tener en cuenta que estos procederes se caracterizan por un error de estimación de la pendiente bastante pequeño, y por lo tanto, diferencias entre los valores real y estimado del orden de las centésimas pueden ser denotadas como significativas estadísticamente, pero sin ninguna repercusión en los objetivos del experimento analítico (¿qué significa, analíticamente hablando, que la pendiente fuera de 0.999 o de 0.995?). Esto refuerza, una vez más, el precepto de que es más importante que el analista establezca las especificaciones de

Tabla 3. Factor de rechazos. Se presenta la frecuencia de rechazos de la hipótesis nula $H_o : \beta = 1$ observada con cada proceder, expresada como un cociente respecto de la frecuencia nominal de rechazos, para un coeficiente de confianza estadística del 5%. Para más detalles, consulte la Sección “Material y Métodos” de este artículo.

Proceder	Tamaño Muestral	Analito		
		Sodio	Albúmina	Glucosa
P1	30	2.84	1.56	1.64
	50	4.92	1.60	2.08
	100	7.92	3.12	2.76
P2	30	1.28	0.68	1.12
	50	0.68	0.92	1.16
	100	1.20	0.76	0.46
P3	30	1.00	0.92	1.36
	50	1.24	1.04	1.36
	100	1.40	0.88	1.08
P4	30	0.92	0.96	1.36
	50	1.20	1.04	1.32
	100	1.32	0.92	1.16
P5	30	0.88	1.00	1.24
	50	1.12	1.04	1.24
	100	1.20	0.88	1.04

Este trabajo confirma que la Regresión de Mínimos Cuadrados no es una solución adecuada al problema de comparación de métodos, coincidiendo así con resultados citados anteriormente, porque devuelve siempre estimados sesgados negativamente de la pendiente de la recta de comparación de métodos, y su uso resulta en una elevada frecuencia de rechazos de la hipótesis nula $H_o : \beta = 1$. La razón para ello parece ser la inclusión de un componente sistemático en el error de estimación de la pendiente, como puede concluirse de la comparación de la raíz del error cuadrado medio de las pendientes, el error real de estimación de las pendientes, y el promedio de los errores de estimación de las pendientes obtenidas en cada iteración del algoritmo de simulación estadístico-matemático.

calidad del método analítico como paso previo de la etapa de validación, y no descanse ciegamente en las inferencias estadísticas para dictaminar sobre el rendimiento del método.

hubiera conducido a una mayor tasa de rechazos de la hipótesis nula H_o : $\beta = 1$.

Tabla 4. Utilidad del proceder. La utilidad del proceder se estableció en virtud del sesgo incurrido en la estimación de la pendiente de la recta de comparación de métodos y la frecuencia de rechazos de la hipótesis nula H_o : $\beta = 1$. Para más detalles, consulte la Sección "Material y Métodos" de este artículo.

Proceder	Tamaño muestral	Analito		
		Sodio	Albúmina	Glucosa
P1	30	Inadecuada	Inadecuada	Inadecuada
	50	Inadecuada	Inadecuada	Inadecuada
	100	Inadecuada	Inadecuada	Inadecuada
P2	30	Dudosa	Adecuada	Dudosa
	50	Adecuada	Adecuada	Adecuada
	100	Dudosa	Adecuada	Dudosa
P3	30	Adecuada	Adecuada	Dudosa
	50	Inadecuada	Adecuada	Inadecuada
	100	Inadecuada	Adecuada	Adecuada
P4	30	Adecuada	Adecuada	Adecuada
	50	Inadecuada	Adecuada	Dudosa
	100	Dudosa	Adecuada	Dudosa
P5	30	Adecuada	Adecuada	Dudosa
	50	Dudosa	Adecuada	Dudosa
	100	Dudosa	Adecuada	Adecuada

Otras conclusiones son obvias de este estudio, y una de ellas ataca un apotegma muy común entre los analistas cuando se trata de "mejorar" el rendimiento de un método analítico mediante el incremento del número de observaciones. En efecto, con cualquiera de los procederes validados en este estudio, el incremento del número de observaciones no se tradujo en una mejoría de los estimados de las pendientes, ni en una disminución de la frecuencia de rechazos de la hipótesis nula H_o : $\beta = 1$. Si ello hubiera sido así, los mejores resultados se hubieron obtenido con las series de 100 observaciones, lo que no es el caso. Sí es de notarse que un aumento en el número de observaciones redunda en un error de estimación de la pendiente cada vez más pequeño, pero ello, contraproducentemente,

Llama la atención el buen desempeño de una forma axiomática de implementación de la Regresión Deming. Algunos autores han recomendado que, en ausencia de estimados fehacientes de las varianzas intrarreplicados para la construcción del coeficiente λ (como fue descrito originalmente), éste se construya a partir de las varianzas de las series de valores. Este estudio valida este axioma, siempre y cuando se asegure que las varianzas de los métodos analíticos en comparación sean similares, y por lo tanto, se prevea un coeficiente λ unitario.

Sin embargo, fue decepcionante comprobar que en más de la mitad de las ocasiones, estos tres procederes resultaron en una frecuencia elevada de rechazos de la hipótesis nula H_o : $\beta = 1$ (aún cuando se asegurara que $Y = X$) en contraste con estudios anteriores.[4,7] Los trabajos citados en la literatura consultada presentan realmente la tasa de rechazos de la hipótesis nula $H_o : X = Y$, que se computa del número de veces en que la pendiente estimada es diferente de uno; el intercepto es diferente de cero; o una concurrencia de ambos eventos. Por el contrario, en este trabajo se computa el número de veces que la pendiente estimada fue diferente de uno.

Ahora bien, cuando se examinan los trabajos de los autores que validaron originalmente estos procederes, se puede constatar que, aunque teóricamente se preveían tasas unitarias de rechazo de la hipótesis nula $H_o : X = Y$, los resultados obtenidos realmente no permitieron sostener esta previsión. Esta circunstancia valida, de cierta manera, los hallazgos presentados en este trabajo. Por otro lado, las tasas observados de rechazo de la hipótesis nula $H_o : \beta = 1$ no fueron mayores de 2 (como sí ocurrió en el caso de la Regresión de Mínimos Cuadrados). De hecho, una tasa de rechazo igual o menor de 2 implicaría un nivel de significación del 10% (en lugar del 5% prefijado). Por consiguiente, el criterio de la frecuencia de rechazos de la hipótesis, aun cuando se utilizó en este estudio para juzgar el rendimiento del proceder, no debe desalentar al analista en la selección de un proceder estadístico para el tratamiento del problema de la comparación de métodos.

Finalmente, cuando se juzga el rendimiento del proceder estadístico según el sesgo del estimado de la pendiente y la tasa de rechazo de la hipótesis nula $H_o : \beta = 1$, se observó que la Regresión de Mínimos Cuadrados no fue una solución adecuada del problema de comparación de métodos en ninguno de los escenarios definidos; y la utilidad de cualquiera de los tres procederes restantes es Dudosa, sino Inadecuada, cuando el rango de concentraciones del analito es estrecho (el caso del Sodio), Adecuada cuando el rango de concentraciones del analito es intermedio (el caso de la Albúmina); y entre Adecuada y Dudosa para rangos de concentraciones extendidos (el caso de la Glucosa).

No constituyó un objetivo de este estudio evaluar las posibles dependencias entre el rango analítico de concentraciones del método analítico y el rendimiento del proceder estadístico particular. Desafortunadamente, y como se ha señalado anteriormente, tampoco existen trabajos previos que permitan explicar satisfactoriamente los resultados obtenidos. Sin embargo, y en virtud de la evidencia recomendada, se puede recomendar que el tratamiento del problema de comparación de métodos cuando se trata de analitos con un rango de concentraciones igual o menor de 1.5 (Límite Superior:Límite Inferior) debe contemplar el uso de estrategias no paramétricas como la de Bland y Altman.[13] Si a pesar de todo, fuera necesario proveer sendos estimados de la pendiente y el intercepto en escenarios con un rango reducido de concentraciones del analito, éstos deberían asumirse con mucha cautela.

El estudio reseñado en este artículo interesa especialmente la calidad de la estimación de la pendiente de la recta de comparación de métodos. En la experiencia de los autores, a los analistas les inquieta particularmente que el proceder estadístico que se aplique en la solución de este problema no falle en devolver un valor unitario de la pendiente, si éste existe, y relegan el estimado del intercepto, no importa que éste sea diferente de cero. Para ellos, es más fácil de aceptar un intercepto significativamente diferente de cero, que una pendiente inferior a la unidad.

CONCLUSIONES

Se ha evaluado el rendimiento de varios procederes estadísticos promovidos como soluciones del problema de comparación de métodos. De acuerdo con los resultados presentados en este estudio, la Regresión de Mínimos Cuadrados debería rechazarse como solución, debido a que resulta en estimados sesgados de la pendiente de la recta de comparación de métodos y una elevada tasa de frecuencia de rechazos de la hipótesis nula $H_o : \beta = 1$. Los otros procederes considerados (Regresión de Passing y Bablok, Regresión de Deming en cualquiera de tres implementaciones posibles) se caracterizaron por devolver estimados insesgados de la pendiente, pero a costa de tasas variables de rechazo de la hipótesis nula $H_o : \beta = 1$. Deben considerarse otras estrategias no paramétricas (como la de Bland y Altman) cuando el rango de las concentraciones es estrecho, y no es obligatorio mostrar estimados de la pendiente y el intercepto de la recta de comparación de métodos.

ADDENDUM

En el tiempo transcurrido desde la publicación de este artículo, han aparecido nuevas soluciones estadísticas del problema de la comparación de métodos.[14-15] Sin embargo, los desarrollos alcanzados no han encontrado salida en programas informáticos para la aplicación de la solución estadística tenida como la más efectiva. Por otro lado, las soluciones propuestas demandan de un conocimiento estadístico íntimo del problema, circunstancia que puede lastrar la difusión de las mismas. Se anticipa entonces que el problema de la comparación de métodos analíticos seguirá ocupando la atención de analistas, informáticos y matemáticos por igual en los años venideros, en particular, ante las cada vez más exigentes normas de los organismos evaluadores de tecnologías biomédicas.

SUMMARY

The performance of 4 solutions to the methods comparison problem: Ordinary Least Squares Regression (P1), Passing-Bablok Regression (P2), Deming Regression with the coefficient λ estimated from the variances of replicate measurements (P3) and the Deming Regression with the coefficient λ estimated from the variances of each serie of observations (P4), when the corresponding analytical errors can be assumed to be of equal size and constant throughout the analytical range of interest, is discussed in this article. The performance of the Deming Regression with coefficient λ equal to 1 was also evaluated (P5). According to the range of the observations, three different analytical scenarios were constructed: narrow (the Sodium case), intermediate (the Albumin case), and extended (the Glucose case) ranges. The performance of each procedure was documented from series of 30, 50 and 100 pairs of numbers (pseudo)randomly generated with a statistical simulation algorithm written in VisualBasic®© for EXCEL®© (Microsoft, USA). To be considered as an adequate solution to the methods comparison problem, the procedure must assure that the slope estimate of the regression line be statistically equal to one; the slope estimation error be the lowest possible; and the number of rejections of the null hypothesis $H_o : \beta = 1$ be equal to (or less than) a nominal value of 5%. The P1 procedure returned slopes estimates different from (lower than) one. The slope estimation error of the P1 procedure was distorted by a systematic error. The use of the P1 procedure resulted in an elevated frequency of rejections of the null hypothesis $H_o : \beta = 1$. The three others procedures returned unbiased estimates of the slope of the methods comparison's regression line, though at the cost of an increased rejection rate of the null hypothesis $H_o : \beta = 1$. In spite of this finding, these three procedures can be used as solutions of the methods comparison

problem, instead of the Ordinary Least Squares Regression procedure, when the analytical errors are equal. Regarding the P4 procedure, its use is recommended when there are single measurements from each method, and it can be assumed, from the results of method validation studies, that the analytical errors are of equal size. ***Delgado Ramos A, Ramos Salazar R, Martínez Canalejo H, Santana Porbén S****. Linear regression procedures as solutions of the methods comparison problem. I. Constant and equal analytical errors. RCAN Rev Cubana Aliment Nutr 2010;20(1):152-167. RNP: 221. ISSN: 1561-2929.*

Subject headings: Methods comparison / Methods simulation / VisualBasic / Least-squares / Passing-Bablok / Deming.

REFERENCIAS BIBLIOGRÁFICAS

1. Hollis S. Analysis of method comparison studies [Editorial]. Ann Clin Biochem 1996;33:1-4.
2. Westgard JO, Hunt MR. Use and interpretation of common statistical tests in method comparison studies. Clin Chem 1973;19:49-57.
3. Draper NR, Smith H. Applied Regression Analysis. New York: John Wiley & Sons, 1966:44-103.
4. Passing H, Bablok W. A new biometrical procedure for testing the equality of measurements from two different analytical methods. Aplication of linear regression procedures for method comparison studies in Clinical Chemistry. Part I. J Clin Chem Clin Biochem 1983;21:709-20.
5. Deming WE. Statistical adjustment of data. New York: Wiley, 1943:184.
6. MacTaggart DL, Farwell SO. Analytical use of linear regression. Part I: Statistical errors in both variables. J AOAC International 1992;75:608-14.
7. Linnet K. Evaluation of regression procedures for method comparison studies. Clin Chem 1993;39:424-32.
8. Payne RP. Method comparison: evaluation of least squares, Deming and Passing/Bablok regression procedures using computer simulation. Ann Clin Biochem 1997;34:319-20.
9. Ramos Salazar R, Delgado Ramos A, Martínez Canalejo H, Santana Porbén S. Procederes de regresión lineal como soluciones al problema de la comparación de métodos. II. Errores analíticos constantes pero diferentes. Rev Mex Patol Clin 2001;48:223-32.
10. Martínez Canalejo H, Santana Porbén S. Manual de Procedimientos Bioestadísticos. La Habana: Editorial de Ciencias Médicas, 1990.
11. Linnet K. Estimation of the linear relationship between the measurements of the methods with proportional error. Stat Med 1990;9:1463-73.
12. Hartmann C, Massart DL. Display methods for visual comparison of the results of two measurement methods. J AOAC International 1994;77:1318-25.
13. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. The Lancet 1986;**i**:307-10.
14. Martin RF. General Deming regression for estimating systematic bias and its confidence interval in method-comparison studies. Clin Chem 2000;46: 100-4.
15. Ludbrook J. Statistical techniques for comparing measurers and methods of measurement: a critical review. Clin Exp Pharmacol Physiol 2002;29:527-36.