

La necesidad de mejores evaluaciones metodológicas y nuestra exigencia ante los criterios de aceptabilidad de los resultados

F. López Azorín

Sr. Director:

En 1986 una revisión de artículos publicados sobre evaluaciones de métodos e instrumentos analíticos utilizados en Química Clínica y Hematología, detectó numerosos defectos en su realización experimental y en las pruebas estadísticas utilizadas, alertando sobre la necesidad de mejorar las evaluaciones (1). Evaluar los métodos exige una sobrecarga de trabajo importante que ahora puede abordarse mejor gracias a los avances producidos en automatización e informatización, pero que obliga a diseñar los experimentos de evaluación y a decidir las pruebas estadísticas más adecuadas a utilizar. Una revisión sobre los métodos utilizados y propuestos nos puede ayudar a unificar criterios con el objetivo de que los métodos sean evaluados y ello se realice de la mejor manera posible.

La evaluación de un método persigue detectar los errores aleatorios y sistemáticos que origina, para cuantificarlos mediante cálculos estadísticos y determinar si son inferiores a los límites permisibles que no comprometen a la utilización clínica de los resultados. Estos límites, establecidos en función de la variabilidad biológica (2), pueden consultarse tabulados para la mayoría de analitos (3). Ellos nos dicen hasta dónde podemos equivocarnos. La evaluación nos permite conocer el error que cometemos en nuestras condiciones de trabajo y valorar modificaciones del procedimiento ó futuros cambios por otros métodos. La existencia de esos límites nos obliga a intentar conocer nuestro error con la mayor exactitud posible. Las exigencias legales y nuestra propia responsabilidad profesional nos deben impedir aceptar la incertidumbre de desconocer el error que estamos cometiendo ó incluso dudar si nuestros métodos son aceptables.

En el esquema experimental de evaluación propuesto por el equipo de Westgard en 1978 y todavía aceptado (4), la detección de errores aleatorios mediante experimentos de replicación intraserial e interserial, así como la detección de errores sistemáticos constantes o proporcionales mediante series de interferencia o de recuperación, respectivamente en una primera fase, no plantea problemas especiales en su ejecución ni en su interpretación estadística. Si los errores sistemáticos detectados están en un límite aceptable, la evaluación se debe completar con un experimento de comparación de métodos frente a los resultados obtenidos con un método de referencia, surgiendo en esta fase los mayores problemas. La utilización de la estadística en los experimentos de comparación de métodos es, posiblemente, la parte más difícil del análisis estadístico en los estudios de evaluación de métodos. La pregunta a contestar es ¿Qué método estadístico es el que comete menos error para calcular el error que comete un método analítico?

En los últimos años se ha producido una discusión importante sobre el tema.

Inicialmente el método más utilizado era el cálculo de la regresión lineal por mínimos cuadrados, con la exigencia de linealidad de datos, pequeños errores aleatorios, intervalo suficientemente amplio de concentraciones y ausencia de valores aberrantes («outliers») (5). Este cálculo de la regresión también exige que la variable independiente «x» no tenga error, es decir que el método de referencia utilizado sea exacto, y esto en la mayoría de ocasiones no es posible, por lo que los resultados del método de comparación también presentarán errores. En el caso de que presenten errores aleatorios ha sido propuesto como más aceptable el método de regresión ortogonal establecido por Deming en 1943 (6), incluso considerado adecuado si existen valores aberrantes (7). Si el error que presenta esta variable es proporcional, cuando la varianza de los resultados aumenta con la concentración, ha sido considerado adecuado el cálculo de la correlación lineal mediante el cálculo de mínimos cuadrados ponderado (8).

En los años ochenta, para evitar los inconvenientes del método simple de mínimos cuadrados, Passing y Bablok desarrollan el cálculo del método de regresión no paramétrica, que fue defendido como «método de elección debido a su robustez» y a permitir «obtener una estimación no sesgada de la recta de regresión» (9), siendo ampliamente utilizado gracias a la publicación de sus instrucciones en BASIC y la distribución del programa en soporte magnético. Con él parecía que teníamos resuelto el problema, sin embargo no debía ser así cuando en 1991 Zaydman et al., en su estudio sobre modelos de regresión lineal para comparación de métodos, rechazan el método de Passing-Bablok porque no cumple la exigencia de que la razón de varianzas en «y» y «x» iguale el cuadrado de la pendiente, y rechazan el método de regresión lineal simple debido a que no cumple dos requisitos: ausencia de error en la variable independiente «x» y que el error en la variable dependiente y debe ser homogéneo en todo el intervalo de concentraciones. Obtienen los mejores resultados con el método de mínimos cuadrados ponderado cuando los errores en la variable independiente «x» son pequeños frente al intervalo de concentraciones medido (v.g. en Glucosa), mientras que consideran mejor el método de Deming si los errores en «x» son significativos frente al intervalo ensayado, cuando el intervalo de concentraciones es estrecho (v.g. en Potasio) (10).

Considerando que después Bland y Altman proponen su método de evaluación de las diferencias, con el atractivo de su sencillez de cálculo y cuya utilización frente al uso de los métodos de regresión ya hemos comentado en estas páginas (11), podemos resumir de lo expuesto que, en los casos que deba utilizarse el cálculo de la regresión lineal, los métodos que parecen más aceptables son el de mínimos cuadrados ponderado y el método de Deming. Ambos aplican cálculos complejos y requieren programa informático. El paquete informá-

tico SPSS no los contiene en su módulo básico, pero si contiene el de mínimos cuadrados ponderado en el módulo de modelos de regresión. El método de Deming está incluido en el programa «EP Evaluator» desarrollado por D.G. Rhoads (12) y también puede ser solicitada una versión en BASIC a K. Linnet, posiblemente el autor que ha hecho un mayor esfuerzo por su estudio e implantación (13). Ante la posible conveniencia y necesidad de su aplicación, nuestra Sociedad debería plantearse la posibilidad de conseguirlos y facilitar su distribución.

Correspondencia:
F. López Azorín
C/ Tomás Maestre, 1 - 3º-B
30004 - Murcia.

BIBLIOGRAFÍA

1. Hackney JR, Cembroski GS. Need for improved instrument and kit evaluations. *Am J Clin Pathol* 1986; 86: 391-3.
2. Fraser CG. Data on biological variation: Essential prerequisites for introducing new procedures? *Clin Chem* 1994; 40: 1671-73.
3. Sociedad Española de Bioquímica Clínica y Patología Molecular. Comité de garantía de la calidad y acreditación de laboratorios. Comisión de calidad analítica. Especificaciones de la calidad analítica en laboratorios clínicos con distintos niveles de recursos. *Quim Clin* 1998; 19: 219-36.
4. Garber CC, Carey RN. Evaluation of methods. En: Kaplan LA, Pesce AJ *Clinical Chemistry: Theory, Analysis and Correlation*. 3ª ed. Mosby. 1996, 402-23.
5. Westgard JO, Hunt MR. Use and interpretation of common statistical test in method-comparison studies. *Clin Chem* 1973; 19: 44-57.
6. Wakkers PJM, Hallendoorn HBA, Op de Weegh GJ, Heerspink W. Applications of statistics in clinical chemistry. A critical evaluation of regression lines. *Clin Chim Acta* 1975; 64: 173-84.
7. Cornbleet PJ, Gochman N. Incorrect least-squares regression coefficients in method-comparison analysis. *Clin Chem* 1979; 25: 432-38.
8. Garden JS, Mitchell DG, Mills WN. Nonconstant variance regression techniques for calibration-curve-based analysis. *Anal Chem* 1980; 52: 2310-15.
9. Frey E, Fuentes-Arderiu X, Queraltó JM. Comparación estadística de métodos analíticos. *Educación continuada Química Clínica*. 1988; 1: 69-79.
10. Zaydman L, Lasky FD, Van Nostrand C, Doptis PD. Recommended linear regression models for comparison of method in clinical chemistry. [Abstract]. *Clin Chem* 1991, 37: 960.
11. López, F. Utilización adecuada de las pruebas estadísticas en los estudios de comparación de procedimientos. *Quim Clin* 2000; 19: 237.
12. Rhoads DG. Implementation of NCCLS and alternative method-evaluation protocols. *Clin Chem* 1991, 37: 1528-30.
13. Linnet K. Evaluation of regression procedures for method comparison studies. *Clin Chem* 1993; 39: 424-32.