



DOCUMENTO DE CONSENSO

Verificación de la intercambiabilidad de resultados entre equipos duplicados[☆]



Enrique Prada de Medio^{a,b,*}, Raquel Blázquez Sánchez^{a,c},
Carmen Perich Alsina^{a,d}, Gabriela Gutiérrez Bassini^{a,e}, Daniel Pineda Tenor^b,
Ana Isabel Álvarez Ríos^{a,c}, Ángel Salas García^{a,d}, Josep María Jou Turallas^{a,e},
María Ángeles Cuadrado Cenxual^{a,b}, Jorge Moranco Zaragoza^{a,c},
Carmen Ricós Aguilá^{a,d}, María José Cobo del Hoyo^{a,b}, Rafael Calafell Clar^{a,c}
y Francisco Ramón Bauzá^{a,d}

^a Grupo de Trabajo Interdisciplinar de Calidad en el Laboratorio Clínico, Madrid-Barcelona, España

^b Asociación Española de Biopatología Médica, Madrid, España

^c Asociación Española de Farmacéuticos Especialistas, Madrid, España

^d Sociedad Española de Bioquímica Clínica y Patología Molecular (SEQC), Barcelona, España

^e Sociedad Española de Hematología y Hemoterapia, Madrid, España

Recibido el 21 de febrero de 2014; aceptado el 24 de febrero de 2014

Disponible en Internet el 4 de abril de 2014

PALABRAS CLAVE

Equipos duplicados;
Intercambiabilidad
de resultados;
Equipo virtual único

Resumen Una situación bastante común en los laboratorios clínicos es la existencia de distintos analizadores que, usados indistintamente, determinan las mismas pruebas. Debemos asegurar que el grado de concordancia entre los resultados obtenidos sea independiente de los sistemas o plataformas analíticas y que, por tanto, las mediciones sean comparables. Si no podemos garantizarlo, puede crearse confusión en el clínico y una posible mala interpretación de los resultados, viéndose comprometida la seguridad del paciente.

Por este motivo, este grupo de trabajo plantea 2 situaciones diferenciadas: Implantación de un nuevo equipo y revisión de equipos previamente verificados.

En el primer caso se debe realizar un análisis de las diferencias, utilizando la variabilidad biológica (VB) como fuente para establecer la diferencia máxima permitida; en caso de que no existieran datos de VB, se utilizará el criterio estadístico para este cálculo. Una segunda posibilidad sería la utilización de la regresión lineal para el estudio de intercambiabilidad entre analizadores.

[☆] Documento realizado en colaboración con: Comité de Calidad, Gestión, Evidencia y Seguridad de AEBM y con la Comisión de Calidad Analítica de la SEQC.

Este documento es un consenso entre las siguientes comisiones: Comisión de Certificación y Acreditación de AEFA, Comisión de Calidad Analítica de SEQC, Comité de Calidad, Gestión, Seguridad y Evidencia de AEBM y Comité de Estandarización en Hematología de SEHH, fundamentándose en documentación de ámbito público incorporando el concepto de equipo virtual único.

* Autor para correspondencia.

Correo electrónico: eprada@sescam.jccm.es (E. Prada de Medio).

KEYWORDS

Duplicate equipment;
Interchangeability of
results;
Single virtual
equipment

En el caso de equipos ya verificados previamente, se propone la utilización del protocolo del CLSI, que permite la comprobación de hasta 10 analizadores con un máximo de 2 series analíticas.

Si los resultados proporcionados por los distintos analizadores no superan la desviación máxima establecida, se considerarán como un equipo virtual único; en caso contrario, los resultados no serán intercambiables, por lo que el informe de resultados deberá indicar de forma clara el analizador empleado en cada paciente, así como el intervalo de referencia adecuado. © 2014 AEBM, AEFA y SEQC. Publicado por Elsevier España, S.L. Todos los derechos reservados.

Verification of the interchangeability of patient results between duplicate equipment

Abstract A fairly common situation in clinical laboratories is the existence of different analyzers used interchangeably for determining the same tests. It must be ensured that the level of agreement between the results obtained is independent of the systems or analytical platforms used, and that the measurements are comparable. Otherwise, we can create clinician misunderstanding and a possible misinterpretation of the results. Safety of the patient may be at risk.

For this reason, this workgroup established two different situations: implementation of new equipment and review of previously verified equipment. In the first case, an analysis of the differences should be performed using the biological variability (VB) as a means to establish the maximum difference allowed. When VB data are not available, the statistical criterion should be used for this calculation. A second possibility would be the use of linear regression for the study of comparability of analyzers.

The use of the Clinical and Laboratory Standards Institute Protocol (CLSI), which allows up to 10 analyzers to be verified with a maximum of 2 analytical series, is proposed in the case of previously checked equipment.

If the results provided by the different analyzers do not exceed the maximum deviation established, they would be considered as a single virtual equipment, otherwise, the results are not interchangeable, and the laboratory report should indicate clearly the analyzer used in each patient, as well as the appropriate reference range.

© 2014 AEBM, AEFA y SEQC. Published by Elsevier España, S.L. All rights reserved.

Introducción

La existencia de distintos analizadores que, determinando las mismas pruebas, están trabajando de forma simultánea, es una situación bastante común en los laboratorios clínicos de nuestro país. Además, los pacientes a lo largo de su historia clínica, ya sea en diferentes episodios o bien dentro del mismo evento médico, pueden requerir atención desde la primera asistencia en un servicio de urgencias, ingreso con posibilidad de traslado a una unidad de cuidados críticos que disponga de pruebas a la cabecera del paciente y seguimiento posterior en atención primaria tras el alta de la hospitalización. Esto implica que las pruebas de laboratorio solicitadas pueden haber sido realizadas desde múltiples localizaciones y por diferentes equipos de medida.

Los datos proporcionados por el laboratorio clínico tienen un fuerte impacto en la atención sanitaria, por lo que la trazabilidad y equivalencia de los resultados proporcionados debe ser considerada primordial por parte del profesional del laboratorio, porque en el caso de que estas propiedades no se cumplieran, podemos crear confusión y una posible mala interpretación de los resultados por parte del clínico, con una repercusión directa e inmediata sobre el paciente, afectando a su

seguridad e incluso provocando un evento adverso sobre el mismo.

Además, según la norma ISO 15189:2012 (apartado 5.6.6) para la acreditación de los laboratorios clínicos, el asegurar la veracidad de los procedimientos de medida es requisito imprescindible¹.

Por todo lo expuesto anteriormente, este grupo de trabajo, formado por profesionales de las comisiones o comités de calidad de 4 sociedades del laboratorio clínico de ámbito nacional (AEBM, AEFA, SEHH y SEQC), plantea 2 situaciones diferenciadas en el abordaje de los estudios de transferibilidad de resultados cuando existen distintos analizadores que realizan las mismas pruebas en un laboratorio:

- Implantación de un *nuevo equipo* en el laboratorio, para complementar el equipamiento actual. El nuevo equipo puede ser un mismo modelo de analizador, usando los mismos métodos de ensayo, y calibrado con los mismos materiales, o un analizador distinto y/o métodos de ensayo y/o características técnicas y/o procesos de calibración distintos.
- Verificación de la intercambiabilidad de los resultados de equipos *ya implantados* en el laboratorio.

Si los resultados proporcionados por los distintos analizadores no superan la desviación máxima establecida, se considera que dichos resultados son intercambiables y que los equipos comparados pueden ser considerados como un *equipo virtual único*.

En caso contrario, los resultados no son intercambiables, por lo que el informe de resultados debe indicar de forma clara el analizador empleado en cada paciente, así como el intervalo de referencia adecuado.

Este documento de consenso pretende servir como guía y herramienta para los laboratorios al afrontar este tipo de escenario de cara a su práctica clínica habitual, así como de cara a un posible proceso de implantación de un sistema de gestión de la calidad.

Objeto y campo de aplicación

El objeto de este documento es presentar un protocolo que asegure la intercambiabilidad de los resultados obtenidos cuando un laboratorio utiliza varios instrumentos para medir una misma magnitud biológica en muestras de pacientes, y por tanto, informados de forma independiente del método o analizador empleado (*equipo virtual único*)².

Este documento aplica a los instrumentos del laboratorio clínico con técnicas cuyos resultados son cuantitativos y continuos.

Definiciones

Intercambiabilidad: el término *comparability*, tal y como lo define el *Clinical and Laboratory Standards Institute* (CLSI), expresa la concordancia entre resultados obtenidos para una magnitud, en un mismo paciente, utilizando diferentes procedimientos de medida dentro de un sistema de salud. Se considera que los resultados son comparables si las diferencias entre ellos no exceden un valor crítico establecido basado en criterios de aceptación definidos³.

Al no existir el término comparabilidad en el Diccionario de la Real Academia de la Lengua Española, se utiliza el término «intercambiabilidad».

Equipo virtual único: conjunto de instrumentos que cumplen los criterios de intercambiabilidad que los capacita para su empleo de forma compartida, indistinta o alternativa, para determinar una misma magnitud biológica en un mismo paciente y el mismo tipo de muestra.

Procedimiento

El presente grupo de trabajo propone 2 procedimientos según las situaciones descritas anteriormente, que se esquematizan en el anexo 1:

Implantación de un nuevo analizador en el laboratorio clínico

Se debe establecer una metodología rigurosa, tal como describe el CLSI⁴, que consiste en los siguientes pasos.

Selección de muestras

Se deben preparar varias muestras para cubrir un intervalo de concentraciones de interés clínico, en el cual sea válido el concepto de equipo virtual único.

En relación con el material de las muestras, se debe aplicar el siguiente orden de prioridad:

- Muestras de pacientes recién obtenidas, recogidas en el contenedor adecuado y conservadas según los requisitos de estabilidad de las magnitudes a determinar.
- Materiales de control conmutables. Permitirían comparar diversos sistemas analíticos, asumiendo que los resultados obtenidos serán aplicables a los especímenes humanos.
- Materiales de control sin conmutabilidad conocida. Pueden considerarse adecuados solo si se cumplen las siguientes condiciones: utilización de instrumentos idénticos adaptados al mismo método analítico, mismos lotes de reactivos, mismo lote de calibrador y mismo lote de material control.

El número mínimo de muestras a considerar, siempre que sea posible, es de 40, ensayadas a lo largo de como mínimo 5 series analíticas⁵. La selección de un mayor tamaño de muestra mejora la fiabilidad del proceso. Las muestras procederán de pacientes distintos y deben cubrir el intervalo de medición del analizador. Es importante considerar que al menos el 50% de los valores muestrales representen concentraciones representativas del espectro de enfermedades a diagnosticar o, en su defecto, fuera del intervalo de referencia propuesto por el laboratorio.

Se debe identificar la presencia de muestras con interferentes conocidos para alguno de los procedimientos de medida comparados y constatar si dicha interferencia afecta a uno solo de los equipos o a todos.

Las muestras deben ser procesadas de acuerdo con los procedimientos habituales del laboratorio. En el caso de que el almacenaje sea inevitable, su duración y características dependerán de la estabilidad de la magnitud que se debe ensayar⁴. Todas las alícuotas deben almacenarse en condiciones similares para evitar introducir variables ajenas a la medición.

Mediciones

El personal responsable de efectuar el ensayo debe estar familiarizado con los procedimientos de medida a evaluar.

Las medidas deben ser realizadas por duplicado, empleando 2 alícuotas diferentes por muestra y ensayadas éstas en 2 series analíticas distintas (en total, 4 alícuotas por muestra, 2 para cada equipo a evaluar); por tanto, el número mínimo de alícuotas será de 160.

Con objeto de evitar efectos de arrastre, el orden en el que se analizan las alícuotas ha de invertirse. Es decir, las muestras deben ser ensayadas con el siguiente orden: 1, 2, 3, 4. N para las primeras alícuotas y N...4, 3, 2, 1 para el duplicado. Esta metodología de inversión ha de ser seguida en ambos equipos, aunque la secuencia inicial de muestras no tiene que ser necesariamente la misma.

Es recomendable realizar el análisis a lo largo de un mínimo de 5 días, siendo lo ideal emplear un total de 20. De esta forma, se minimizan los posibles errores sistemáticos que pueden derivarse de una única serie analítica.

Para evitar sesgos por deterioro de la muestra, se recomienda procesar todas las muestras de forma simultánea en los 2 analizadores, o con un máximo de 2 h de diferencia entre los 2 métodos, en función de la estabilidad de las magnitudes⁴.

Evaluación de errores y valores aberrantes

Se deben identificar posibles resultados erróneos y valores aberrantes:

- Se aceptarán los resultados si se cumplen las reglas establecidas en el control de calidad aplicado en la rutina del laboratorio⁶.
- Eliminar los valores aberrantes (anexo 2). El análisis de comparación debe ser realizado con todos los puntos, eliminando en todo caso los valores aberrantes (*outliers*), mediante la observación de todos los puntos en las gráficas descritas en el apartado siguiente. Sin embargo, la aplicación de los procedimientos descritos en el anexo 2 dota de un mayor rigor a la detección y eliminación de dichos valores aberrantes.

El máximo de puntos aberrantes admisibles para la comparación es del 2,5% de los datos. En el caso de que se identifiquen más de un 2,5% de valores aberrantes, es necesario investigar las posibles interferencias, errores humanos, mal función de los equipos y controles de calidad. En el caso de que el problema no pueda ser identificado de forma inequívoca y resuelto satisfactoriamente, se debe iniciar el procedimiento de nuevo, recopilando nuevas muestras.

Análisis de datos

El análisis de datos se puede realizar siguiendo 2 metodologías diferentes, en orden de prioridad:

- Análisis de las diferencias.
- Análisis de regresión lineal.

Análisis de las diferencias. Permite determinar la presencia de sesgo, tanto constante como proporcional, entre las mediciones de los equipos a comparar. Para ello se emplea una representación gráfica denominada diagrama de diferencias, también conocido como diagrama de Bland-Altman, en el que se representan:

- Eje de ordenadas: diferencia (D_i) o diferencia relativa porcentual (D'_i), según se esté valorando sesgo constante o proporcional entre las mediciones obtenidas, empleando ambos métodos para cada una de las muestras.

Por convención, se considera al método x como el de referencia:

$$D_i = Y_i - X_i$$

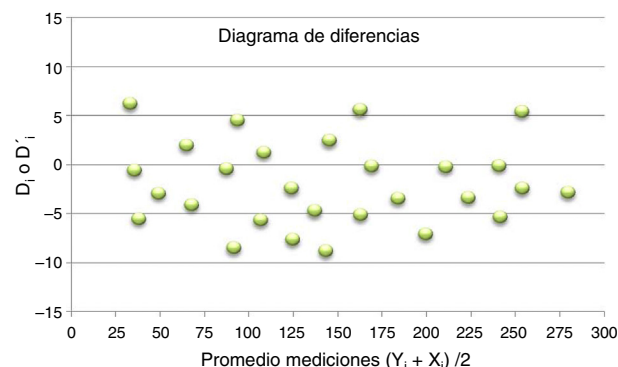
$$D'_i = 100 \frac{Y_i - X_i}{(Y_i + X_i)/2}$$

Siendo Y_i el promedio del resultado de las alícuotas de la muestra «i» en el analizador Y.

Siendo X_i el promedio del resultado de las alícuotas de la muestra «i» en el analizador x.

- Eje de abscisas: promedio de cada pareja de resultados (promedio de las alícuotas) para cada una de las muestras:

$$\frac{(Y_i + X_i)}{2}$$



La interpretación de los resultados se realiza comparando la media de las diferencias, expresada en forma absoluta (D_m) o relativa (D'_m), con respecto a la desviación máxima (D) admisible.

Esta debería establecerse inicialmente con criterio biológico y, si no es posible, se aplicará el criterio estadístico.

- Criterio biológico, en que se utiliza la media relativa de las diferencias (D'_m).

Debido a que en la mayoría de los casos es difícil asegurar que no hay error sistemático, siempre que la imprecisión de los 2 instrumentos sea la deseable ($0,5 CV_i$)⁷, es recomendable que la especificación para error sistemático de cada instrumento, expresada en términos de porcentaje, se mantenga dentro del límite deseable o mínimo^{8,9}:

$$D'_m < 0,25 (CV_i^2 + CV_G^2)^{1/2} \text{ (desviación deseable)}$$

$$D'_m < 0,375 (CV_i^2 + CV_G^2)^{1/2} \text{ (desviación mínima)}$$

Donde, CV_i : coeficiente de variación intraindividual, CV_G : coeficiente de variación grupal.

En el anexo 3 se ilustra la importancia de los dígitos significativos al utilizar estas fórmulas.

En aquellos casos en los que la media de las diferencias sea menor que la desviación máxima (D), se considera que los resultados son intercambiables y ambos equipos pueden ser considerados como un equipo virtual único.

- Criterio estadístico, se aplica para las magnitudes sin datos de variabilidad biológica (VB) conocidos.

Se tienen en cuenta la media de las diferencias (D_m o D'_m) y la desviación estándar de estas medias (S_D o $S_{D'}$). De esta forma, se considera que el intervalo de confianza del 95% de la media de las diferencias será:

- Para las diferencias absolutas: $D_m \pm 1,96S_D$.
- Para las diferencias relativas: $D'_m \pm 1,96S_{D'}$.

En aquellos casos en los que el intervalo de confianza para la media de las diferencias incluya el valor cero, se considera que no existen diferencias estadísticamente significativas entre los resultados proporcionados por ambos equipos ($p \geq 0,05$), por lo que los resultados son intercambiables y ambos equipos pueden ser considerados como un equipo virtual único.

Si el intervalo de confianza de la media de las diferencias no incluye al valor cero, no podemos constatar la ausencia de diferencias estadísticamente significativas, por lo que los resultados de ambos equipos no son intercambiables y la consideración de equipo virtual único no puede ser aplicada. En estos casos, el método comparado presenta un error sistemático (ES) o sesgo del tipo:

- Constante: el intervalo de confianza de las diferencias absolutas no incluye el valor cero.
- Proporcional: el intervalo de confianza de las diferencias relativas no incluye el valor cero.

Regresión lineal. Antes de realizar el análisis de regresión, es necesario calcular el coeficiente de correlación, que permite decidir si el análisis de regresión se puede o no efectuar. El más empleado es el coeficiente de correlación de Pearson (r), que asume una distribución normal de los datos. Si bien r puede tomar valores comprendidos entre -1 y $+1$, en la comparación entre resultados solo tienen sentido los valores positivos. Un $r = 1$ indica una correlación perfecta entre ambos analizadores ($Y = x$).

Si $r < 0,975$, el tamaño muestral debería ampliarse incorporando muestras adicionales. En caso de que, tras realizar la citada ampliación de la muestra, los valores de r no alcancen los niveles requeridos, se considera que los resultados aportados por ambos equipos no son intercambiables.

Si $r \geq 0,975$, se puede efectuar el análisis de regresión.

El análisis de regresión establece la forma en la que los valores de Y (considerado este el equipo a evaluar) varían en función de los valores de x (considerado por convención el equipo de referencia).

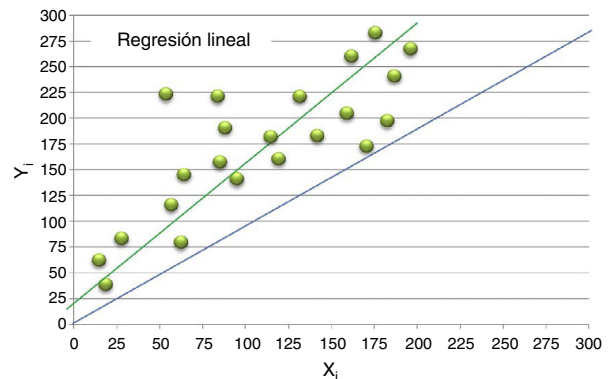
Hay que tener en cuenta que como resultado se utiliza el promedio de ambas alicuotas.

Se asume que los resultados aportados por ambos equipos pueden ser representados gráficamente en un eje de coordenadas y se relacionan en función la siguiente ecuación:

$$Y = bX + a$$

Donde « b » es la pendiente de la recta de regresión y « a » el valor de la ordenada en el origen (corte de la recta en el eje de ordenadas). Mediante el análisis de regresión lineal es posible obtener los valores de « a » y « b », así como los intervalos de confianza del 95% para cada uno de ellos. Existen diferentes aproximaciones para su estimación: regresión lineal simple, regresión ortogonal o método de Deming, regresión no paramétrica de Passing-Bablok. Si bien se considera que el último es el más consistente desde un punto de vista estadístico, la práctica diaria indica que los valores aportados por cualquiera de estos métodos son muy similares, no alterando la decisión de intercambiabilidad de resultados.

Dado que los cálculos necesarios para la estimación de la pendiente, el punto de corte en ordenadas y, sobre todo, de los intervalos de confianza son complejos y tediosos, se recomienda el uso de una aplicación informática para su estimación. En cualquier caso, las ecuaciones necesarias pueden consultarse en el documento EP09-A2⁴ del CLSI.



Interpretación: la recta de regresión ideal es aquella que presenta una pendiente « b » igual a 1 y una ordenada en el origen « a » igual a cero. De esta forma, la ecuación obtenida ($Y = X$) implica una identidad perfecta entre los resultados.

La observación directa de la recta obtenida y su comparación con la recta ideal permite detectar los puntos aberrantes, la dispersión de los valores y la existencia de sesgos y tendencias. Sin embargo, para un análisis más exhaustivo, se han de considerar los intervalos de confianza del 95% obtenidos para la pendiente y la ordenada en el origen.

En el caso en el que los intervalos de confianza contengan el valor 0 para la ordenada en el origen « a » y el valor 1 para la pendiente « b », se considera que no existen diferencias estadísticamente significativas ($p \geq 0,05$) entre ambos equipos, por lo que los resultados proporcionados por ambos son intercambiables y podemos aplicar el concepto de *equipo virtual único*.

Cuando el intervalo de confianza para la pendiente « b » no contiene el valor 1, el equipo Y presenta un error sistemático proporcional con respecto al equipo de referencia x . Si el valor de $b > 1$, los resultados aportados por el analizador Y serán superiores a los del x . Si el valor de $b < 1$, los resultados del analizador Y serán inferiores a los del x .

Cuando el intervalo de confianza para la ordenada en el origen « a » no contiene el valor 0, el equipo Y presenta un error sistemático constante con respecto al equipo de referencia X . Los valores aportados por el analizador Y presentan una diferencia constante con respecto al equipo X , que será superior en caso de $a > 0$ o inferior en caso de $a < 0$.

Verificación de la intercambiabilidad de los resultados de equipos ya implantados en el laboratorio

Este grupo de trabajo considera que, una vez realizado el estudio estadístico completo del apartado «Implantación de un nuevo analizador en el laboratorio clínico», el laboratorio debe verificar si la intercambiabilidad se mantiene a lo largo del tiempo. Para ello, propone utilizar el protocolo del CLSI³.

Es un protocolo sencillo de utilizar, ya que utiliza ya que utiliza un máximo de 2 series de pacientes para comparar

hasta 10 instrumentos. Incluye una tabla para determinar el número de replicados de cada muestra en función del número de instrumentos a comparar y de los CV medios interinstrumento obtenidos, que permite detectar la diferencia entre resultados con una potencia del 80% y un nivel $\alpha = 0,05$.

La concentración de las muestras de pacientes debe ser lo más próxima posible a los niveles de decisión clínica. Se puede calcular a partir del valor medio de los resultados de los 2 niveles de control, por los 2 sistemas de medida $\pm 20\%$.

La intercambiabilidad entre instrumentos se establece tal como se ha explicado en el apartado 4.1.4.1: calculando la media de las diferencias, expresada en forma absoluta (D_m) o relativa (D'_m), y comparándola con la desviación admisible (D), establecida con criterio biológico. Si este no es posible, se aplica el criterio estadístico.

Se puede seguir un ejemplo de aplicación de este protocolo en el anexo 3 disponible como material adicional (sólo accesible desde la versión electrónica de este artículo).

Frecuencia

La verificación de resultados comparables entre sistemas de medida alternativos debe hacerse cuando se modifique el protocolo o procedimiento de trabajo por alguna de las siguientes circunstancias:

- Fallo persistente del control interno de la calidad. Cuando se desvíen los resultados del mismo material de control utilizado en los sistemas de medida alternativos, puede ser necesario verificar que los resultados son intercambiables utilizando muestras de pacientes.
- Fallo persistente del programa de garantía externa de la calidad o prueba de aptitud. Resultados de la participación en programas de garantía externa de la calidad fuera de los objetivos del laboratorio, para los 2 instrumentos compartidos.

La frecuencia de verificación de intercambiabilidad puede variar en función de la estabilidad de los sistemas de medida, reactivos y materiales de calibración y control. El laboratorio, en sus revisiones del sistema de gestión de la calidad, debe incluir la verificación de equipos duplicados como uno de los aspectos claves del sistema.

Otras características, como la rotación del personal, la estabilidad de stocks de lotes, el coste, etc., deben ser considerados por el laboratorio.

En cualquier caso, se recomienda la realización de todo el proceso como mínimo cada 2 años.

Consideraciones finales

A pesar de que se trata de una situación muy habitual en los laboratorios, la existencia de equipos por duplicado, triplicado, etc., no está bien resuelta en la bibliografía actual y por eso estas 4 sociedades científicas consideran necesario elaborar este documento consenso para que sirva como instrumento para abordar esta problemática.

La consideración de diferentes analizadores como un equipo virtual único implica que los resultados proporcionados por ambos sean intercambiables e indistinguibles desde el punto de vista clínico. La validación de las técnicas

empleadas y la planificación e implementación de un sistema de gestión de la calidad en el laboratorio constituyen la base en la que se sustenta la fiabilidad de los resultados.

Una situación que puede aparecer tras la evaluación es que los 2 equipos puedan comportarse como equipo virtual único en la mayor parte del intervalo, pero que en determinados valores extremos eso no sea así. La pregunta es: ¿deben definirse como equipos distintos, en lugar de como equipo virtual único, teniendo en cuenta que un porcentaje bajo de resultados, habitualmente menos del 2%, se comportan así?

La actitud más recomendable es definir el equipo virtual único cuando los resultados estén dentro del intervalo de valores en los que se cumple la intercambiabilidad. Para situaciones con valores fuera de este intervalo, las medidas deberían realizarse en un único equipo que se ha asignado como referencia para valores extremos.

En cualquier caso, si tras la evaluación se constata que los 2 equipos no se comportan como equipo virtual único para todo el intervalo de concentraciones, el laboratorio deberá investigar las causas y tomar las medidas correctivas pertinentes.

Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.

Anexos. Material adicional (Anexos 1 y 4)

Se puede consultar material adicional a este artículo en su versión electrónica disponible en <http://dx.doi.org/10.1016/j.labcli.2014.02.0041>.

Anexo 2. Detección de valores aberrantes

Detección de valores aberrantes intraequipo

Una de las ventajas de usar datos duplicados para cada una de las muestras participantes en la comparación de equipos es la de detectar errores aleatorios puntuales. Gracias al siguiente procedimiento, podemos detectar la presencia de valores aberrantes para cada uno de los equipos o métodos a comparar, denominados x (equipo ya implantado) e Y (equipo a evaluar) a partir de este momento.

Evaluar las *diferencias absolutas individuales* entre los duplicados de cada muestra aplicando las siguientes ecuaciones:

$$DX_i = |X_{i1} - X_{i2}|$$

$$DY_i = |Y_{i1} - Y_{i2}|$$

Donde i es igual al número de muestra (desde 1 hasta N , siendo N el número total de muestras).

Evaluar la *diferencia absoluta media* del conjunto de todos los duplicados, mediante el sumatorio de las

diferencias absolutas individuales de cada muestra dividido por el número total de muestras en cada equipo:

$$\overline{DX} = \frac{\sum DX_i}{N}$$

$$\overline{DY} = \frac{\sum DY_i}{N}$$

Definir los *límites de aceptación* para cada equipo (x e Y), multiplicando por 4 la diferencia absoluta media y redondeando al alza en caso necesario. Comparar cada una de las diferencias absolutas individuales con los límites de aceptación correspondientes.

En el caso de que alguna diferencia absoluta individual exceda los límites de aceptación, se realizará un análisis adicional, pero normalizando los valores. Para ello deben calcularse:

- Las diferencias absolutas individuales normalizadas (o relativas):

$$DX'_i = \frac{|X_{i1} - X_{i2}|}{\bar{X}_i}$$

$$DY'_i = \frac{|Y_{i1} - Y_{i2}|}{\bar{Y}_i}$$

- La diferencia absoluta media normalizada:

$$\overline{DX'} = \frac{\sum DX'_i}{N}$$

$$\overline{DY'} = \frac{\sum DY'_i}{N}$$

- Los *límites de aceptación normalizados*, multiplicando por 4 la diferencia absoluta media normalizada para cada equipo (x e Y).

Interpretación: en el caso de que *ninguna diferencia absoluta individual exceda los límites de aceptación*, consideraremos que no existen valores aberrantes intraequipo. El análisis normalizado no es necesario. Aquellos puntos que superen el límite de aceptación en el test absoluto, pero no los superen en el normalizado, serán considerados válidos. En el caso de que un único punto exceda tanto los *límites de aceptación* como los *límites de aceptación normalizados*, se debe investigar la causa y eliminar en cualquier caso esa medición. En el caso de que más de un punto exceda ambos límites de aceptación, debe realizarse una investigación exhaustiva:

- Si el problema es inequívocamente determinado y se tiene la certeza de que solo ha afectado a las muestras que excedieron los límites, se procede a la sustitución de las citadas muestras.
- Si el problema es inequívocamente determinado, pero no podemos asegurar que se limita a las muestras que

excedieron los límites, el proceso debe iniciarse desde el principio, obteniendo nuevas muestras.

- Si el problema no puede ser identificado, interrumpir el proceso y contactar con la casa comercial.

Detección de valores aberrantes interequipos La detección de valores aberrantes entre los diferentes equipos o métodos (x e Y) puede ser realizada de forma intuitiva al observar los gráficos de comparación de métodos que describiremos en los apartados siguientes. Sin embargo, al igual que en el análisis intraequipo, podemos realizar un estudio más riguroso siguiendo el siguiente procedimiento estadístico:

- Evaluar las *diferencias absolutas* entre los métodos. En este caso, es necesario definir un método de referencia. Como ya se ha comentado, consideraremos al x como el método de referencia:

$$E_{ij} = |Y_{ij} - \bar{X}_i|$$

Donde i es igual al número de muestra (desde 1 hasta N, siendo N el número total de muestras) y j la alícuota empleada (duplicado 1 o 2 de cada muestra). Nótese que el valor perteneciente al método de referencia (x) es la media de ambas alícuotas.

- Calcular la *diferencia absoluta media* entre métodos:

$$\bar{E} = \frac{1}{2N} \sum_i^N \sum_j^2 E_{ij}$$

- Definir el valor *límite del test* (LT_E), multiplicando por 4 la diferencia absoluta medias y redondeando al alza cuando sea necesario. Comparar cada E_{ij} con el LT_E.

En el caso de que alguna diferencia absoluta exceda el LT_E, se realizará un análisis adicional similar al anterior, pero normalizando los valores. Para ello deben calcularse:

- Las diferencias absolutas normalizadas (o relativas) entre métodos:

$$\bar{E}'_{ij} = \frac{|Y_{ij} - \bar{X}_i|}{\bar{X}_i}$$

- La diferencia absoluta media normalizada entre métodos:

$$\bar{E}' = \frac{1}{2N} \sum_i^N \sum_j^2 \bar{E}'_{ij}$$

El *límite del test normalizado* se obtiene multiplicando por 4 la diferencia absoluta media normalizada entre

métodos. Comparar con los valores de diferencias absolutas normalizadas entre métodos.

Interpretación: en el caso de que la diferencia absoluta media no exceda los límites de aceptación, consideraremos que no existen valores aberrantes interequipo. El análisis normalizado no es necesario.

Aquellos puntos que superen el límite de aceptación en el test absoluto, pero no los superen en el normalizado, serán considerados válidos.

Aquel o aquellos punto/s que no superen ambos test estadísticos se considera un valor aberrante y debe ser eliminado del análisis.

Anexo 3. Determinación de los dígitos significativos

Es importante tener en cuenta el número de posiciones decimales utilizado en la expresión de los resultados de una magnitud al calcular la diferencia máxima admisible (D). Esta debe adecuarse al número de dígitos significativos, ya que en el redondeo puede suceder que, al comparar resultados entre sí, la menor diferencia que se pueda encontrar entre 2 instrumentos sea superior a la especificación deseable.

Un ejemplo claro ocurre con HbA1c:

- Si se utiliza un solo decimal para expresar los resultados:

Sistema 1: HbA1c = 6,1%.

Instrumento 2: HbA1c = 6,0%.

La diferencia entre los 2 instrumentos = $(6,1 - 6,0) \times 100/6,05 = 1,65\%$, que supera la DM deseable (1,12%).

- Sin embargo, si los resultados se expresan con 2 decimales:

Instrumento 1: HbA1c = 6,06%.

Instrumento 2: HbA1c = 6,04%.

La diferencia entre los 2 instrumentos = $(6,06 - 6,04) \times 100/6,05 = 0,33\%$, que está perfectamente dentro de la especificación deseable (1,12%).

Bibliografía

1. Asociación Española de Normalización. Laboratorios clínicos. Requisitos particulares relativos a la calidad y a la competencia. UNE-EN-ISO 15189. Madrid: AENOR; 2012.
2. AEBM. Comité de Calidad, Gestión, Seguridad y Evidencia. Métodos estadísticos en la comparación de equipos de laboratorio [consultado 19 Feb 2014]. Disponible en: <http://www.aebm.org/documentos-y-publicaciones/documentos-de-la-aebm.html>
3. Clinical, Laboratory Standards Institute. Verification of comparability of patient results within one health care system. Proposed guideline. Document CLSI EP31-A-IR. Pennsylvania. 2012.
4. Clinical, Laboratory Standards Institute. Method comparison and bias estimation using patients samples; approved guideline. 2nd ed. Document EP9-A2. CLSI. Pennsylvania. 2002.
5. SEQC. Comisión de Metrología. Documento 2010-K. Recomendaciones para el estudio de la veracidad en el laboratorio clínico mediante la comparación de procedimientos de medida. SEQC. Barcelona, 2010 [consultado 19 Feb 2011]. Disponible en: .
6. Perich C, Álvarez AI, Blázquez R, Calafell R, Cobo MJ, Cuadrado MA, et al. Aplicación práctica del control interno de la calidad en los procedimientos de medida cuantitativos. Rev Lab Clin. 2014. En prensa.
7. Petersen PH, Fraser CG, Westgard JO, Larsen ML. Analytical goal-setting for monitoring patients when two analytical methods are used. Clin Chem. 1992;38:2256-60.
8. Kenny D, Fraser CG, Hyltoft Petersen P, Kallner A. Consensus agreement. En: Hyltoft Petersen P, Fraser CG, Kallner A, Kenny D, editors. Strategies to set global analytical quality specifications in laboratory medicine. Scan J Clin Lab Invest. 1999;59,7:585.
9. Minchinela J, Ricós C, Perich C, Fernández-Calle P, Alvarez V, Doménech MV, et al. Base de datos de los componentes de Variación Biológica, con las especificaciones de la calidad analítica (deseable, mínima y óptima). Actualización del año 2014 [consultado 19 Feb 2014]. Disponible en: http://www.seqc.es/es/Comisiones/18/9/102/Base_de_datos_de_Variacion_biologica_%7CComision_de_Calidad_Analitica_%7CComite_Cientifico/