



Salud Pública de México

ISSN: 0036-3634

spm@insp.mx

Instituto Nacional de Salud Pública
México

Gortari, Eduardo de; Herrera, Manuel; Loría, Alvar; Terrés, Arturo; González, Marco Antonio;
Hernández, María Antonieta

Programa piloto en laboratorios clínicos mexicanos. III. Estrategia para evaluar la calidad de los
resultados

Salud Pública de México, vol. 36, núm. 5, septiembre-octubre, 1994, pp. 484-491

Instituto Nacional de Salud Pública
Cuernavaca, México

Disponible en: <http://www.redalyc.org/articulo.oa?id=10636504>

- ▶ Cómo citar el artículo
- ▶ Número completo
- ▶ Más información del artículo
- ▶ Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica

Red de Revistas Científicas de América Latina, el Caribe, España y Portugal
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

PROGRAMA PILOTO EN LABORATORIOS CLÍNICOS MEXICANOS. III. ESTRATEGIA PARA EVALUAR LA CALIDAD DE LOS RESULTADOS

EDUARDO DE GORTARI, M.C.,⁽¹⁾ MANUEL HERRERA, M.C.,⁽¹⁾ ALVAR LORÍA, Q.B.P.,⁽²⁾
ARTURO TERRÉS, M.C.,⁽³⁾ MARCO ANTONIO GONZÁLEZ-SALAYANDIA, M.C.,⁽¹⁾
MARÍA ANTONIETA HERNÁNDEZ, Q.F.B.⁽¹⁾

De Gortari E, Herrera M, Loría A,
Terrés A, González-Salayandia MA,
Hernández MA.

Programa piloto en laboratorios
clínicos mexicanos. III. Estrategia
para evaluar la calidad de los resultados.
Salud Pública Mex 1994;36:484-491.

RESUMEN

Objetivo: explorar la validez de una estrategia sencilla y de retroalimentación rápida que discrimina precisión y exactitud de mediciones de laboratorios. **Participantes:** el grupo original de 18 laboratorios. **Métodos:** se colectaron sueros y resultados de mediciones de cuatro componentes séricos (glucosa, urea, creatinina y úrico) en la rutina del día en que los participantes recibieron una visita de verificación de la Dirección General de Regulación de Servicios de Salud de la Secretaría de Salud. Los sueros fueron transportados el mismo día a un laboratorio de referencia (LR) para repetir a la brevedad posible y en forma ciega, las mediciones del participante. El LR utilizó un sistema automatizado que, bajo un programa interno con control diario y uno externo de prestigio reconocido (programa internacional de la Organización Mundial de la Salud), mostró excelente precisión y buena exactitud durante los 200 días del estudio. **Estrategia de análisis:** la estrategia permite discriminar entre precisión y exactitud usando el coeficiente de correlación (CORR) para precisión, y las relaciones

De Gortari E, Herrera M, Loría A,
Terrés A, González-Salayandia MA,
Hernández MA.
Pilot studies in Mexican clinical
laboratories. III. Strategy
to evaluate quality of results.
Salud Pública Mex 1994;36:484-491.

ABSTRACT

Objective: to explore the validity of a simple and rapid strategy which discriminates precision and accuracy of assays done in sera. **Participants:** the same group of 18 laboratories. **Methods:** the sera and results of glucose, urea, creatinine and urates assayed in the routine work of a single day were collected by auditors of an agency of the Ministry of Health. Six laboratories used autoanalyzers and 12 manual kits in the assays. The sera were transported that same day (stored in ice) to a reference laboratory (REF LAB) to repeat the assays as soon as possible. The REF LAB used a spectrum analyzer which showed, throughout the seven months of the study, excellent precision in its internal program and good accuracy in the international EQA program of the World Health Organization. **Strategy of analysis:** precision and accuracy were independently evaluated with two correlates per analyte and participant: the coefficient of correlation of paired results (REF LAB/participant) as a correlate of precision, and the mean of ratios of paired results (participant/REF LAB) as a correlate of accuracy. A correlation below 0.70 and a mean

(1) Dirección General de Regulación de Servicios de Salud, Secretaría de Salud, México.

(2) Instituto Nacional de la Nutrición Salvador Zubirán, México.

(3) Hospital The American British Cowdray, México.

Fecha de recibido: 11 de enero de 1994 Fecha de aprobado: 23 de agosto de 1994

participante-referencia (REL) para exactitud; se consideró inaceptable un CORR abajo de 0.70 y una REL fuera de 85 a 115 por ciento (el participante sistemáticamente se alejó más del 15 por ciento de los valores del LR). Resultados: hubo 62 sistemas analizables: la imprecisión fue baja ya que hubo datos inaceptables sólo en el 8 por ciento de los sistemas (5 de 62), lo cual es menor al 30 por ciento observado en laboratorios mexicanos. La inexactitud fue más frecuente ya que fue inaceptable en el 44 por ciento de los sistemas (28 de 62). Conclusiones: se considera que esta estrategia permite discriminar entre precisión y exactitud. La baja frecuencia de imprecisión podría indicar que la estrategia sufrió un efecto del observador sobre el observado (la visita de verificación fue anunciada). La alta tasa de inexactitud pone de manifiesto la conveniencia de contar, localmente, con una mayor disponibilidad de calibradores y de programas externos con controles confiables en estabilidad y en valores asignados. Ambos son necesarios para laboratorios que no pueden participar en programas internacionales debido a los costos y/o por carecer de las credenciales académicas que algunos programas exigen. Se debe continuar la búsqueda de estrategias de programas externos que discriminen precisión de exactitud, y evitar el uso de índices que no lo hacen.

Palabras clave: control de calidad, sistemas de salud, laboratorio clínico, México

Solicitud de sobretiros: Q.B.P. Alvar Loría. Departamento de Control de Calidad, Instituto Nacional de la Nutrición Salvador Zubirán. Vasco de Quiroga 15, 14000 Tlalpan, México, D.F.

EN ESTE TRABAJO se presentan los resultados de una estrategia para evaluar la calidad de los resultados de mediciones químicas en sueros, realizadas como parte del trabajo rutinario del grupo de 18 laboratorios del Programa Piloto en Laboratorios Clínicos Mexicanos.

MATERIAL Y MÉTODOS

En una visita de verificación de la Dirección General de Regulación de Servicios de Salud (DGRSS) de la Secretaría de Salud, los laboratorios entregaron a los verificadores los sueros a los que se les había medido, ese mismo día, alguno de los cuatro componentes séricos siguientes:

ratio outside of 1 ± 0.15 were considered inadequate. Results: only 62 systems had information (10 either lacked reagents or had no samples that day). Imprecision was observed in only 8 per cent of the analytes (5/62) which is lower than the 30 per cent usually seen in Mexican EQA programs. On the contrary, the frequency of inaccuracy was high, i.e. the ratios were inadequate in 44 per cent of the systems (28/62). Conclusions: our strategy of evaluation was able to discriminate imprecision from inaccuracy. The low frequency of imprecision in the study suggests the possibility of an observer's effect on the observed (date and hour of the audit visit was known in advance). The high rate of inaccuracy highlights the convenience of having a larger local availability of reference materials and EQA programs with reliable control material. The search for indexes discriminating imprecision from inaccuracy should continue for use in laboratories with very little or no experience in EQA.

Key words: quality assessment, health care, clinical laboratory, Mexico

glucosa, urea (o nitrógeno ureico), úrico y creatinina, así como una copia de los resultados de estas mediciones.

Los métodos usados por 12 participantes fueron manuales con *kits* comerciales. Los otros seis participantes usaron sistemas automatizados (dos del Departamento del Distrito Federal, uno de la Secretaría de Salud y los tres privados).

Los sueros se empacaron en hielo y se transportaron ese mismo día al laboratorio de referencia (LR), que fue el Departamento de Control de Calidad del Instituto Nacional de la Nutrición Salvador Zubirán. Los sueros se almacenaron en congelación a -20°C para que en el LR se repitiera, a la brevedad posible, la medición de los mismos componentes séricos que había hecho el par-

LABORATORIOS CLÍNICOS MEXICANOS. III. RESULTADOS

ticipante. En el LR se empleó un sistema automatizado (modelo Spectrum marca Abbott) para repetir las mediciones en forma ciega, ya que el operador ignoraba los datos del participante.

PRECISIÓN Y EXACTITUD DEL APARATO DE REFERENCIA DURANTE EL ESTUDIO

Desde septiembre de 1990 hasta marzo de 1991 el aparato estuvo bajo un programa interno de control de precisión en el que diariamente se midieron los cuatro componentes en un mismo suero control liofilizado, que se reconstituía los jueves de cada semana y se partía en seis alícuotas que se congelaban para medirse del lunes al sábado de la siguiente semana. La precisión en términos del coeficiente de variación en el lapso de estudio para los cuatro componentes se presenta en el cuadro I. La precisión del sistema de referencia fue excelente para tres analitos (3-6% a lo largo de siete meses) y fue buena en urico.

Durante el lapso del estudio el aparato se utilizó también en el programa externo de control de exactitud de la Organización Mundial de la Salud (OMS). En el cuadro II se presentan los resultados de las mediciones hechas cada dos semanas en sueros de la OMS; los datos están en términos del porcentaje del valor asignado por la OMS, de modo que una media del 100 por ciento significa que el promedio del LR es idéntico al promedio de la OMS. En el cuadro II se aprecia que la exactitud interna del LR (*versus* sus propios datos) fue excelente, ya que no mostró diferencias interaños, y la exactitud externa (*versus* los valores asignados por la OMS) fue buena, ya que ningún analito se alejó más del 10 por ciento del valor asignado por

CUADRO I
Precisión (en términos de CV) de las mediciones en el LAB REF durante los 7 meses del estudio

| Componente | N | Media | CV % |
|------------|-----|-----------|------|
| Glucosa | 198 | 111 mg/dL | 3 |
| Urea | 197 | 59 " | 5 |
| Creatinina | 199 | 4.5 " | 6 |
| Urico | 197 | 6.1 " | 10 |

CUADRO II
Exactitud* en las mediciones del LAB REF durante los 7 meses del estudio

| Componente | Periodo | N | Media** |
|------------|--------------|---|---------|
| Glucosa | sep/dic 1990 | 6 | 92 |
| | ene/mar 1991 | 7 | 93 |
| Urea | sep/dic 1990 | 6 | 108 |
| | ene/mar 1991 | 8 | 109 |
| Creatinina | sep/dic 1990 | 5 | 107 |
| | ene/mar 1991 | 8 | 110 |
| Urico | sep/dic 1990 | 6 | 102 |
| | ene/mar 1991 | 6 | 101 |

* En términos de % del valor asignado por la OMS

** Media ideal= 100%

la OMS, si bien cabe notar que midió urea y creatinina ligeramente altas durante el estudio. Al parecer no hubo problema de exactitud absoluta en urico y tampoco en glucosa (con el tiempo, la glucosa baja en sueros almacenados).

ESTRATEGIA DE ANÁLISIS

El diseño del estudio genera para cada participante una serie de pares de resultados por analito y por suero; por ejemplo, dos glucosas del suero X y dos ureas del suero Y (en cada caso una del participante y otra del LR), etcétera. Con las parejas de datos de cada analito se calcularon dos datos de resumen por analito y por participante:

- El coeficiente de correlación (CORR) que se clasificó como aceptable al alcanzar un valor mínimo de 0.70, e inaceptable cuando resultó inferior a dicho valor.
- El promedio porcentual de las relaciones participante/referencia (REL). El resultado del participante se multiplicó por 100 y se dividió entre el resultado correspondiente (mismo suero y mismo componente) del LR; posteriormente se calculó el promedio de estas relaciones para cada componente y participante. Se consideró aceptable la REL si el promedio oscilaba entre 85 y 115 por ciento, e inaceptable si estaba fuera de estos límites.

Consecuentemente, cada participante tenía ocho datos de resumen al medir el CORR y la REL de los cuatro analitos. Las pruebas estadísticas fueron el análisis de correlación¹ y la prueba U de Mann-Whitney.²

RESULTADOS

Hubo un total de 62 sistemas analizables (18 de glucosa, 17 de urea, 18 de creatinina y nueve de úrico). Los sistemas faltantes se debieron a que 10 participantes no midieron el analito el día de la visita, ya sea por carecer de solicitudes o de reactivos. En el cuadro III se muestran los datos de CORR y REL de los 18 participantes.

En lo que respecta al número de datos, hubo 10 sistemas que aportaron menos de cinco datos por analito, 26 con cinco a nueve datos, y 31 que aportaron 10 o más mediciones por analito (cuadro III). Los sistemas de pocos datos son de participantes con una carga de trabajo baja.

Los CORR fueron mayoritariamente aceptables ya que 57 de los 62 sistemas alcanzaron un valor mínimo de 0.70 (cuadro III). De los cinco sistemas con CORR inaceptable (de 0.43 a 0.64), tres son de creatinina, lo cual probablemente se deba a que sus concentraciones en los sueros fueron normales; esto es, que las comparaciones se hicieron a niveles cercanos a 1 mg/dL y, por ende, las discrepancias de décimas de mg/dL se vuelven muy importantes. Asimismo el CORR de creatinina alcanzó un valor casi perfecto (0.90 o más) en sólo 12 de 18 sistemas (67%) pero lo alcanzó en 40 de 44 sistemas (91%) de glucosa, urea y úrico (cuadro III).

Así, los CORR no discriminaron entre los laboratorios participantes ya que en este estudio de un día único, la gran mayoría logró alcanzar niveles de concordancia aceptables *versus* el LR.

Por el contrario, 27 de 62 sistemas (44 por ciento de las veces) tuvieron una REL inaceptable, pues hubo una mayor frecuencia de REL con valores altos (19 sistemas por arriba del 115 por ciento) que bajos (ochos por abajo del 85 por ciento); sin embargo, parecen distribuirse al azar, ya que en los cuatro componentes se dan casos de REL altas y bajas. En 25 de estos 27 escapes, la REL no bajó del 56 por ciento ni sobrepasó el 13 por ciento: las dos excepciones se observaron en glucosa y urea de uno de los laboratorios, que informó concentraciones superiores al doble de las que midió el LR; sus REL fueron de 241 y 248 por ciento en glucosa y urea, respectivamente (cuadro III).

En resumen, a diferencia del CORR, la REL apareció como un índice que permite una partición de los labora-

tarios en función de la proporción de REL aceptables. En el cuadro IV se presenta un ordenamiento de los laboratorios en función de la REL; sólo hubo dos con REL aceptable en todas sus mediciones, y un solo participante con REL inaceptable en todas sus mediciones. Los 15 participantes restantes muestran combinaciones de REL aceptables e inaceptables. De esta manera se dividieron los 18 participantes en ocho niveles, ya que varios alcanzaron los lugares 3, 7 y 13. Si bien hubo seis laboratorios automatizados en los primeros lugares del cuadro IV y fueron estadísticamente superiores a los sistemas manuales (prueba U, $p<0.02$), cabe notar, por otra parte, que hubo dos sistemas automatizados con problemas de inexactitud (ocuparon los lugares 7 y 13 en el cuadro II).

La decisión de emplear una estrategia heterodoxa para calificar y ordenar a los participantes obedece a tres motivos:

1. Las observaciones hechas en el Programa de los Laboratorios Químico-Clínicos de los Institutos Nacionales de Salud (INS) que opera desde 1984.³⁻⁷ En el programa participa un grupo pequeño de laboratorios como un modelo de estudio de detección de problemas de laboratorio en México. Se ha observado que a pesar de su excelencia, los sistemas de medición de los INS presentan problemas de exactitud en forma intermitente. Por lo tanto, se considera que cualquier intento de evaluación debe discriminar si se trata de un sistema con ambos problemas o sólo el de exactitud.
2. La exactitud sólo es posible si se tiene buena precisión. De ahí la importancia de detectar los sistemas con imprecisión para erradicarla como primera medida correctiva.^{8,9}
3. Existen pocos métodos para calificar y/o clasificar los laboratorios químico-clínicos. Uno muy conocido es el índice de variancia (*variance index*) que usa la OMS y que han adoptado los programas externos de aseguramiento de calidad en varios países, entre ellos los dos programas externos mexicanos que operan actualmente.^{10,11}

El índice de la OMS tiene la gran ventaja de calificar la medición en una sola cifra (de cero a 400). Ello permite clasificar al sistema de medición en categorías que van desde la excelencia de una puntuación abajo de 25, hasta las puntuaciones de 200 o más que, según la OMS, ameritan suspender las mediciones y no reanudarlas hasta no haber erradicado los problemas del sistema. En medio

LABORATORIOS CLÍNICOS MEXICANOS. III. RESULTADOS

CUADRO III
Coeficiente de correlación y relación participante/referencia
por participante y por analito (N= No. resultados comparados)

| Participante | Glucosa | | | Urea | | |
|--------------|------------|-------|------|-------|-------|------|
| | N | CORR | REL | N | CORR | REL |
| 01 | 5 | 1 | 67* | 5 | 0.95 | 56* |
| 02 | 8 | 1 | 105 | 8 | 0.97 | 106 |
| 03 | 6 | 0.99 | 96 | 5 | 1 | 102 |
| 04 | 17 | 1 | 91 | 17 | 1 | 87 |
| 05 | 11 | 0.95 | 241* | 9 | 0.58* | 268* |
| 06 | 5 | 0.97 | 94 | 4 | 0.99 | 83* |
| 07 | 3 | 1 | 121* | 3 | 0.87 | 110 |
| 08 | 13 | 0.99 | 136* | 13 | 1 | 100 |
| 09 | 3 | 0.80 | 101 | - | - | - |
| 10 | 12 | 0.97 | 80* | 8 | 0.97 | 102 |
| 11 | 13 | 0.96 | 126* | 11 | 0.90 | 92 |
| 12 | 11 | 0.97 | 100 | 10 | 0.99 | 116* |
| 13 | 13 | 0.98 | 89 | 13 | 0.96 | 76* |
| 14 | 6 | 0.91 | 95 | 6 | 0.99 | 108 |
| 15 | 9 | 1 | 107 | 9 | 1 | 90 |
| 16 | 12 | 1 | 100 | 12 | 0.99 | 96 |
| 17 | 12 | 1 | 101 | 11 | 0.96 | 99 |
| 18 | 10 | 0.97 | 95 | 10 | 0.96 | 103 |
| <hr/> | | | | | | |
| Participante | Creatinina | | | Urico | | |
| | N | CORR | REL | N | CORR | REL |
| 01 | 5 | 0.70 | 128* | - | - | - |
| 02 | 3 | 0.91 | 135* | 2 | 1 | 131* |
| 03 | 5 | 1 | 128* | - | - | - |
| 04 | 17 | 1 | 89 | 17 | 0.93 | 118* |
| 05 | 11 | 0.98 | 107 | - | - | - |
| 06 | 5 | 1 | 148* | - | - | - |
| 07 | 4 | 0.64* | 124* | - | - | - |
| 08 | 13 | 1 | 96 | - | - | - |
| 09 | 3 | 0.73 | 70* | - | - | - |
| 10 | 8 | 1 | 143* | - | - | - |
| 11 | 12 | 0.97 | 94 | 5 | 0.91 | 120* |
| 12 | 10 | 0.73 | 103 | 3 | 1 | 121* |
| 13 | 13 | 0.43* | 95 | 3 | 0.99 | 71* |
| 14 | 6 | 0.97 | 93 | - | - | - |
| 15 | 9 | 1 | 130* | 9 | 0.99 | 115 |
| 16 | 12 | 1 | 92 | 11 | 0.99 | 93 |
| 17 | 12 | 1 | 116* | 11 | 0.93 | 109 |
| 18 | 10 | 0.47* | 74* | 10 | 0.60* | 61* |

* Datos inaceptables: CORR <0.70 y REL fuera de 85-115%

| Lugar | Lab | Clasificación de REL | | | Proporción REL buenas** |
|-------|-----|----------------------|-------|-------|----------------------------|
| | | Gluc | Urea | Creat | |
| 1 | 16* | Buena | Buena | Buena | 4/4=100% |
| 2 | 14 | Buena | Buena | - | 3/3=100% |
| 3 | 04* | Buena | Buena | Mala | 3/4= 75% |
| 3 | 15* | Buena | Buena | Mala | 3/4= 75% |
| 3 | 17* | Buena | Mala | Buena | 3/4= 75% |
| 6 | 03 | Buena | Buena | Mala | 2/3= 67% |
| 7 | 02 | Buena | Buena | Mala | 2/4= 50% |
| 7 | 11 | Mala | Buena | Mala | 2/4= 50% |
| 7 | 12 | Buena | Mala | Buena | 2/4= 50% |
| 7 | 13 | Buena | Mala | Buena | 2/4= 50% |
| 7 | 18* | Buena | Buena | Mala | 2/4= 50% |
| 12 | 09 | Buena | - | Mala | 1/2= 50% |
| 13 | 05* | Mala | Mala | Buena | 1/3= 33% |
| 13 | 06 | Buena | Mala | Mala | 1/3= 33% |
| 13 | 07 | Mala | Buena | Mala | 1/3= 33% |
| 13 | 08 | Mala | Buena | Mala | 1/3= 33% |
| 13 | 10 | Mala | Buena | Mala | 1/3= 33% |
| 18 | 01 | Mala | Mala | Mala | 0/3= 0% |

* Sistema automatizado

** REL buena= entre 85 y 115% del valor del laboratorio de referencia

quedan los sistemas buenos (índices de 25 a 99) y los regulares (100 a 199).

Sin embargo, la sencillez del índice de la OMS tiene una desventaja: la calificación no discrimina la inexactitud de la imprecisión, lo cual complica la evaluación de resultados en laboratorios con poca trayectoria en programas externos. Como ejemplo del problema que puede crear un índice afectado por imprecisión e inexactitud es que, si se aplica el índice de la OMS a este estudio, 29 de los 62 sistemas (47%) no deberían hacer estas mediciones por tener un índice arriba de 200 (cuadro V). Lo anterior iría en contra del hecho de que en 14 de los 18 participantes operaban programas internos de control de estas cuatro mediciones en el momento del estudio, lo cual sugiere que consideraban emitir datos confiables en mayor o menor grado.

En resumen, el sistema de la OMS al parecer no refleja la realidad y contrasta con la visión más discriminativa del presente análisis. Así, a pesar de que este estudio destaca una alta frecuencia de inexactitud, debe tenerse en cuenta que la presente estrategia detecta los problemas de exactitud externa (la comparada contra un valor de referencia confiable), lo cual no se opone a que puedan tener una buena exactitud interna, esto es, que aunque sus sistemas midan en forma inexacta lo hagan consistente y sistemáticamente, lo cual tiene la enorme ventaja de no desorientar a los clínicos usuarios. La imprecisión de un laboratorio provoca desconcierto y desconfianza y los problemas de inexactitud externa tienen mucha menor repercusión (si ésta no es de gran magnitud y se mantiene una buena exactitud interna).

| CUADRO V Puntuaciones que se obtendrían con el índice de variancia de la OMS en este estudio | | | | | |
|---|---------|------|------------|-------|----------|
| Lab | Glucosa | Urea | Creatinina | Urico | Promedio |
| 16* | 40 | 126 | 133 | 129 | 143 |
| 03 | 96 | 124 | 304 | - | 175 |
| 14 | 168 | 297 | 62 | - | 176 |
| 15* | 89 | 171 | 312 | 187 | 190 |
| 08 | 241 | 213 | 167 | - | 207 |
| 17* | 33 | 257 | 187 | 190 | 222 |
| 07 | 277 | 166 | 240 | - | 228 |
| 12 | 118 | 187 | 145 | 240 | 230 |
| 10 | 243 | 101 | 374 | - | 239 |
| 04* | 126 | 274 | 161 | 199 | 253 |
| 06 | 85 | 364 | 310 | - | 253 |
| 05* | 310 | 370 | 135 | - | 272 |
| 18* | 72 | 141 | 246 | 387 | 282 |
| 09 | 173 | - | 392 | - | 283 |
| 13 | 136 | 346 | 118 | 288 | 296 |
| 02 | 60 | 160 | 313 | 376 | 303 |
| 01 | 268 | 400 | 295 | - | 321 |
| 11 | 319 | 221 | 192 | 234 | 322 |

* Sistema automatizado

Se considera que cualquier estrategia que discrimine la imprecisión de la inexactitud y, de ésta, la interna de la externa, puede ser más útil para evaluar a los laboratorios de países no desarrollados, y que por ahora no es conveniente aplicar el índice de la OMS en países que no han gozado del beneficio de haber aplicado, desde hace lustros, programas externos a nivel nacional que facilitan la erradicación de los problemas de exactitud externa. En las circunstancias actuales, el índice de la OMS sólo puede servir para descalificar y, por lo tanto, desalentar.

Finalmente, cabe destacar que sólo el 8 por ciento de los sistemas que participaron en este estudio mostró imprecisión *versus* la tasa de 30 por ciento observada en un programa mexicano que ha explorado este punto.¹¹ Lo anterior sugiere que esta estrategia tiene una potencial fuente de error, cuando menos en la forma en que se

realizó el estudio: los laboratorios sabían con antelación acerca del día y la hora de la visita de los verificadores de la DGRSS, lo que pudo haber causado el llamado efecto del observador sobre el observado, o sea que los laboratorios pudieron haber intentado mostrar su mejor cara ante el órgano oficial de vigilancia de los laboratorios clínicos del país; por ejemplo, presentar al mejor analista, abrir un nuevo estuche de reactivos, pipetejar con mayor meticulosidad, etcétera. Es necesario evaluar la persistencia de la buena precisión cuando la visita de verificación no es anunciada al participante.

La calidad del sistema del LR mostró excelente precisión y buena exactitud durante el estudio, lo cual le da validez a las comparaciones.

La mayoría de los participantes logró un CORR bueno, lo cual se puede interpretar como ausencia de problemas de precisión el día de la visita de la DGRSS. Sin embargo, lo anterior podría haber sido subestimado por el llamado efecto del observador sobre el observado.

Por otra parte, 16 de los 18 participantes tuvieron uno o más sistemas con problemas de exactitud externa, ya que mostraron discrepancias sistemáticas mayores al 15 por ciento *versus* el LR.

La alta frecuencia de inexactitud externa (que no implica necesariamente inexactitud interna) sugiere la necesidad de contar con más calibradores de programas externos con valores asignados confiables.

CONCLUSIÓN

El análisis presentado tiene la ventaja de permitir discriminar las imprecisiones de las inexactitudes. Debe notarse, sin embargo, que evalúa exactitud externa (*versus* en LR en este caso) pero no explora para nada la exactitud interna (la del propio participante). En otras palabras, el hecho de que un participante tenga buena inexactitud externa no invalida el que pudiera tener una buena exactitud interna, de modo que el promedio de N mediciones en un mismo control no cambia a lo largo del tiempo en su sistema de medición. Desde este punto de vista, un laboratorio con buena precisión y buena exactitud interna puede operar razonablemente bien ya que su consistencia no provoca la desorientación del clínico usuario.

REFERENCIAS

1. Yamane T. Statistics: An introductory analysis. Nueva York: Harper & Row/J Weatherhill, 1967:443.
2. Siegel S. Estadística no paramétrica. 2a. ed. español. México, D.F.: Trillas, 1979:143-155.
3. Loría A. Programa INS de control de calidad. I. Precisión y exactitud relativas en cuatro mediciones de química sanguínea. *Rev Invest Clin* 1984;36:293-303.
4. Loría A. Programa INS de control de calidad. II. Estudios de seguimiento (fase 2). *Rev Invest Clin* 1985;37:385-390.
5. Loría A. Programa INS de control de calidad. III. Estudios de seguimiento (fase 3). *Rev Invest Clin* 1986;38:435-440.
6. Loría A. Programa INS de control de calidad. IV. El efecto de usar estándares de una sola fuente en sistemas imprecisos o inexactos. *Rev Invest Clin* 1987;39:385-389.
7. Loría A. Programa INS de control de calidad. V. Uso de una estrategia de programa interno/externo. *Rev Invest Clin* 1988;40:317-323.
8. Loría A, Zaltzman S, Cristerna JM. Programa permanente químico-clínico de los INS. I. La primera fase del programa. *Rev Invest Clin* 1993;45:353-362.
9. Loría A, Rosas-Baruch A, Moreno AJ, Villavicencio Ferral P, Rojas L. Programa permanente químico-clínico de los INS. II. Las fuentes de variación de dos analizadores imprecisos. *Rev Invest Clin* 1994;46:45-52.
10. Vargas de Cabral M, Castillo de Sánchez ML, Alva Estrada S. Programa de evaluación externa de la calidad de la Asociación Mexicana de Bioquímica Clínica. Resultados de la evaluación externa de la calidad en las determinaciones de glucosa y colesterol. *Bioquímica* 1989;14:27-34.
11. Curiel López P, Fuentes Mancilla LM, Cabaña Cortés EM, Lara UCM, Alva Estrada SI, Valles de Bourges V *et al.* Programa de evaluación de la calidad entre laboratorios. VII. Resumen de resultados de dos años. *Lab-Acta* 1993;5:37-42.