

Como Mejorar la Confiabilidad de las Mediciones para los Pacientes

Dr Hugo Gasca-Aragón
QFB Guiomar Pabello-Poegner
CENAM

16.05.2014

The logo for AACC (Asociación Argentina de Control de Calidad) is displayed in a bold, red, sans-serif font. The letters are thick and blocky. Below the main text, there is a faint, semi-transparent reflection of the same text, creating a subtle double-effect. The logo is positioned in the bottom right corner of the slide, which is a white triangular area.

Contenido

- Las normas oficiales y recomendaciones ISO
- Descripción del Programa de EA pilotado por el CENAM (2002-2006)
- Materiales de Referencia Certificados
- Evaluación de los Participantes
- Comparación de Valores por Consenso y Valores de Referencia Certificados y sus incertidumbres
- Estudio de Simulación para Inferir el comportamiento de la población
- Resultados del Estudio de Simulación
- Conclusiones

NOM-064-SSA1-1993 [18]

Especificaciones sanitarias de los equipos de reactivos utilizados para diagnóstico

1.5.1 Reproducibilidad intralote para validar precisión Intralote

El coeficiente de variación de los resultados debe ser menor del 5%

2.5.2 Reproducibilidad interlote para validar precisión interlote

El coeficiente de variación de los resultados debe ser menor del 5%

3.5.3 Exactitud

El “índice” (sesgo) de exactitud de los estudios debe ser menor del 5%

Recomendaciones ISO (17043,13528)

Valor asignado utilizado en una comparación

Ordenados de mayor a menor orden metrológico [4,5]

1. Valor Conocido, si no es viable (4.4.5.3 en [4]) entonces
2. Valor de Referencia Certificado, si no es viable entonces
3. Valor de Referencia, si no es viable entonces
4. Valor de Consenso de Participantes Expertos, si no es viable entonces
5. Valor de Consenso de Participantes

[4] 4.4.5.3 “Para los programa de EA en áreas distintas de la de calibración, se deben determinar la necesidad y **viabilidad** de la trazabilidad metrológica, la relevancia e incertidumbre del valor asignado,...”

Programa de EA en Análisis Clínicos del CENAM

Motivos

1. Diseminar las mejores prácticas metrológicas a los Participantes
2. Impactar directamente en la salud de la población
3. Alinear esfuerzos para diseminar la trazabilidad metrológica en la medición a raíz de la implantación de la Directiva Europea en el año 2000 sobre la medición de reactivos clínicos
4. Desarrollar y hacer disponibles nuevos métodos de análisis estadístico y para evaluación de los Participantes

Programa de EA en Análisis Clínicos del CENAM

Descripción

1. Ejecutado durante los años 2002-2004-2006
2. Enfocado en los principales marcadores de química sanguínea (calcio, glucosa, colesterol, creatinina, urea y ácido úrico)
3. Se utilizaron Materiales de Referencia Certificados (MRC) producidos por el CENAM
4. Un total de 63 participaciones (Laboratorio-comparación) durante los 3 años (2002:44, 2004:13, 2006:6)

Material de Referencia Certificado

(5.13) Material suficientemente **homogéneo** y **estable** respecto a una o más propiedades, usada en calibración, en la asignación del valor de otro material o en aseguramiento de calidad [15]. (5.14) acompañado por su **documentación** emitida por un **instituto autorizado**, mencionando los procedimientos válidos usados para obtener el valor de la propiedad especificada con **incertidumbre** y **trazabilidad** [15].

Material de Referencia Certificado

Certificados mediante la técnica de dilución isotópica, un método reconocido internacionalmente como potencialmente primario

1. Materiales altamente homogéneos y estables bajo las condiciones recomendadas de almacenamiento y traslado. La incertidumbre estándar relativa de los valores asignados es menor al 3% en un 94% de los valores certificados [1,2,3].
2. La incertidumbre asignada considera todas las fuentes probables que intervienen en la medición [1,2,3].
3. Los valores de los mensurandos varía en un rango de 3 ordenes de magnitud (0.7 a 162 mg/dl) [1,2,3].

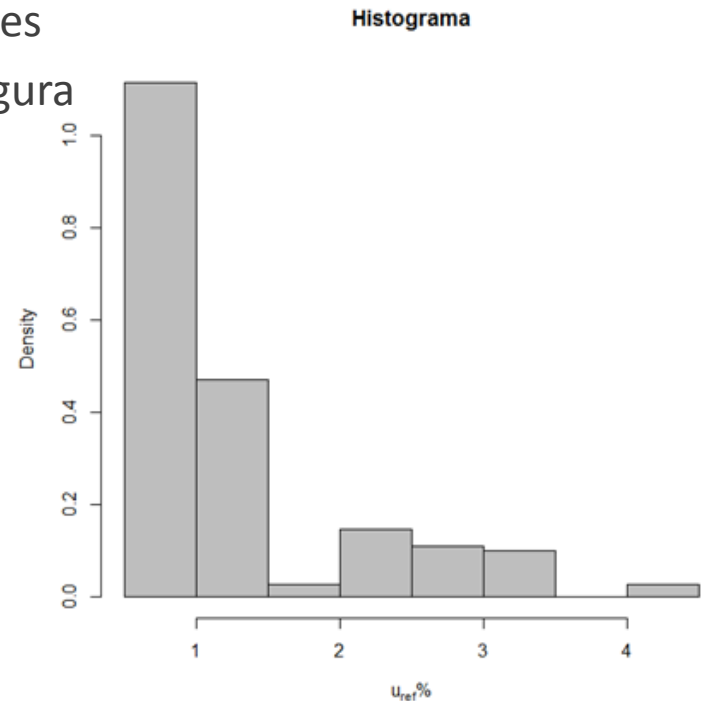
Material de Referencia Certificado

Que la incertidumbre estándar relativa de los valores asignados sea menor al 5% en todos los casos asegura con alta probabilidad que...

1. Los MRC son suficientemente **homogéneos y estables**.
2. Una desviación significativa (>5%) de estos valores por parte de un participante:

No puede ser atribuida al material mismo.

No puede ser explicada mediante el argumento de variabilidad biológica [8].



Evaluación de los Participantes

El Análisis original fue mediante la raíz del error cuadrático medio [9], el cual es relativamente poco conocido

En esta revisión se utilizó la estadística de diferencia porcentual $D\%$ [4,5], que es una estimación del sesgo relativo respecto del valor de referencia

$$D\% = \frac{X_{\text{lab}} - X_{\text{MRC}}}{X_{\text{MRC}}} \times 100$$

En una segunda etapa (estudio de simulación) se analizó la participación utilizando la estadística zeta [4,5] utilizando también el consenso de los participantes

$$\zeta = \frac{X_{\text{consenso}} - X_{\text{MRC}}}{\sqrt{u_{\text{consenso}}^2 + u_{\text{MRC}}^2}}$$

Evaluación de los Participantes

1. La evaluación fue condicionada bajo dos condiciones

A cada mensurando en cada muestra

A cada participante

2. Limitantes

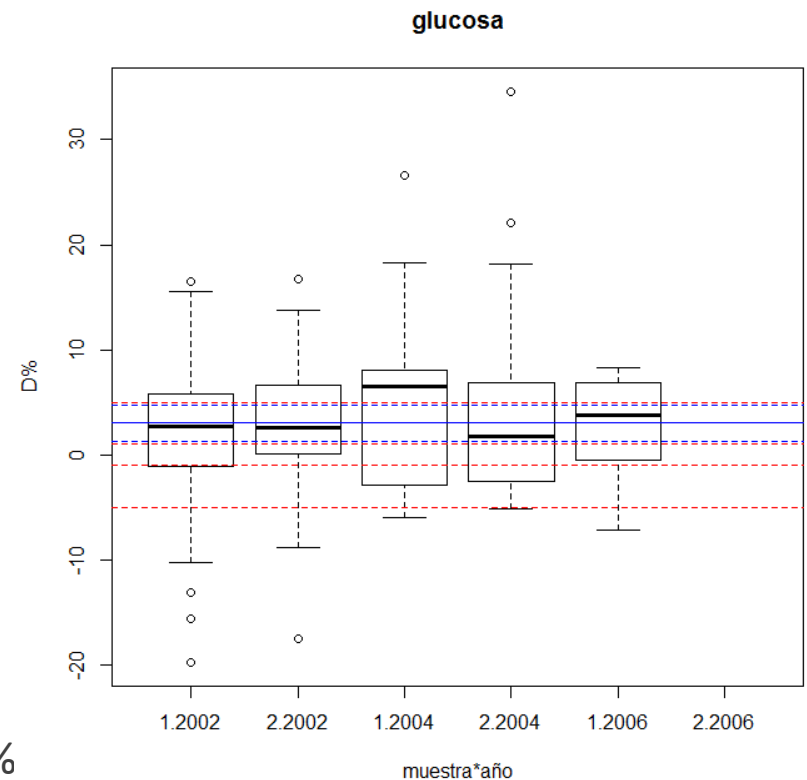
Las condiciones para un análisis de tendencias por laboratorio no se cumplieron debido a la confidencialidad que restringió poder asociar los resultados entre diferentes años

Resultados condicionados a cada mensurando

Existe la tendencia a sobre estimar el mensurando

Las líneas rojas indican el -5%, -1%, 1% y 5%

Las líneas azules representan el valor promedio y su intervalo de confianza del 95%

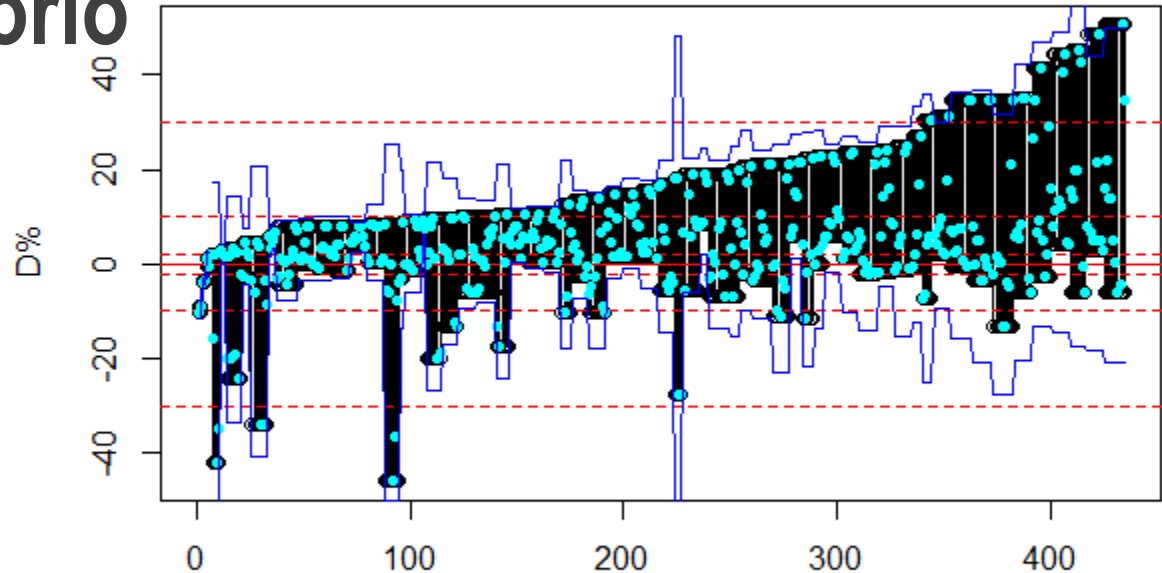


Resultados condicionados a cada mensurando

Mensurado	Valor promedio $D_{\%}$	Incertidumbre estándar del valor promedio $D_{\%}$	RECM de $D_{\%}$
Calcio	-0.00007	0.9	0.94
Glucosa	3.0	0.9	3.16
Colesterol	3.7	1.4	3.97
Creatinina	14.0*	2.1	14.19
Urea	2.2	1.1	2.44
Ácido úrico	2.3	1.1	2.52

* El promedio de los participantes tuvo resultados no conformes con los requerimientos de la norma NOM-064-SSA1-1993 [18].

Resultados condicionados a cada laboratorio



Existe la tendencia a sobre estimar el mensurando $\text{lab} \cdot \text{analito}$

El 10% de los laboratorios tiene una diferencia porcentual del 10% o menor

Casi el 8% de los laboratorios muestran evidencia fuerte de un error sistemático significativo. Su banda de confianza no contiene el cero

Comparativo de Valores Asignados

Se procedió a comparar la diferencia entre valores asignados por valor de referencia certificado y el valor de consenso de los participantes.

Se utilizó la estadística zeta [4,5] por claridad de la interpretación.

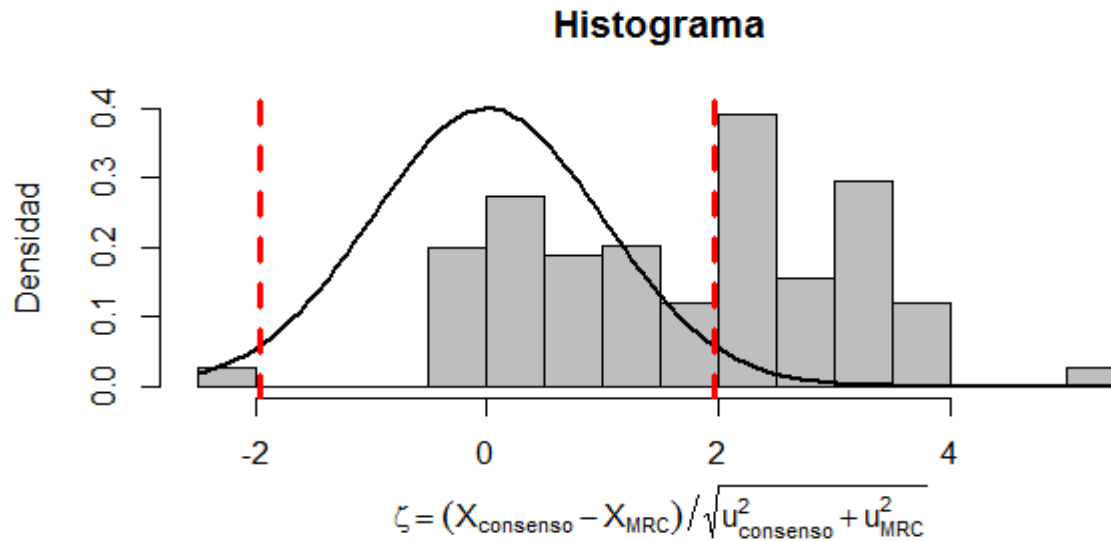
$$\zeta = \frac{X_{consenso} - X_{MRC}}{\sqrt{u_{consenso}^2 + u_{MRC}^2}}$$

1. La estadística zeta se comporta como una normal estándar bajo los supuestos de

normalidad de los términos de error
no hay presencia de error sistemáticos

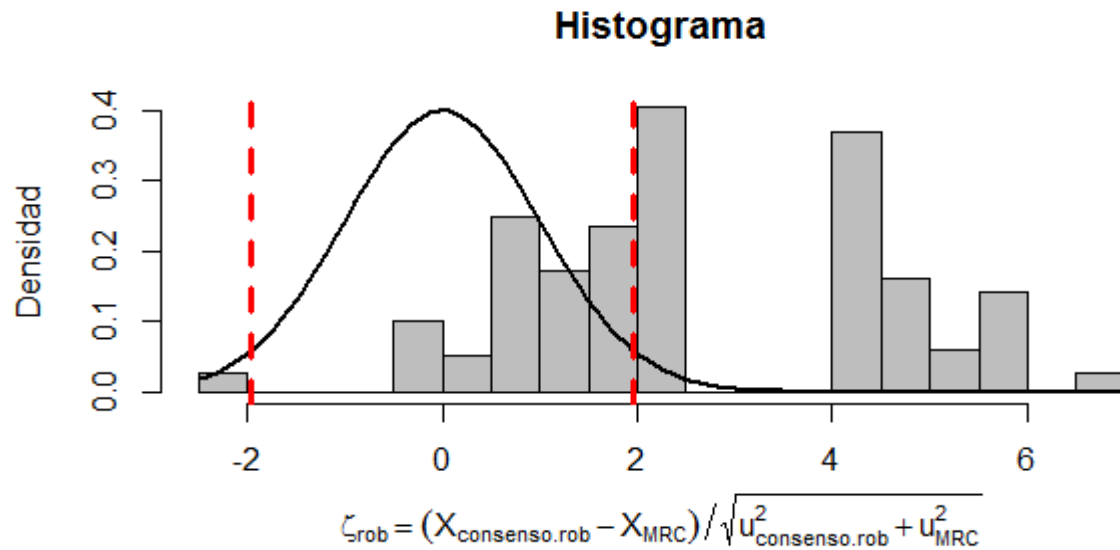
2. El valor por consensos de participantes se estimó mediante
Estadísticas no robustas (media y desviación estándar)
Estadísticas robustas [14] (mediana y mediana de la desviación absoluta)

Comparativo Valores de Consenso estimado con estadísticas no robustas



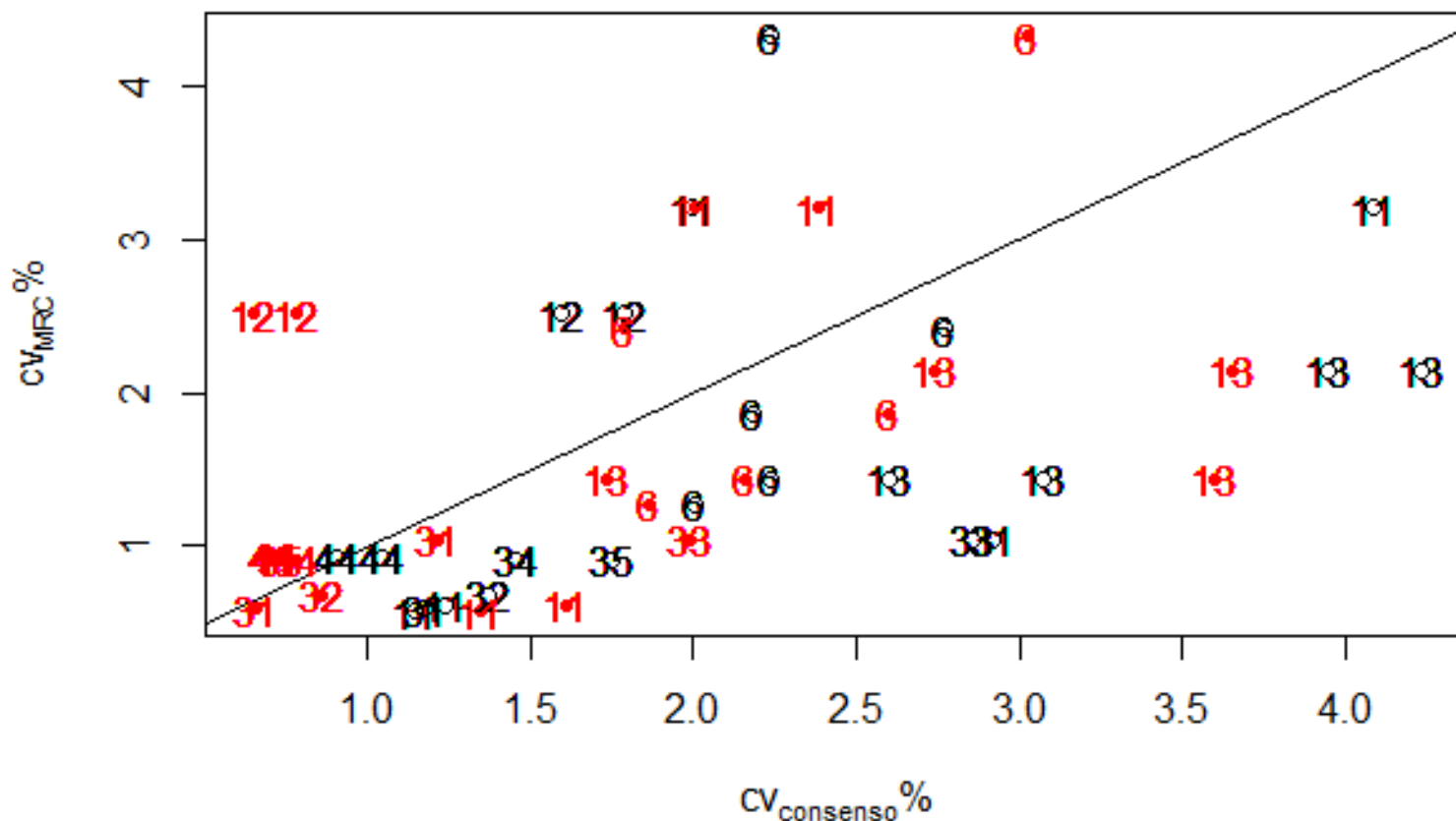
1. Un 51% de los valores asignados por el consenso de los participantes utilizando estadísticas no robustas aparecen sesgados (no satisfactorios).
2. La evaluación puede ser distorsionada por la presencia de algunos valores atípicos.

Comparativo Valores de Consenso estimado con estadísticas robustas

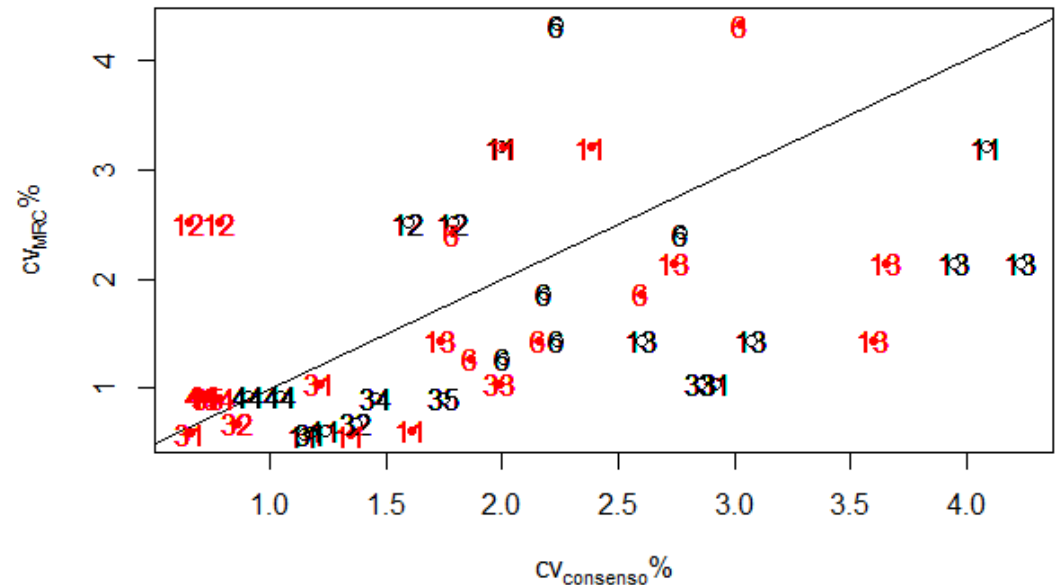


1. Un 60% de los valores asignados por el consenso de los participantes utilizando estadísticas robustas aparecen sesgados (no satisfactorios). Las estimaciones son robustas ante la presencia de un 50% de valores atípicos.
2. Este resultado sugiere que la mayoría de los participantes tienen un desempeño deficiente.

Comparativo Incertidumbre de los Valores Asignados por Consenso



Comparativo Incertidumbre de los Valores Asignados por Consenso



1. Los valores asignados por consenso de los participantes tienden a tener mayor incertidumbre que los valores asignados de un MRC (aparecen con mayor frecuencia debajo de la línea)
2. Los valores asignados por consenso de pocos participantes son menos realistas y tienden a subestimar la incertidumbre real

Estudio de Simulación Diseño

Supuestos

1. La muestra es representativa de la población de participantes
2. La incertidumbre de a mediano plazo es constante entre los participantes

Para asegurar que no se indujera ninguna propiedad de distribución en los resultados, se consideró una técnica de re-muestreo no-paramétrico y explotamos esta propiedad en la prueba de hipótesis

Estudio de Simulación Diseño

El estudio se limitó a los mensurandos:

Glucosa (63 participantes en 5 comparaciones durante 3 años),

Colesterol (53 participantes en 5 comparaciones durante 3 años) y

Creatinina (52 participantes en 5 comparaciones durante 3 años).

Las muestras de los siguientes mensurandos parecen no ser representativas:

Calcio (sólo hay datos de 2 comparaciones en un periodo de 1 año)

Urea (sólo hay 17 participantes en 3 comparaciones durante 2 años)

Ácido Úrico (sólo hay 17 participantes en 3 comparaciones durante 2 años)

Estudio de Simulación

Procedimiento de Evaluación

Para cada mensurando se formaron 5 grupos:

1. Participantes sesgados $|D_{\%}| > \Phi^{-1}(0.975)u[D_{\%}]$
2. Participantes no sesgados con incertidumbre relativa menor al 10%
3. Participantes no sesgados con incertidumbre relativa entre 10% y 15%
4. Participantes no sesgados con incertidumbre relativa entre 15% y 20%
5. Participantes non sesgados con incertidumbre relativa mayor al 20%

La evaluación se realizó bajo dos condiciones (sensibilidad a valores atípicos [14])

- i. Usando todas las sub-muestras
- ii. Excluyendo las sub-muestras de participantes con alta incertidumbre o valores sesgados (Este es un criterio para eliminar valores atípicos en general).

Estudio de Simulación

Algoritmo

Para cada mensurando se generaron un número grande (>1000) de re-muestras aleatorias considerando:

1. Conservar la estructura del experimento, dos datos relacionados en el año 2002, dos datos relacionados en 2004 y un dato aislado en 2006.
2. Para cada re-muestra se calcula el consenso mediante el promedio y la desviación estándar del promedio (como medida de precisión intermedia).
3. Se grafican los datos simulados en una gráfica de objetivo modificada y se catalogan los casos en las 5 familias.
4. Se analiza la distribución de los resultados bajo las dos condiciones (i e ii) en base a la categorización del paso previo.

Estudio de Simulación

Gráficos de Objetivo [12]: muestran una medida de sesgo (error sistemático) y la una medida de dispersión/incertidumbre (error aleatorio) de cada participante. Organizados en semicírculos concéntricos. Un resultado aceptable aparece cerca del origen/objetivo. Los semicírculos representan 1, 2 y 3 desviaciones estándar del valor asignado.

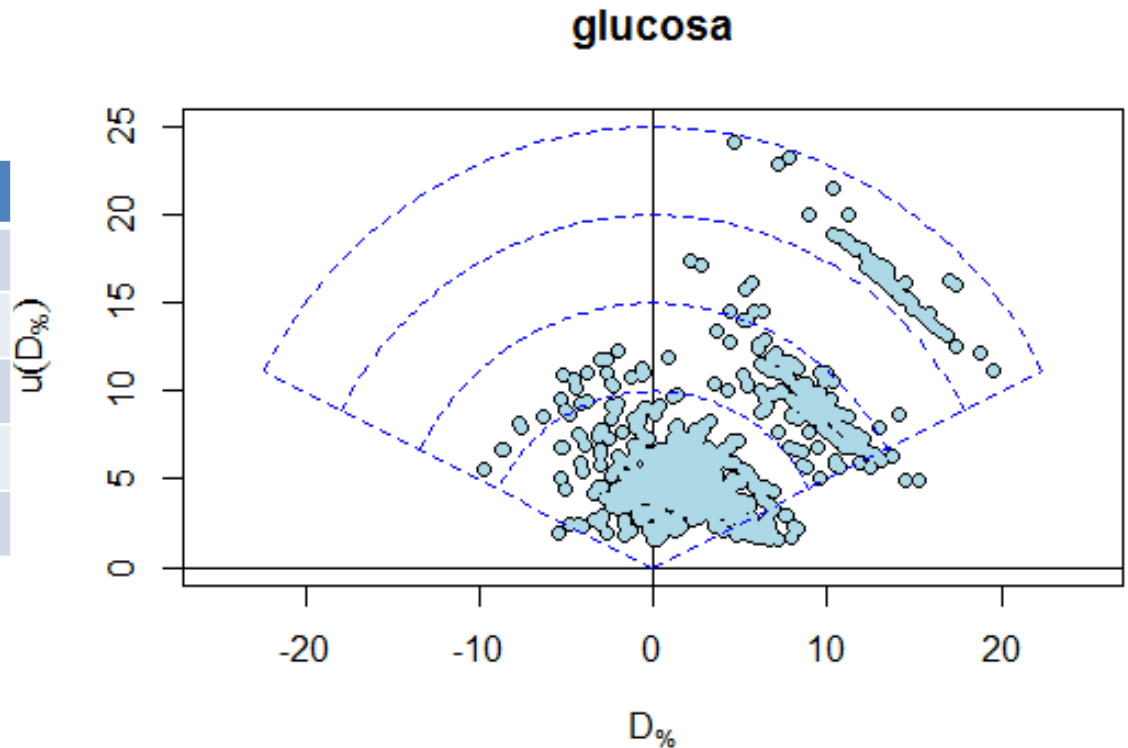
1. En la gráfica modificada las cantidades son relativas al valor de referencia del mensurando para hacer comparables los resultados en el tiempo y entre participantes.
2. La gráfica modificada utiliza abanicos en vez de semicírculos concéntricos. Los regiones en forma de abanicos son una aproximación a la región de mayor probabilidad de la distribución conjunta de ambos parámetros bajo normalidad e independencia [13]. Las líneas diagonales satisfacen la ecuación

$$|D_{\%}| = \Phi^{-1}(0.975)u[D_{\%}], \quad u[D_{\%}] \approx \frac{\sqrt{\sigma_{mt,lab}^2 + u_{MRC}^2}}{X_{MRC}} \times 100$$

Estudio de Simulación

Resultados para Glucosa

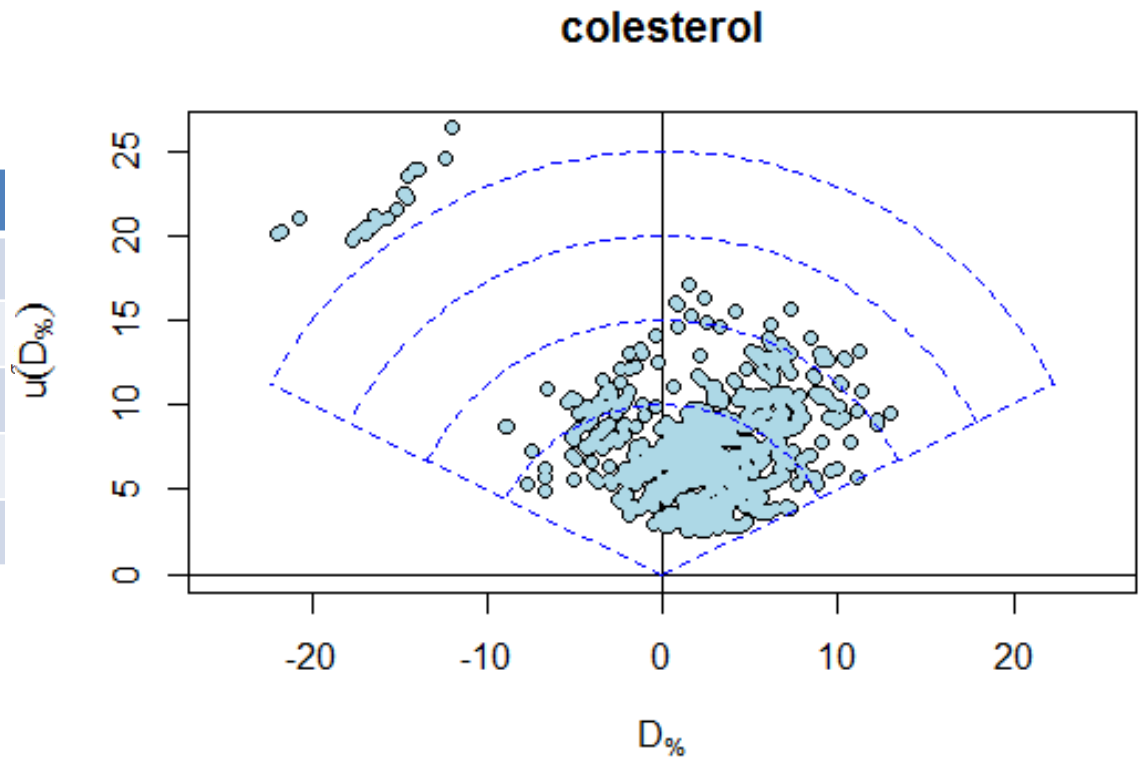
Participantes	$u[D\%]$	%
sesgados		6
no-sesgados	> 20	9
	> 15	1
	> 10	20
	≤ 10	64



Estudio de Simulación

Resultados para Colesterol

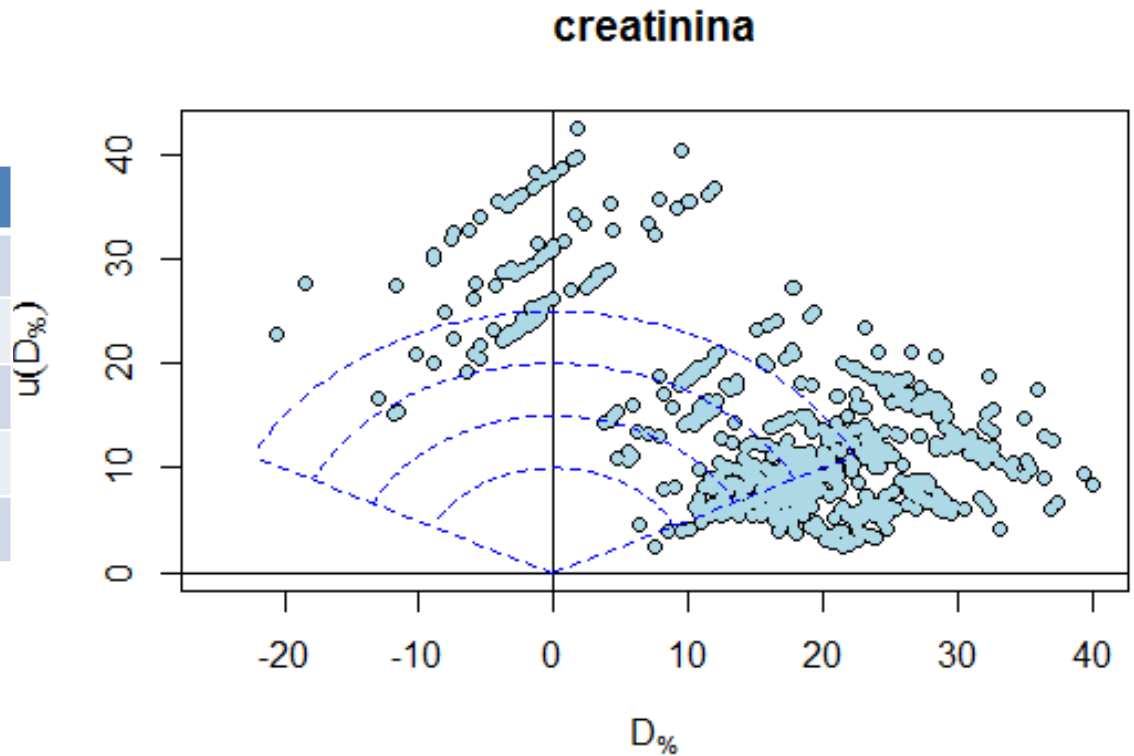
Participantes	$u[D\%]$	%
sesgados		0
no-sesgados	> 20	3
	> 15	4
	> 10	22
	≤ 10	71



Estudio de Simulación

Resultados para Creatinina

Participantes	$u[D\%]$	%
sesgados		40
no-sesgados	> 20	43
	> 15	13
	> 10	5
	≤ 10	<1



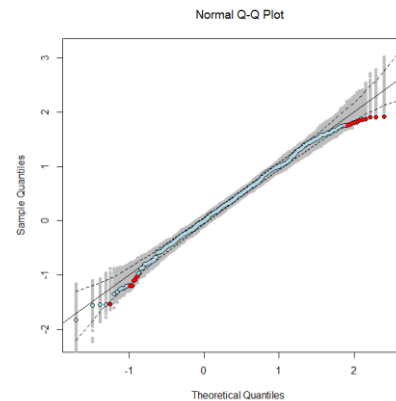
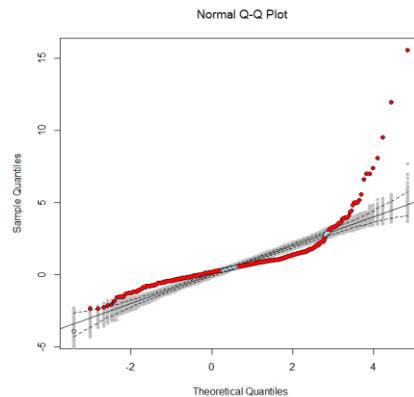
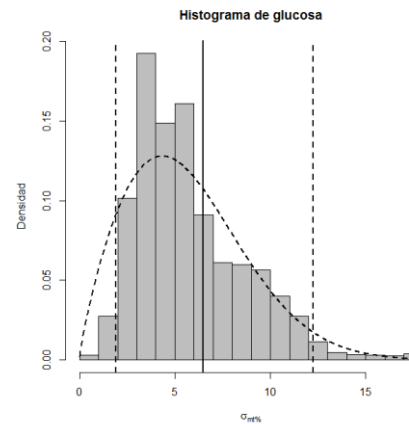
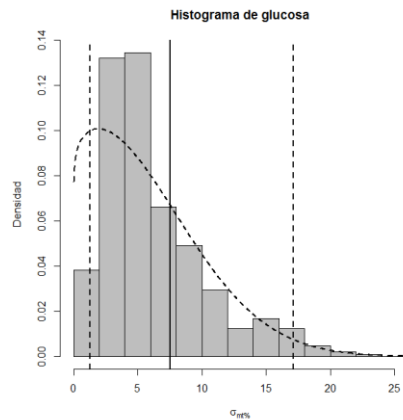
Estudio de Simulación

Resultados de Evaluación

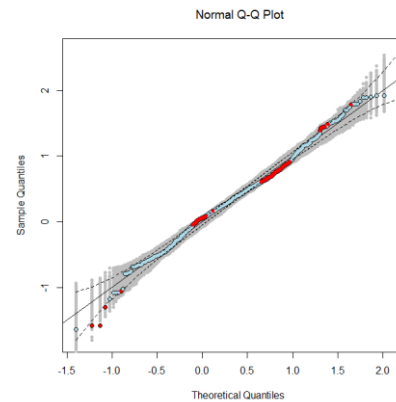
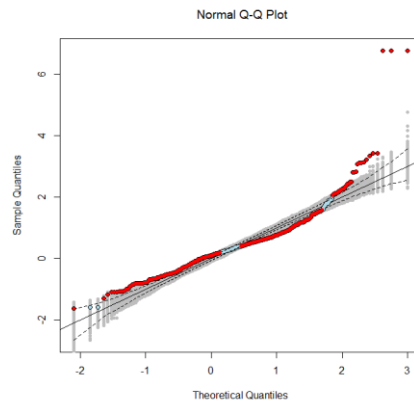
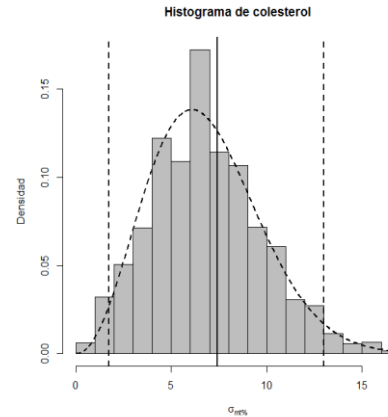
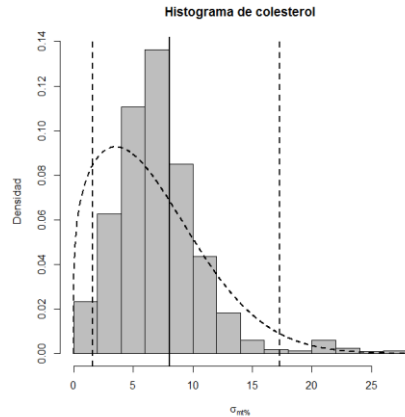
Para cada mensurando se graficaron un histograma de la incertidumbre de precisión intermedia y una gráfica de cuantiles-cuantiles, bajo las condiciones (i) e (ii)

Bajo la condición (ii) (filtrando los probables valores atípicos) se infiere que los datos siguen una distribución normal (no se tiene evidencia para rechazar la hipótesis).

Estudio de Simulación Resultados de Evaluación: Glucosa

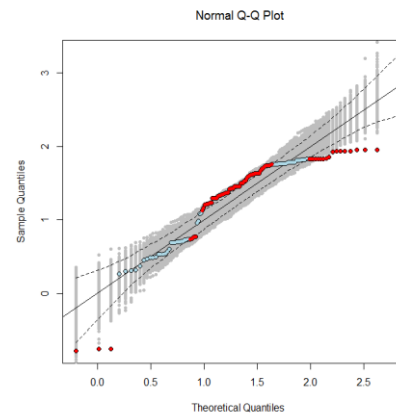
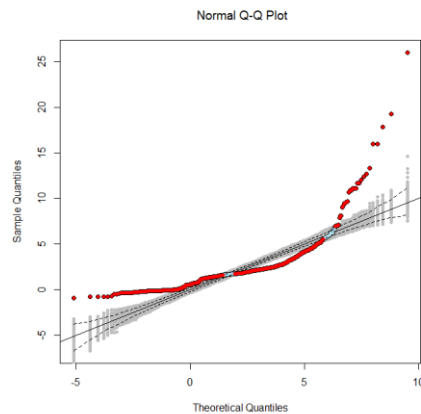
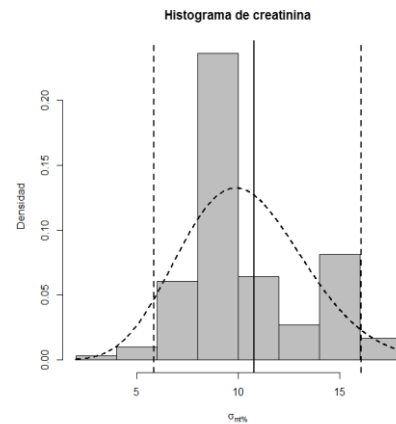
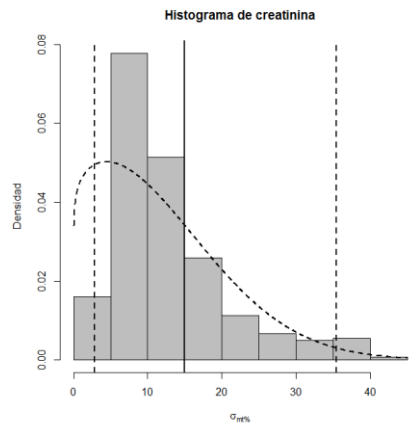


Estudio de Simulación Resultados de Evaluación: Colesterol



Estudio de Simulación

Resultados de Evaluación: Creatinina



Conclusiones

- Los valores de consenso tienden a generar evaluaciones sesgadas con un riesgo no controlado y este sesgo es heredado a los servicios ofrecidos al paciente final, en los valores de referencia certificados cualquier fuente probable de sesgo es analizada y corregida de origen mediante comparaciones con otros INM.
- La incertidumbre de un valor de consenso tiende a ser mayor que la incertidumbre de un valor de referencia certificado y tiende a subestimar la incertidumbre real si el número de participantes es relativamente pequeño.
- Las técnicas de detección de valores atípicos no garantizan un valor de consenso confiable para evaluar a los participantes.
- El estudio de simulación con la técnica presentada sugiere que los resultados finales de una evaluación excluyendo los valores sesgados y los valores con alta variabilidad tienden a comportarse como variables con distribución normal y por consecuencia podemos aplicar los resultados de la teoría estadística en estimaciones, pruebas de hipótesis e inferencias probabilísticas.

Reconocimiento

Agradecemos a los laboratorios participantes por las mediciones realizadas y a los metrólogos y directivos del CENAM por el soporte analítico durante el programa de EA:

Marcos Antonio Ávila Calderón

Laura Regalado Contreras

Mauricio Maldonado Torres

Miryan Balderas Escamilla

Gabriela Salazar Briones

Judith Rivera Mellado

Esther Castro Galván

Melina Perez Urquiza

Yoshito Mitani Nakanishi

Bibliografía

- [1] Mitani N. Y. & Pérez U. M. (2002). Informe de Resultados obtenidos en la prueba de aptitud técnica 630-IL-1002b para los laboratorios químico-clínicos que realicen mediciones de glucosa, colesterol, creatinina y calcio en suero. CENAM. Querétaro, México.
- [2] Mitani N. Y. & Pérez U. M. (2004). Informe del estudio comparativo 630-Q023-1003 cuantificación de glucosa, creatinina, colesterol, ácido úrico y urea en suero humano. CENAM. Querétaro, México.
- [3] Mitani N. Y. & Pérez U. M. (2006). Informe de la prueba de aptitud técnica 630-Q023-0054-PA Suero Humano. CENAM. Querétaro, México.
- [4] ISO/IEC 17043/NMX-CH-17043-IMNC-2010 (2010). Evaluación de la Conformidad – Requisitos generales para los ensayos de aptitud.
- [5] ISO/IEC 13528/NMX-CH-13528-IMNC-2007 (2007). Métodos estadísticos para su uso en ensayos de aptitud por comparaciones interlaboratorios.
- [6] Efron (1979). Bootstrap methods: Another look at the Jackknife. The Annals of Statistics, Vol 7, No 1, pp 1-26.
- [7] BIPM-JCGM (2008). Evaluation of measurement data - Guide to the expression of uncertainty in measurement.
- [8] Harris et al (1970). Biological and Analytic Components of Variation in Long-Term Studies of Serum Constituents in Normal Subjects. II. Estimating Biological Components of Variation. Clin. Chem; 16, 1022-1027.

Bibliografía (cont)

- [9] Castelazo & Mitani (2012). On the use of the mean square error as a proficiency index. Accred. Qual. Assur, 17, 95-97.
- [10] Gasca et al (2013). Notas del taller de ensayos de aptitud. SIM week 2013. CENAM. Querétaro, México.
- [11] Thompson & Ellison (2011). Dark uncertainty. Accred. Qual. Assur. 16, pp 483-387.
- [12] Duewer D.L. et al (1999). Micronutrients Measurement Quality Assurance Program: Helping Participants Use Interlaboratory Comparison Exercise Results To Improve Their Long-Term Measurement Performance. Anal. Chem. 71, pp 1870-1878.
- [13] Vangel & Rukhin. (1999). Maximun Likelihood Analysis for Heteroscedastic One-Way Random Effects ANOVA in Interlaboratory Studies. Biometrics, Vol 55, No 1, 129-136.
- [14] Huber P.J. (1981). Robust Statistics. John Wiley. NY.
- [15] ISO/IEC 99. NMX-Z-055-IMNC (2009). Vocabulario Internacional de Metrología.
- [16] The BIPM key comparison database <http://kcdb.bipm.org>
- [17] Comunicación electrónica con Marco Antonio Ávila Calderón.
- [18] NOM-064-SSA1-1993. Especificaciones sanitarias de los equipos de reactivos utilizados para diagnóstico.