

NOTA IMPORTANTE:

La entidad sólo puede hacer uso de esta norma para si misma, por lo que este documento NO puede ser reproducido, ni almacenado, ni transmitido, en forma electrónica, fotocopia, grabación o cualquier otra tecnología, fuera de su propio marco.

ININ/ Oficina Nacional de Normalización

CERTIFICACION DE MATERIALES DE REFERENCIA PRINCIPIOS GENERALES Y ESTADISTICOS

Certification of reference materials
General and statistical principles

Descriptores: Material de referencia; certificación;
conceptos básicos

1. Edición

1998

ICS: 01.120

REPRODUCCION PROHIBIDA

Oficina Nacional de Normalización (NC) Calle E No. 261 Vedado, Ciudad de La Habana. Te-
léf.: 30-0835 Fax: (537) 33-8048 E-mail: ncnorma@ceniai.inf.cu

Prefacio

La Oficina Nacional de Normalización (NC), es el Organismo Nacional de Normalización de la República de Cuba que representa al país ante las Organizaciones Internacionales y Regionales de Normalización.

La preparación de las Normas Cubanas se realiza generalmente a través de los Comités Técnicos de Normalización. La aprobación de las Normas Cubanas es competencia de la Oficina Nacional de Normalización y se basa en evidencias de consenso.

Esta norma:

- Ha sido elaborada por la Oficina Nacional de Normalización.
- Esta norma es idéntica a la **ISO Guía 35 “Certificación de materiales de referencia- Principios generales y estadísticos”**. Segunda edición 1989.
- Sustituye a la NC-ISO Guía 35:1994.
- Las referencias normativa que aparecen en el texto con respecto a la norma ISO se sustituyen por las relativas a las normas cubanas que correspondan con dichas normas, en los casos en que éstas existan.
- Consta de un Anexo informativo.

© NC, 1998

Todos los derechos reservados, a menos que se especifique, ninguna parte de esta publicación podrá ser reproducida o utilizada por alguna forma o medios electrónicos o mecánicos, incluyendo las fotocopias o microfilmes, sin el permiso previo escrito de:

Oficina Nacional de Normalización (NC).

Calle E No. 261 Ciudad de La Habana, Habana 3. Cuba.

Impreso en Cuba

Indice

1. Objetivo.....	1
2. Definiciones	2
3. Papel de los materiales de referencia en la ciencia de la medición	3
4. Incertidumbre de medición.....	7
5. Homogeneidad de los materiales	15
6. Principios generales de certificación	22
7. Certificación por un método definitivo	25
8. Certificación por ensayos interlaboratorios	29
9. Certificación basada en un enfoque metrológico.....	42
ANEXO A. Bibliografía.....	58

Introducción

El Comité de Materiales de Referencia (REMCO) está comprometido con las pautas para la preparación, certificación y uso de los materiales de referencia. Esta Guía se propone describir los principios generales y estadísticos para la certificación de materiales de referencia.

Varias secciones de esta Guía fueron preparadas por diferentes delegados al REMCO. Este proyecto fue coordinado con representantes del TC 69 de la ISO, "Aplicaciones de métodos estadísticos".

Agradecimientos para J.D. Cox (BSI, UK) por la preparación de la sección sobre el papel de los materiales de referencia en sistemas de medición (sección 3). Las secciones 4, 5 y 6 están basadas sobre material contenido en tres fuentes publicadas previamente:

- a) Cali, J.P. et al. El papel de los materiales de referencia patrones en los sistemas de medición, NBS, Monografía 148, Washington, DC, Buró Nacional de Normas, 1975 (especialmente Capítulo III, por H.H. Ku);
- b) Uriano, G.A. y Gravatt, C.C. El papel de los materiales de referencia y los métodos de referencia en análisis químico. Crit. Rev. Química Analítica **6**, 1977. **361**;
- c) Marschal, A. Materiales de referencia. Buró Nacional de Metrología, Laboratorio Nacional de Ensayos, París.

K.R. Eberhardt (ANSI, USA) preparó la sección 7 sobre el uso de un método definitivo para certificar materiales de referencia. R. Sutarno y H. Steger (SCC, Canada) prepararon la sección 8 sobre el uso de un programa de ensayo interlaboratorios para certificar materiales de referencia. H. Marchandis (Buró General de Referencia, Comisión de las Comunidades Europeas) preparó la sección 9 sobre introducción metrológica a la certificación, incluida por primera vez en la segunda edición de esta Guía. G. Uriano (ANSI, USA) sirvió como editor de esta Guía.

Especial reconocimiento para los miembros del TC 69/SC6 de la ISO y a su secretario, K. Petrick (DIN, Alemania, RF), por su cooperación en la preparación de aquellas secciones del documento relacionadas con el análisis estadístico de los datos. En particular se reconocen gratamente las numerosas contribuciones del Prof. P. T. Wilrich (DIN, Alemania, RF) y del Dr. T. Miyasu (JISC, Japón) del TC 69/SC 6 de la ISO, para la revisión y edición de la Guía.

Las guías anteriores^[1-3] preparadas por REMCO han tratado los siguientes aspectos de los materiales de referencia:

- a) mención de los materiales de referencia en Normas Internacionales;
- b) términos y definiciones usados en relación con los materiales de referencia;
- c) los contenidos de los certificados de los materiales de referencia.

El propósito de esta Guía es proporcionar una introducción básica a los conceptos y aspectos prácticos relacionados con el uso de los materiales de referencia. La NC ISO Guía 33^[29] trata más ampliamente los conceptos y aspectos prácticos relacionados con el uso de los materiales de referencia.

CERTIFICACION DE MATERIALES DE REFERENCIA. PRINCIPIOS GENERALES Y ESTADISTICOS

1 Objetivo

De acuerdo con la definición dada en 2.1, los materiales de referencia (MR) pueden ser usados en diversas funciones de medición relacionadas con la calibración de instrumentos, métodos de evaluación y asignación de valores propios. El propósito de la sección 3 es discutir estas funciones de medición y mostrar como la trazabilidad¹ de la medición puede ser asegurada por el uso de los MR, por lo tanto proporcionar compatibilidad universal de la medición.

Lo mismo que los materiales de referencia certificados (MRC) se prefieren sobre otras clases de MR en las citas de Normas Internacionales¹, así también se prefieren generalmente los MRC sobre otras clases de MR en la ciencia de las mediciones, dado que existen los MRC necesarios para un tipo particular de medición. La asistencia en la localización de fuente(s) de suministros de MRC para distintos campos técnicos se realiza por medio del Directorio de Materiales de Referencia Certificados de la ISO^[4].

Es evidente que la calidad de una medición basada en el uso de un MRC dependerá en parte del esfuerzo y el cuidado que brinde el órgano de certificación en la determinación de los valores propios del candidato a MRC. De aquí, que el proceso de certificación^[2] debe ser llevado a cabo usando métodos de medición bien caracterizados que tengan una alta exactitud y precisión y suministren valores propios trazables a las unidades fundamentales de medición. Además, los métodos deben producir valores con incertidumbres que sean apropiadas para el uso final esperado del MRC. Las secciones 4 y 5 están en correspondencia con dos de las más importantes consideraciones técnicas en la certificación de los MR: las incertidumbres de medición y la homogeneidad del material. La sección 6 estipula los principios generales para la certificación del MR.

Dos enfoques generales comúnmente usados para asegurar técnicamente la certificación válida del MR se exponen en las secciones 7 y 8. La sección 7 describe el uso de un método único de la más alta exactitud (es decir, algunas veces referido como un método "absoluto" o "definitivo") y usualmente usado por un laboratorio único para la certificación del MR. La sección 8 describe el uso de un enfoque de ensayo interlaboratorios para la certificación de MR, el cual pudiera involucrar más de un método .

El enfoque metrológico expuesto en la sección 9 tiene como objetivo la producción de valores certificados, cuya exactitud e incertidumbre son demostradas por evidencias experimentales.

En resumen, el propósito de esta Guía es contribuir a la comprensión de métodos válidos para la certificación de los MR y también ayudar a los usuarios potenciales a definir mejor sus requisitos técnicos. La Guía también puede ser útil en el establecimiento de un potencial completo de MRC como una ayuda para asegurar la exactitud y compatibilidad interlaboratorios de mediciones realizadas a escala nacional o internacional.

1.) Una definición acordada internacionalmente de "trazabilidad" en la ciencia de las mediciones es dada en la referencia [5]:

trazabilidad: La propiedad del resultado de una medición a través de la cual ésta puede ser relacionada a los patrones adecuados, generalmente patrones nacionales o internacionales, a través de una cadena ininterrumpida de comparaciones.

2 Definiciones

Las definiciones de los términos básicos “material de referencia” y “material de referencia certificado” fueron establecidas al final de 1977^[1] y más tarde ligeramente enmendadas^[2] hasta llegar a las siguientes.

2.1 material de referencia; MR: Un material o sustancia de la cual una o más propiedades están suficientemente bien establecidas para ser usadas en la calibración de un aparato, la evaluación de un método de medición, o para asignar valores a materiales.

NOTA Un MR puede estar en forma de un gas puro o mezclado, líquido o sólido, o incluso un simple objeto manufacturado. Algunos MR son certificados en un lote, donde cualquier parte razonablemente pequeña del mismo puede exhibir el valor propio establecido para el lote completo dentro de los límites de incertidumbre establecidos. Existen otros MR como objetos manufacturados individuales, los cuales también son certificados individualmente.

Numerosos MR tienen propiedades que, debido a que las mismas no pueden ser correlacionadas con una estructura química establecida o por otras razones, no pueden ser medidas en unidades de masa o en de cantidad de sustancia o determinadas por métodos de medición físicos o químicos exactamente definidos. Tales MR incluyen ciertos MR biológicos (por ejemplo una vacuna para la cual ha sido asignada una unidad internacional por la Organización Mundial de la Salud) y ciertos MR tecnológicos (por ejemplo bloques de goma para la determinación de abrasividad o placas de acero para la determinación de dureza). Es reconocido que la definición de “material de referencia” dada anteriormente puede implicar un solapamiento con el término “medida materializada” como está definido en el *Vocabulario Internacional de Términos Generales y Básicos en Metrología*^[5]; consecuentemente, algunos materiales pueden ser caracterizados tanto como materiales de referencia o como medidas materializadas.

2.2 material de referencia certificado; MRC: Un material de referencia, donde uno o más de sus valores propios son certificados por un procedimiento técnicamente válido, acompañado de un certificado u otra documentación trazable, la cual es expedida por un órgano de certificación.

NOTA Un MRC puede constar de unidades, cada una de las cuales es certificada individualmente o certificada por examen de muestras representativas de un lote.

3 Papel de los materiales de referencia en la ciencia de la medición

La metrología es el campo del conocimiento relacionado con la medición. La metrología o ciencia de la medición² incluye todos los aspectos, tanto teóricos como prácticos, con relación a las mediciones, cualquiera que sea el nivel de exactitud, y cualquiera que sea el campo de la ciencia o la tecnología donde ellos ocurran^[6]. Esta sección describe el papel de los materiales de referencia en mediciones cuantitativas.

3.1 Papel de los materiales de referencia en el almacenamiento y transferencia de información o valores propios

Por definición (ver 2.1), un material de referencia tiene una o más propiedades, cuyo(s) valor(es) está(n) bien establecido(s) por mediciones. Cada valor(es) propio(s) de un MR particular ha(n) sido establecido(s) y "almacenado(s)" por el MR (hasta su fecha de expiración) y son transferidos cuando el mismo MR es conducido de un lugar a otro. Por extensión el valor propio de un MR puede ser determinado con una incertidumbre bien definida, por tanto el valor propio puede ser usado como un valor de referencia para propósitos de intercomparación o de transferencia. De aquí que los MR ayuden en la transferencia de mediciones, en tiempo y espacio, de forma similar a los instrumentos de medición³ y a las medidas materializadas^[6].

Un esquema general para construir un sistema jerárquico de medición se muestra en la sección 6.5 del *Vocabulario de Metrología Legal*^[6]. La concatenación de varios niveles y posiciones en un sistema de medición mediante "patrones de referencia" puede, en principio, efectuarse tanto por instrumentos de medición como por medidas materializadas o materiales de referencia.

Un MR tendrá que ser adecuado para el papel exacto para el cual se concibió en cuanto a almacenamiento y transferencia de información de los valores propios medidos. El siguiente criterio técnico (también puede ser adecuado el criterio legal o comercial) se aplica en general para la idoneidad del propósito de los MR:

- a) el MR mismo y los valores propios incorporados deberán ser estables por un espacio de tiempo aceptable, bajo condiciones reales de almacenamiento, transportación y uso;
- b) el MR debe ser suficientemente homogéneo para que los valores propios medidos en una porción del lote, puedan ser aplicados en cualquiera otra porción del lote, dentro de límites acepta-

2) "Ciencia de la medición" es el sinónimo de "metrología" de acuerdo con la definición internacional del último término^[6], puede notarse, sin embargo, que generalmente el uso corriente del término "metrología" se restringe a mediciones físicas de gran exactitud. El término "metrología" está sin embargo, siendo usado cada vez más en el contexto de las mediciones químicas, ingenieriles, médicas y biológicas.

3) Algunos instrumentos de medición no son realmente transportables (por razones de tamaño, masa, fragilidad, inestabilidad o costo), en cualquier caso el objeto a medir debe ser llevado al instrumento para efectuar la transferencia de medición. Pero todos los MR y medidas materializadas son realmente transportables y así pueden ser llevadas hasta el objeto a medir.

bles de incertidumbre; en casos de no homogeneidad de un lote grande, puede ser necesario certificar cada unidad del lote separadamente;

- c) los valores propios del MR deben haber sido establecidos con una precisión y una exactitud suficientes para el fin en que van a ser utilizados;
- d) la documentación clara relacionada con el MR y sus valores propios establecidos, debe estar asequible. Preferentemente los valores propios deben haber sido certificados, así la documentación debe entonces incluir un certificado, preparado de acuerdo con la NC ISO Guía 31^[3].

La palabra "exactitud" fue usada deliberadamente en c) para indicar que, siempre que sea posible, la medición de un valor propio dado debe haber sido hecha por un método que tenga errores sistemáticos despreciables o sesgo relativo a los requerimientos del uso final (o donde el resultado ha sido corregido por un sesgo conocido) y por medio de instrumentos de medición o medidas materializadas, las cuales son trazables a los patrones de medición nacionales. El uso posterior de un MR con valores propios trazables asegura que la trazabilidad se reproduce hasta el usuario. Puesto que la mayoría de los patrones de medición nacionales están armonizados internacionalmente, resulta que los patrones de medición de un país pueden ser compatibles con mediciones similares en otro país. En muchos casos, los MRC son adecuados para las intercomparaciones de patrones de medición nacionales.

3.2 El papel de los materiales de referencia en el Sistema Internacional de Unidades (SI)

3.2.1 Dependencia de las unidades básicas del SI con sustancias y materiales

La mayoría de las mediciones hechas en el mundo actual están dentro del marco del Sistema Internacional de Unidades.^[7] En este momento, el SI reconoce siete unidades básicas, a saber unidades de longitud (metro, símbolo m), masa (kilogramo, kg), tiempo (segundo, s), corriente eléctrica (ampere, A), temperatura termodinámica (kelvin, K), cantidad de sustancia (mole, mol) e intensidad luminosa (candela, cd). Las definiciones^[7] de estas unidades básicas aluden las siguientes sustancias: kriptón-86⁴ (para definir el metro), platino-iridio (para la fabricación del prototipo del kilogramo), cesio-133 (para la definición del segundo), agua (para la definición del kelvin) y carbono-12 (para la definición del mole). Las opiniones difieren en cuanto a si las sustancias designadas están sujetas a la definición de material de referencia (2.1). El uso de estas sustancias en la metrología básica es consecuente con el uso de materiales de referencia en otros tipos de aplicaciones de la medición.

Ciertamente, tales materiales tienen un status especial como sustancias definidas sobre las cuales se basa el SI. La dependencia se aplica estrictamente a la definición de unidad, ya que la obtención de las unidades puede involucrar a otras sustancias/materiales. Esto es especialmente cierto en lo que respecta a la obtención del mole^[8] y del kilogramo.

3.2.2 La obtención de unidades derivadas del SI con la ayuda de materiales de referencia

De las 7 unidades básicas se obtiene un ilimitado número de unidades derivadas del SI combinando unidades básicas como productos y/o cocientes. Por ejemplo, una unidad derivada de concentra-

⁴) Recientemente, la Conferencia General de Pesas y Medidas redefinió el metro como la distancia recorrida por la luz en el vacío durante 1/299 792 458 de un segundo.

ción másica está definida como $\text{kg}\cdot\text{m}^{-3}$ y la unidad derivada de presión (a la cual se ha dado el nombre especial de pascal, símbolo Pa) está definida como $\text{m}^{-1}\cdot\text{kg}\cdot\text{s}^{-2}$. Formalmente hablando, las unidades derivadas últimamente dependen de las sustancias de las cuales dependen las unidades básicas mismas (ver 3.2.1). En la práctica, las unidades derivadas a menudo no se obtienen de las unidades básicas, sino de los MR con valores propios aceptados. Así una variedad de sustancias/materiales puede estar involucrada en la obtención de unidades derivadas (ejemplos 1 y 2 siguientes) o incluso de las unidades básicas (ejemplos 3 y 4 siguientes).

Ejemplo 1: La unidad SI de viscosidad dinámica, el pascal segundo ($\text{Pa}\cdot\text{s} = \text{m}^{-1}\cdot\text{kg}\cdot\text{s}^{-1}$) puede ser obtenida^[9] tomando el valor de una muestra de agua bien purificada como 0,001 002 Pa.s a 20 °C .

Ejemplo 2: La unidad SI de capacidad calorífica molar, el joule por mole.kelvin ($\text{J}\cdot\text{mol}^{-1}\cdot\text{K}^{-1} = \text{kg}\cdot\text{m}^2\cdot\text{s}^2\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$) puede ser obtenido^[10] tomando el valor de la α -alúmina purificada como 79,01 $\text{J}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$ a 25 °C .

Ejemplo 3: La unidad SI de cantidad de sustancia, el mole, puede ser obtenida tomando 0,069 72 kg de galio metálico altamente purificado^[11].

Ejemplo 4: La unidad SI de temperatura, el kelvin, puede ser obtenida a cualquier temperatura T_1 ($273,15 \text{ K} < T_1 < 903,89 \text{ K}$) a partir de mediciones de la resistencia de un alambre de platino de alta pureza a T_1 , al punto triple del agua purificada, al punto de solidificación del estaño purificado, y al punto de solidificación del zinc purificado, asociado a una relación matemática específica^[12] La palabra “termodinámica” ha sido omitida deliberadamente para evitar la controversia sobre si las temperaturas termodinámicas son, o no son, las mismas que las Temperaturas Prácticas Internacionales de 1968: la intención del Comité Internacional de Pesas y Medidas fue armonizar los dos tipos de temperatura exactamente, en el marco de conocimientos disponibles durante 1968-1975.

3.2.3 Relación de la química analítica con el Sistema Internacional de Unidades

Se notará que las sustancias químicas purificadas (frecuentemente llamadas “puras”) fueron citadas en los ejemplos del 1 al 4, (ver 3.2.2). La medición del grado de pureza, o más generalmente de la composición química de los materiales, está dentro del dominio de la química analítica. Además de la dependencia del SI con relación a las sustancias químicas, es digna de examen la dependencia de la química analítica con relación al SI. Actualmente la mayor parte de los químicos analíticos emplea las unidades del SI (todas las unidades básicas, excepto la candela y también muchas unidades derivadas) en sus mediciones. Sin embargo, el análisis composicional depende de un concepto adicional, esto es que las especies químicas puras existen para las cuales las composiciones químicas de otras sustancias o materiales están referidas al invocar las leyes de cambio químico y la estequiometría.

De una o más especies químicas puras, consideradas como patrones de medición primarios, es posible construir jerarquías de medición para la química analítica similares a las usadas en las mediciones físicas^[6] . Ejemplos de tales patrones de medición, son:

a) el electrón, al cual otras especies pueden estar relacionadas por análisis electroquímico^[13] ;

- b) el carbono-12, al cual otras especies pueden en principio estar relacionadas por espectrometría de masa, mediciones según la ley de Raoult, o mediciones volumétricas con gases de baja densidad, etc;
- c) un elemento o compuesto altamente purificado, al cual otras sustancias pueden estar relacionadas por métodos electroquímicos, gravimétricos, titrimétricos, espectrométricos, etc.

Las “otras especies” citadas en estos ejemplos podrán ser usadas en muchos casos como MR. Muchas sustancias podrán desempeñar este papel de intermediarios entre patrones primarios y patrones analíticos de trabajo, usando la diversidad de técnicas y reacciones químicas que un analista puede emplear. El concepto de trazabilidad se aplica a la química analítica tal como se hace en otras ramas de la ciencia de la medición. La calidad de los resultados de un análisis químico será elevada si la trazabilidad del resultado puede estar evidentemente establecida en términos de la trazabilidad de los instrumentos, medidas materializadas y MR empleados. En la mayoría de los casos, la trazabilidad también dependerá de los valores de las masas atómicas relativas (antes llamadas “pesos atómicos”) usadas en los cálculos; la fuente de estas debe ser registrada por los analistas (por ejemplo [11]).

3.2.4 El papel de los materiales de referencia en la obtención de unidades fuera del SI

Donde los componentes de un sistema de medición (por ejemplo, el Sistema Imperial) puedan ser relacionados exactamente a los componentes correspondientes del SI, no es necesario tener medios independientes para la obtención de sistemas de medición no pertenecientes al SI. Donde las magnitudes no puedan ser relacionadas a aquellas del SI, entonces en principio es necesaria la obtención independiente de las unidades no pertenecientes al SI. En la práctica, sin embargo, algunos de tales sistemas permanecen en uso y por eso son curiosidades históricas principalmente.

3.3 Uso de los materiales de referencia

REMCO pretende publicar una guía separada que cubra los principios generales y estadísticos para el uso de los materiales de referencia. Existen muy pocos documentos publicados que hablen de los problemas generales asociados con el uso de los materiales de referencia. El lector es remitido a los documentos y recomendaciones publicados por la Comisión 1.4 de la IUPAC sobre Patrones y Materiales de Referencia Físico-químicos, la cual está de acuerdo primeramente con el uso de materiales de referencia para la obtención de propiedades físicas. Las siguientes publicaciones de la Comisión 1.4 de la IUPAC en la *Química Pura y Aplicada*, están comprometidas con la certificación y el uso de los materiales de referencia para propiedades físicas:

Propiedad física	Volumen, fecha de publicación y número de la página	
Entalpía	40	1974 : 399
Rotación óptica	40	1974 : 451
Refracción óptica	40	1974 : 463
Densidad	45	1976 : 1
Masa molecular relativa	48	1976 : 241
Absorbancia y longitud de onda	49	1977 : 661
Reflectancia	50	1978 : 1 477
Actividad iónica potenciométrica	50	1978 : 1 485
Viscosidad	52	1980 : 2 393
Permitividad	53	1981 : 1 847
Conductividad térmica	53	1981 : 1 863

4 Incertidumbre de medición

En la discusión sobre incertidumbres de medición, son usados por lo general los términos “precisión”; “error sistemático o sesgo” y “exactitud”. Los significados de estos términos no son inflexiblemente fijos, sino dependen en gran medida de la interpretación y el uso de los datos^[14, 15].

4.1 Un ejemplo ilustrativo

Si dos operarios igualmente entrenados, A y B, realizan cada uno 4 réplicas de una medición de un material uniforme cada día, durante 4 días en un instrumento, y 4 días nuevamente en un instrumento similar, los resultados de las 16 observaciones de cuatro mediciones pueden semejar a los del gráfico 1. ¿Qué se puede apreciar de este gráfico?.

- las dispersiones de cada juego de observaciones de 4 valores son comparables, quizás ligeramente menor para el instrumento 2 que para el instrumento 1;
- se manifiesta que hay mayor variabilidad entre resultados los diarios, que entre los juegos de resultados diarios, particularmente para el instrumento 1;
- el operario B presenta resultados más bajos que el operario A;
- el instrumento 1 brinda resultados más bajos que los del instrumento 2.

La Figura 1 está confeccionado con el propósito de una demostración y las mediciones reales pueden ser mejores o peores que las mostradas. Sin embargo, este proyecto muestra algunos de los 4 tipos de factores que contribuyen a la variabilidad total de estas mediciones:

- factores que actúan dentro de los días;
- factores que actúan entre los días;
- factores debidos a los sistemas de instrumentos;

4) factores debidos a los operarios.

Están a disposición técnicas apropiadas para estimar por separado los efectos de estos cuatro factores y puedan ser calculadas las desviaciones típicas correspondientes a cada una de ellas. Sin embargo, el número limitado de operarios e instrumentos impide el cálculo de las desviaciones típicas de un modo confiables para los factores 3) y 4) como para los factores 1) y 2). El tiempo y el trabajo involucrados sin duda imponen límites sobre cualquier esfuerzo para hacerlo así.

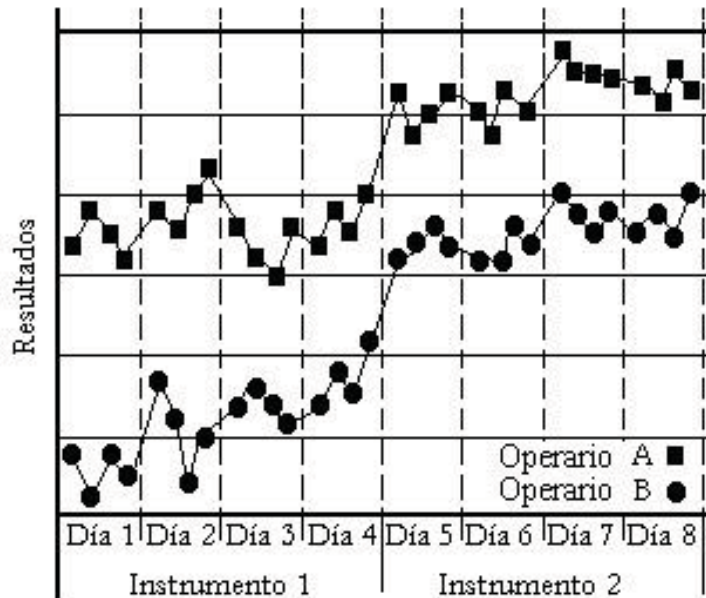


Figura 1. Un ejemplo de los resultados de mediciones de dos operarios usando dos instrumentos en ocho días diferentes.

El omitir tener en consideración factores relacionados con los instrumentos y operarios, es una de las principales causas de diferencias no razonables, usualmente encontradas entre los tipos de ensayos interlaboratorios, o circulares^[16]. Debido a que los instrumentos varían de tiempo en tiempo y los operarios cambian, el resultado de un laboratorio en un tiempo dado, representa sólo uno de los muchos resultados que pueden ser obtenidos, y la variabilidad causada por estas dos fuentes debe ser considerada como parte de la precisión del laboratorio. La desviación típica calculada sin estimar estos efectos, pudiera subestimar la variabilidad verdadera.

Si por el propio uso de los patrones y métodos de referencia^[17], estas dos fuentes de errores fueran eliminadas, la desviación típica calculada a partir de las 16 medias obtenidas de los juegos de observaciones de cuatro mediciones será la propia medida de la precisión. Presumiblemente la gran media de los 16 valores medios será reportada.

La media de muchos valores es más estable que las mediciones individuales. Cuando fuentes externas de variación, tales como los efectos del instrumento y del operario, son eliminadas, la relación entre la desviación típica de las mediciones individuales y la desviación típica de la media de n de tales mediciones puede ser expresada como:

$$\sigma(\bar{X}_n) = \frac{\sigma(X)}{\sqrt{n}} \quad \dots(1)$$

En otras palabras, la desviación típica de la media es más pequeña que la desviación típica de las mediciones individuales con factor de $1/\sqrt{n}$. Una importante condición deberá mantenerse para que esta relación sea verdadera, es decir, que las n mediciones sean independientes unas de otras. La "independencia" puede ser definida en un sentido de probabilidad, pero para los propósitos actuales, las mediciones pueden ser consideradas independientes si ellas no muestran una tendencia o un contorno. Esto verdaderamente no es cierto en el figura 1, y decir que la desviación típica de la media de los 64 valores es $1/8$ ($= 1/\sqrt{64}$) de la desviación típica de las mediciones individuales pudiera subestimar seriamente su variabilidad real. Además, la relación en la ecuación (1) está expresada en términos del valor verdadero de la desviación típica, σ , la cual es usualmente desconocida. Como desviación típica calculada, s , es ella misma un estimado de σ del conjunto de valores medidos, la desviación típica de la media en la ecuación (1) es sólo aproximada cuando s es usada en lugar de σ .

El uso de la desviación típica calculada de los promedios diarios se prefiere sobre los valores individuales, porque lo primero refleja correctamente un componente de variabilidad entre días o de tiempo, lo cual está usualmente presente en las mediciones de precisión.

4.2 Algunos conceptos estadísticos básicos

La información básica disponible sobre los errores de medición está resumida por:

- a) el número de determinaciones independientes o el número a partir del cual fue calculada y reportada la media;
- b) un estimado de la desviación típica, s , definida por:

$$s = \left[\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \right]^{1/2}$$

donde los resultados de n mediciones son denotados por x_1, x_2, \dots, x_n , y su media es:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

A partir de a) y b) pueden ser calculadas algunos estadísticos derivados útiles.

c) desviación típica de la media de n mediciones

$$s(\bar{x}_n) = \frac{s}{\sqrt{n}}$$

Esto es llamado algunas veces error típico de la media, para diferenciarlo de la desviación típica de mediciones individuales.

NOTA Como n viene a ser un número grande, el valor de $s(\bar{x}_n)$ viene a ser muy pequeño, mostrando que el promedio de un número grande de mediciones se aproxima a un valor constante μ que es usualmente el objetivo del procedimiento de medición.

d) el intervalo de confianza para la media (distribución normal). Cada n mediciones hechas, se reporta el valor promedio de las mismas. Estos promedios diferirán, de tiempo en tiempo, dentro de ciertos límites. Asumiendo una distribución normal, un intervalo del tipo $\bar{x} \pm \delta$ puede ser creado[18] de forma tal, que el intervalo desde $\bar{x} - \delta$ hasta $\bar{x} + \delta$ será razonablemente seguro para incluir el valor de μ deseado. El intervalo se calcula por:

$$\delta = t \frac{s}{\sqrt{n}} \quad \dots (2)$$

donde t es un valor tabulado de la distribución de Student y depende del nivel de confianza y los grados de libertad para s ;

e) límites de 2-sigma (ó 2s), 3-sigma (ó 3s). Estos límites describen la distribución del error de medición. Si una medición es hecha por el usuario de un MRC teniendo la misma precisión (es decir, la misma σ) como la obtenida por el laboratorio certificador, sus mediciones deben caer (con una probabilidad aproximadamente de 0,95 a 0,997) dentro de estos límites cuando σ está bien establecida. Por otra parte, existe evidencia de diferencia sistemática.

4.3 Errores de instrumentos y operarios

Los tipos de errores de instrumentos y operarios aún no han sido tratados. Una situación ideal pudiera ser eliminarlos del proceso de medición o usar más instrumentos y más operadores, y entonces estimar las desviaciones típicas inherentes a estas fuentes. Cuando ninguna de las dos anteriores observaciones puedan ser cumplidas, entonces al menos podrá ser hecho si usamos 2 instrumentos y/o operadores. Si el intervalo de confianza para la media de los resultados de los dos instrumentos no se solapan, entonces existe una buena evidencia de diferencia de instrumentos.

Usando su experiencia y criterio, un científico de la medición puede llegar a límites razonables para estos tipos de errores. Si el límite no ha sido calculado de los datos de la medición, entonces su validez no puede ser sustentada por el análisis estadístico. En tales casos, estos límites son "suestimados" y el único recurso es tratarlos como límites de errores sistemáticos.

La detección de diferencias y la diferenciación de la variabilidad total dentro de sus componentes identificables pueden facilitarse a través de un cuidadoso control y diseño estadístico del experimento.

4.4 Diferencias entre los métodos de medición

Cada método de medición significa medir la propiedad deseada de un material, pero raramente un método hace la medida de la propiedad directamente. En muchos casos el método realmente mide alguna otra propiedad que está relacionada con la propiedad en cuestión por teoría, práctica o tradición y entonces es convertido al valor de la propiedad deseada a través de estas relaciones. Las discrepancias entre los resultados de diferentes métodos de medición son comunes aún para mediciones que llevan a la determinación de constantes físicas fundamentales^[19].

En la preparación de un MRC, usualmente 2 ó más métodos de medición se emplean para cada propiedad medida. Si estos métodos están bien establecidos en virtud de experiencias anteriores, los resultados obtenidos por estos métodos concuerdan usualmente con la incertidumbre asignada a cada método.

En algunos casos, estas diferencias son tan grandes, que los resultados no pueden ser conciliados y estos resultados son reportados separadamente para cada método individual. El MR es por tanto no certificado o certificado en base a un método básico dependiente. Un ejemplo histórico de este tipo de reporte lo tiene el NBS MRC 1091, de acero inoxidable. El nitrógeno contenido fue medido por fusión al vacío y por destilación a presión reducida, dando los resultados de 861 y 945 mg/kg, con desviación típica de 3 y 20 mg/kg, respectivamente. Claramente, uno o ambos métodos tiene un error sistemático mayor, comparado con la variabilidad del material o con la incertidumbre de la medición. Un reporte del promedio de los dos métodos, podrá ser altamente engañoso.

La exactitud de la medición, en su sentido absoluto, nunca es realizada. En la práctica, los valores certificados de algunos materiales de referencia están definidos por el uso de un método de arbitraje o asignación de un valor por un procedimiento bien definido de forma que al menos la misma marca exhibida pueda ser usada por cualquiera en ese campo. La importancia de los métodos de referencia para complementar el uso de estos patrones de medición, están siendo también enfatizados^[17]. Un buen ejemplo es el método de referencia de hemoglobina en sangre y el valor asignado como marca exhibida al material de referencia establecido por el Comité Internacional para la Normalización en Hematología (ICSH)^[20 y 21].

4.5 Incertidumbre de valores certificados

La incertidumbre de los valores de un MRC está usualmente hecha de varios componentes, algunos sustentados por datos y otros no:

- a) un intervalo de tolerancia estadística, dando límites a materiales no homogéneos basados sobre datos o cálculos estadísticos;
- b) un intervalo de confianza para la media, dando límites al error de medición basado en datos o cálculos estadísticos;
- c) componentes de la incertidumbre de la medición debido a la variación entre laboratorios y/o operarios y métodos de medición.

- d) una combinación (adición de valores absolutos o la raíz cuadrada de la suma de los cuadrados) de límites estimados para fuentes “conocidas” de posibles errores sistemáticos, basados en la experiencia y los criterios (en otras palabras, donde no hay datos, o un insignificante número de datos, para hacer cálculos estadísticos).

La palabra “conocidas” está entre comillas, para contrastar con los errores sistemáticos que son “desconocidos” o insospechados. Estos errores insospechados podrán ocurrir en un número de formas; un componente en un sistema físico, un defecto menor en una consideración teórica, un error de redondeo en un cálculo. A medida en que más materiales homogéneos estén disponibles y sean desarrollados métodos de medición más precisos, estos tipos de errores serán detectados por diseño o por cambios y seguramente serán eliminados. La exactitud mejorada en la medición de una propiedad es básicamente un proceso iterativo caro y una demanda no garantizada de exactitud, pudiera significar un derroche de recursos.

4.6 Informe de la incertidumbre en los certificados de MRC

Una variedad de informes sobre la incertidumbre pueden ser encontrados en certificados pasados y presentes emitidos para MRC en todo el mundo. Algunos de estos informes están bien formulados y sustentados por datos, otros no; algunos de estos informes contienen un caudal de información que es útil y para usuarios muy precisos pero muy abrumadoras para otros, algunos de estos informes están tan supersimplificados que pierden información. Debido a que el productor del MRC, tiene que mantener a toda clase de usuarios en mente, el uso de una forma simple de informe no le es usualmente posible. La intención es que todos estos informes no sean ambiguos, sean significativos y que contengan toda la información relevante para los usuarios potenciales.

Algunos de los informes usados comúnmente son tomados de certificados existentes y están relacionados en 4.6.1 hasta 4.6.4.

4.6.1 Ejemplo 1: Límites de confianza de 95 % para la media

Cloruro de rubidio

Relación de frecuencia absoluta..... 2,593 ± 0,002

“La incertidumbre indicada son todos los límites de error basados en el 95 % del límite de confianza para la media y asignados para los efectos de fuentes conocidas de posibles errores sistemáticos”.

Porque la relación isotópica es una constante para un lote dado de material y no está sujeto a errores de no homogeneidad del material, un límite de confianza del 95 % para la media se refiere solamente al error de medición. Este se calcula a partir de:

$$t \frac{s}{\sqrt{n}}$$

tal como se describe en la ecuación (2).

Los efectos de fuentes conocidas de un posible error sistemático se exponen en detalle en “Relación de frecuencia isotópica absoluta” y peso atómico del rubidio terrestre”^[22]

4.6.2 Ejemplo 2: Límites de 2-sigma ó 3-sigma

Filtros de vidrio para espectrometría de absorción molecular

Absorbancia $0,500\ 0 \pm 0,002\ 5$

“La incertidumbre es la suma del error aleatorio de $\pm 0,1\ %$ relativo (límite 2σ) y de los sesgos estimados, los cuales son $\pm 0,4\ %$ relativo”.

Cada filtro de vidrio fue calibrado individualmente, y la desviación típica se refiere al error de medición, incluyendo la limpieza de la superficie. Como estos filtros de vidrio serán usados de tiempo en tiempo, un múltiplo de la desviación típica es una medida idónea de la variabilidad.

4.6.3 Ejemplo 3: Incertidumbre expresada en cifras significativas

Acero AISI 4340

Fracción másica elemental

Carbono $3,8_2 \times 10^{-3}$

Manganeso $6,6 \times 10^{-3}$

De acuerdo con la explicación dada en el texto: “No se espera que el valor relacionado espera se desvíe del valor verdadero en más de ± 1 en la última cifra significativa reportada; para una cifra como subíndice, no se espera que la desviación sea mayor de ± 5 ”. Por tanto, la fracción másica del carbono, expresada en por ciento, se halla entre 0,377 y 0,387 y la del manganeso está entre 0,65 y 0,67. Estas incertidumbres incluyen material no homogéneo, mediciones imprecisas y posible sesgo entre laboratorios y redondeo implícito, porque estos valores son “... la mejor estimación presentada del valor verdadero basado en los resultados de un programa analítico eficaz interlaboratorios”.

Cuando 20 ó 30 elementos son certificados de un solo material, este método da un resumen conveniente y conciso de los resultados. Como estos límites están expresados en unidades de 5 y 10, se pierde inevitablemente alguna información para algunos de estos elementos. Sin embargo, cuando se usa el valor certificado, es importante usar todos los dígitos dados, incluyendo los subíndices. La incertidumbre establecida en este certificado depende grandemente del uso del criterio químico.

4.6.4 Ejemplo 4: Desviación típica y número de determinaciones

Método	Oxígeno en metales ferrosos (µg/g)		
	MRC A (Lingote de hierro)	MRC B (Acero inoxidable: AISI 431)	MRC C (Acero fundido al vacío)
Fusión al vacío			
\bar{x}	484	131	28
s	14	8	2
n	216	286	105
Activación de neutrones			
\bar{x}	492	132	28
s	28	7	4
n	6	6	5
Fusión en gas inerte			
\bar{x}	497	129	29
s	13	8	5
n	12	11	20

donde:

\bar{x} valor medio del oxígeno

s desviación típica de una determinación individual

n número de determinaciones.

NOTA La desviación típica incluye tanto los errores debidos a la imprecisión del método analítico como los debidos a la posible heterogeneidad del material analizado.

Un criticismo contra este modo de presentación es que el usuario tendrá que calcular la incertidumbre basado en su propia comprensión de las relaciones.

5 Homogeneidad de los materiales

Muchos MR son sometidos a procedimientos de preparación, los cuales finalmente incluyen una subdivisión en unidades usables. Un subconjunto de unidades individuales de un lote se selecciona para las mediciones, de acuerdo con un plan de muestreo estadísticamente válido. Una incertidumbre de medición es derivada tomando en cuenta la no homogeneidad del material así como otros factores (véase sección 4.). Otros tipos de MR son preparados como mecanismos individuales y la certificación se basa en la medición por separado de cada unidad mejor que en el muestreo estadístico del lote completo. La segunda variante es provechosa cuando el MR puede ser medido no destructivamente.

5.1 Materiales

Los MR preparados como soluciones o compuestos puros se espera sean homogéneos en terrenos físicos (termodinámicos). El objetivo del ensayo para la homogeneidad es principalmente detectar cualquier impureza, interferencia o irregularidad.

Materiales tales como mezclas de polvos, minerales, aleaciones, etc. son por naturaleza heterogéneos en su composición. Por tanto, los MR preparados a partir de tales materiales deben ser ensayados para estimar el grado de homogeneidad.

5.2 Concepto de homogeneidad

En la teoría, un material es perfectamente homogéneo con respecto a una característica dada si no existe diferencia entre el valor de esta característica desde una parte (unidad) a otra. Sin embargo, en la práctica un material es aceptado como homogéneo con respecto a una característica dada, si la diferencia entre el valor de esta característica de una parte (unidad) a otra no puede ser detectada experimentalmente. El concepto práctico de homogeneidad por tanto encierra, tanto una especificidad de la característica, como un parámetro de medición (usualmente la desviación típica) del método de medición usado, incluyendo el tamaño de muestra definido de la porción de ensayo.

5.2.1 Características de interés

Un material puede ser suficientemente homogéneo con respecto a la característica de interés a ser útil como MR aún cuando no sea homogéneo con respecto a otras características, previendo que esta falta de homogeneidad no ejerza una influencia detectable en la exactitud y la precisión de los métodos usados comúnmente en la determinación de la característica de interés.

5.2.2 Método de medición de la homogeneidad

El grado de homogeneidad que debe tener un material para ser usado como MR es medido con la precisión alcanzable con los mejores métodos asequibles para la determinación de la característica para la cual el MR es proyectado. Por tanto, a mayor precisión del método de medición, más alto es el grado de homogeneidad requerido del material.

La precisión alcanzable por el método de medición de la homogeneidad varía con la característica medida y su valor para el MR. Un MR proyectado para más de una característica se describe por un número correspondiente de informes de homogeneidad, cada uno de las cuales debe ser trazable a una precisión determinada experimentalmente. La magnitud de la precisión puede variar ampliamente.

En muchos casos, la precisión alcanzable por un método de medición está afectada por el tamaño de la porción de ensayo tomada del MR. El grado de homogeneidad de un MR está por tanto definido para un tamaño de la porción de ensayo determinada.

5.2.3 Práctica

Idealmente, un MR puede ser caracterizado con respecto al grado de homogeneidad para cada característica de interés. Para MR preparados para un número relativamente grande de características, la valoración del grado de homogeneidad para todas las características es gravoso, tanto física como económicamente y en algunos casos no es factible. En la práctica, por tanto, el grado de homogeneidad de tales MR es valorado sólo para características seleccionadas. Es recomendado que estas características sean apropiadamente seleccionadas sobre la base de relaciones químicas y físicas establecidas; por ejemplo, una concomitancia de interelementos en las fases de mineral de un MR hace razonable el asumir que el MR también tiene un aceptable grado de homogeneidad para los elementos no seleccionados.

5.3 Diseño experimental

5.3.1 Objetivos

Para materiales de referencia de los cuales se espera sean homogéneos en el terreno físico, el propósito principal del ensayo de homogeneidad es detectar problemas no esperados. Algunos ejemplos son la contaminación diferencial durante el empaque final en unidades individuales, o la incompleta disolución o equilibrio de un analito en el solvente (el cual pudiera conducir firmemente a cambios de concentración desde el primer frasco llenado hasta el último). Un análisis direccional estadístico pudiera ayudar en último caso. Si el material es producido en más de un lote, es necesario ensayar la igualdad de los lotes (o certificar los lotes separadamente).

Cuando la naturaleza de un material de referencia conduce a esperar que presente alguna falta de homogeneidad, el fin del programa de ensayo no es una simple detección de la no homogeneidad, sino la estimación de su magnitud. Esto puede requerir de un programa de ensayo mucho más extenso que el requerido para la detección.

La no homogeneidad puede manifestarse al menos en dos formas:

- a) diferentes submuestras de una unidad de MR pueden diferir en la propiedad de interés.
- b) puede haber diferencias entre unidades del MR.

Diferencias entre submuestras pueden usualmente ser reducidas o controladas a un nivel aceptablemente bajo haciendo suficientemente grande el tamaño de la submuestra. A menudo se realiza un estudio para determinar el tamaño de submuestra apropiado antes de comenzar la certificación. Las diferencias que existen entre unidades individuales de un candidato a MR deben ser reflejadas en el informe de incertidumbre del certificado.

En términos estadísticos, el diseño experimental debe satisfacer los siguientes objetivos:

- 1) detectar si dentro de la unidad (rango corto) la variación es estadísticamente significativa en comparación con la variación conocida del método de medición.

- 2) detectar si entre unidades (rango largo) la variación es estadísticamente significativa en comparación con la variación dentro de la unidad.
- 3) determinar si la significación estadística detectada dentro de la unidad o entre unidades, indica una significación física correspondiente de suficiente magnitud para descalificar al candidato a MR para el uso concebido.

Debe ser conocido el grado de homogeneidad de un candidato a MR en su forma final. No obstante, la tarea de valoración de la homogeneidad puede ser realizada en diferentes etapas.

5.3.2 Ensayo preliminar para homogeneidad

Una valoración preliminar de la homogeneidad de un candidato a MR puede ser realizada después de la homogeneización como una parte integral del proceso de preparación. Las propiedades físicas de un MR que pueden causar segregación al ocurrir, por ejemplo el tipo de mezclador, influyen fuertemente la manera de selección de la muestra. Las muestras deberán ser tomadas en regiones donde las diferencias físicas se espera que ocurran. Un muestreo aleatorio deberá ser adoptado solamente cuando las causas de diferencias físicas sean desconocidas o se crean ausentes.

El número de muestras tomadas y las determinaciones replicadas deben ser tales que el ensayo estadístico apropiado debe ser capaz de detectar la posible existencia de no homogeneidad a un nivel predeterminado.

NOTA La ASTM E 826-81 "Práctica normalizada para ensayo de homogeneidad de materiales para el desarrollo de materiales de referencia", brinda un procedimiento detallado para ensayos de homogeneidad en materiales a granel. Esta práctica normalizada está especializada para los casos de ensayo de homogeneidad de metales, tanto en forma sólida como en polvo, y óxidos de minerales finamente pulverizados que se pretenden usar como materiales de referencia en emisión de Rayos X, o Espectroscopía de emisión óptica, o ambos. Para la mayoría de los programas de certificación de MR puede obtenerse un ensayo preliminar apropiado para la homogeneidad por simple adaptación de la práctica dada en ASTM E 826-81.

5.3.3 Ensayo principal para homogeneidad

Este ensayo debe ser realizado al candidato a MR después que el mismo ha sido envasado en la forma final, sin tener en cuenta cualquier ensayo preliminar hecho anteriormente. El propósito de este ensayo es confirmar que la variación entre unidades no es significativa estadística y prácticamente.

Las unidades deberán ser seleccionadas de una reserva, de forma aleatoria, para que cada unidad tenga la misma posibilidad de elección. Debe usarse un diseño experimental en el cual k unidades del material sean seleccionadas y n réplicas de las determinaciones sean realizadas para cada unidad. Se recomienda que las determinaciones sean hechas en orden aleatorio, para evitar posibles variaciones sistemáticas con relación al tiempo. k y n deberán ser lo suficientemente grandes para detectar la posible existencia de no homogeneidad a un nivel predeterminado.

Para ciertos MR, no son posibles las determinaciones replicadas dentro de una unidad porque el uso de la unidad completa es lo que prescrito por el productor. En ese caso, la varianza entre unidades debe ser comparada con la precisión estimada del método de medición para valorar el grado de homogeneidad del MR.

5.4 Resultados posibles del ensayo de homogeneidad

La selección de las muestras y el análisis de los datos son usualmente realizados en consulta con un estadístico. Dependiendo de la forma del material, el énfasis debe ser detectar tendencias o modelos, por ejemplo, de un extremo a otro de una varilla de acero; desde el centro hasta la punta de una placa; desde el tope superior de una porción del material a granel en un tambor; o chequear la variabilidad de material contenido en ámpulas o botellas. Un adecuado experimento diseñado estadísticamente ayuda a asegurar que las conclusiones sean válidas y minimiza el número de mediciones necesarias para alcanzar tales conclusiones.

Los resultados posibles del ensayo de homogeneidad se describen en 5.4.1 hasta 5.4.3.

5.4.1 Material muy homogéneo

La homogeneidad no es un problema, o la variabilidad del material es despreciable en relación al error de medición o al uso del MRC. En este caso, el valor certificado es el mejor estimado de la media del valor propio para el lote y la tolerancia permitida para la incertidumbre describe errores posibles de medición asociados con este estimado.

5.4.2 Material muy poco homogéneo

La variabilidad del material es el factor mayor en la incertidumbre total. En este caso, el lote completo del material es rechazado o reelaborado, o cada espécimen es individualmente medido y certificado.

La reelaboración es una forma de proceder razonable cuando existe razón para creer que la fuente de no homogeneidad puede ser eliminada por preparación de un nuevo lote de material, usando procedimientos perfeccionados. Sin embargo, esto no es siempre posible y algunas veces es necesario tolerar una pequeña cantidad de no homogeneidad entre unidades cuando el material prácticamente no puede ser mejorado.

5.4.3 Material moderadamente homogéneo

La variabilidad del material es de la misma magnitud que el error de medición y debe ser incluida como un componente de la incertidumbre. Este caso se discute en 5.5.

5.5 Algunas ejemplos de ensayo de homogeneidad

De los 3 casos (5.4.1 al 5.4.3) el último es uno de los más frecuentemente encontrados. Dos subclases son aparentes, una donde se detecta una tendencia y otra donde no se detecta tendencia.

Donde ha sido detectada una tendencia, por ejemplo, a lo largo de una varilla de acero que va a ser cortada en piezas, la porción inservible es desechada y, por suerte, la tendencia en la porción remanente es lineal o puede ser de otro modo descrita matemáticamente. En tales casos, una línea

(u otra expresión matemática apropiada) puede ser fijada al valor medido a lo largo de la varilla. La máxima desviación de los puntos promedios de la línea fijada es tomada como una medida de la no homogeneidad, asumiendo que el error de medición es pequeño, en comparación con la tendencia.

Donde no se detecta tendencia, pero los resultados de las mediciones muestran una variabilidad que no es despreciable, un concepto estadístico llamado "intervalo de tolerancia estadística" puede ser usado. Para ilustrar este concepto supongamos una solución que es preparada y envasada en 1 000 ámpulas, de las cuales 30 son medidas para la misma propiedad. Para este ejemplo, el concepto de límite de tolerancia^[18] establece esencialmente que basado en los valores medidos de las 30 ámpulas, casi todas las 1 000 no diferirán del promedio de las 30 ámpulas en más del límite formulado. En términos estadísticos, esto podría leerse: "El intervalo de tolerancia (media $\pm \Delta$) está formulado de forma tal que cubre al menos el 95 % de la población con una probabilidad de 0,99"⁵

Esta declaración no garantiza que el intervalo de tolerancia incluya todas las ámpulas. Se dice que el 99 % del tiempo el intervalo de tolerancia incluirá al menos el 95 % de las ámpulas. El "99 % del tiempo" se refiere a la forma en que este intervalo de tolerancia es formulado, es decir, si 30 ámpulas fueron seleccionadas de una población repetidamente, y los mismos experimentos fueron realizados una y otra vez, el 99 % de los intervalos de tolerancia así formulado cubriría al menos la proporción (95 %) de la población total como se especifica, y el 1 % de los intervalos de tolerancia cubriría menos que el 95 % de la población total.

¿Cómo es formulado este intervalo?. Primero, se calculan la media (ecuación 3) y la desviación típica (ecuación 4) de las 30 ámpulas:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad \dots(3)$$

$$s = \left[\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \right]^{1/2} \quad \dots(4)$$

donde:

$x_1, x_2, \dots, x_i, \dots, x_n$ son los valores medidos, con $n = 30$;

\bar{x} es un estimado de la media, μ , de las 1 000 ámpulas;

s es un estimado de la medida de la dispersión, σ , entre esas ámpulas

⁵) La declaración es verdadera solamente para una población de tamaño infinito; sin embargo, la corrección para una población de tamaño finito es despreciable cuando el tamaño finito es grande.

Los valores \bar{X} y s contienen prácticamente toda la información asequible sobre las 1 000 ámpulas y pueden ser usados para calcular el intervalo de tolerancia $\bar{X} \pm \Delta$.

El valor de Δ es calculado como un múltiplo de s , es decir $\Delta = k_2' s$.

El valor de k_2' depende de tres parámetros:

- a) el número, n , de muestras medidas (30),
- b) la proporción, p , de la población total a cubrir (0,95);
- c) el nivel de probabilidad, $1 - \alpha$, especificado (0,99).

Una tabla de factores para los límites de tolerancia para ambos lados para distribuciones normales da para k_2' el valor de 2,841 para $n = 30$; $1 - \alpha = 0,99$; y $p = 0,95$. Las tablas de estos factores están dadas en la ISO 3207⁶, y en muchos textos estadísticos normalizados^[18].

El término “ambos lados” significa que estamos interesados, tanto en el límite inferior como en el superior del promedio. El término “distribución normal” se refiere a la distribución de todos los valores de interés y es una distribución en forma de campana, simétrica, usualmente hallada en trabajos de medición de precisión.

La Figura 2 es un histograma de las relaciones de la proporción de emisión de ^{137}Cs , en un material de referencia de ^{137}Cs quemado en combustible nuclear, con respecto a un patrón de referencia de radio. Una curva de frecuencia de una distribución normal puede ser adecuada con los siguientes datos: había 98 ámpulas de ^{137}Cs involucradas; cada ámpula fue medida en abril, septiembre y noviembre de 1972. Promediando las tres mediciones, el error de medición fue considerablemente menor que la diferencia de masas de las soluciones activas entre estas ámpulas y el ploteo en la Figura 2 muestra esencialmente la no homogeneidad de la masa de la solución en las ámpulas.

⁶) ISO 3207 Interpretación estadística de datos. Determinación de un intervalo estadístico de tolerancia

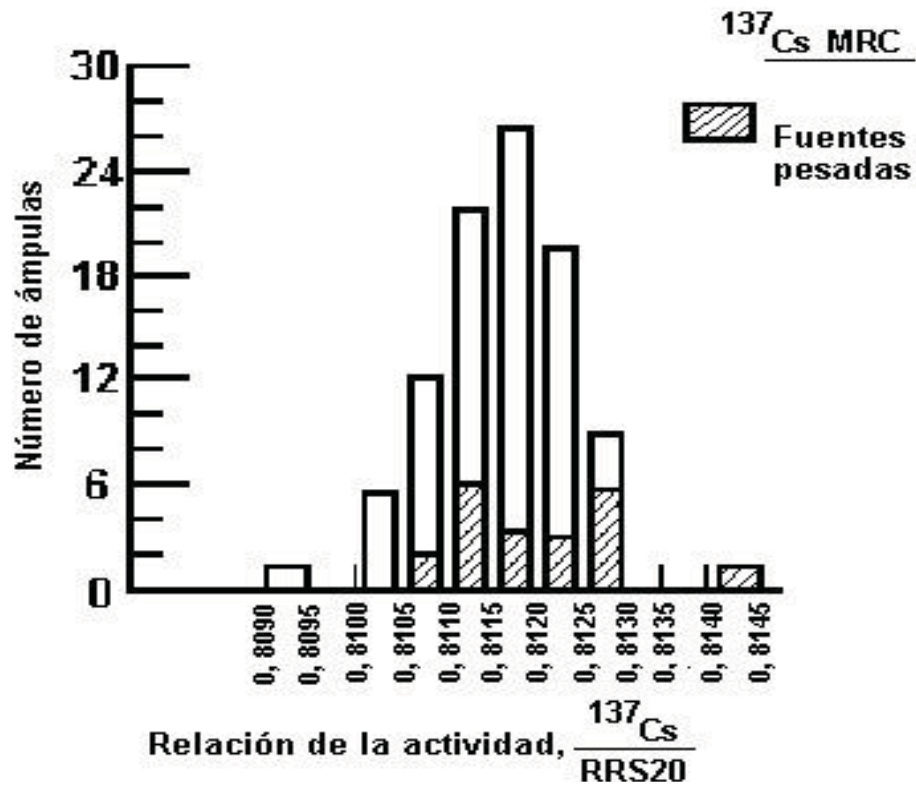


Figura 2. Histograma de la frecuencia (número de ámpulas) contra la relación de la actividad de patrones de ^{137}Cs con el patrón de referencia de radio (RRS20).

Para ilustrar el concepto de las propiedades del intervalo de tolerancia fueron seleccionados los valores de 30 ámpulas de un grupo de 98 valores usando una tabla de números aleatorios, y \bar{x} y s fueron calculadas por las ecuaciones (3) y (4), dando:

$$\bar{x} = 0,811\ 68 \quad s = 0,000\ 80$$

El intervalo de tolerancia para al menos el 95 % cubierto con un nivel de probabilidad de 0,99 es $\bar{x} \pm k_2' s$. Ó $0,811\ 68 - (2,841 \times 0,000\ 80) = 0,809\ 41$ hasta $0,811\ 68 + (2,841 \times 0,000\ 80) = 0,813\ 95$

El intervalo cubre los valores de 96 ámpulas, es decir, todas excepto una sola ámpula en el extremo izquierdo y una en el extremo derecho.

Si los 30 valores fueron seleccionados repetidamente, los valores de \bar{x} y s no serán los mismos, pero los intervalos $\bar{x} - k_2' s$ hasta $\bar{x} + k_2' s$ tendrán la propiedad de cubrir al menos el 95 % de todos los valores el 99 % del tiempo.

Los límites de tolerancia e intervalos son conceptos útiles cuando un número finito de unidades son tomadas de una gran reserva de material. A medida que el número de unidades, n , se incrementa, y $\bar{x} \rightarrow \mu$, $s \rightarrow \sigma$, el valor de k_2 decrece y eventualmente se fija en 2,0 (realmente 1,96) cuando $p = 0,95$. Durante años, sin embargo, algunos científicos de las mediciones se han resistido a usar este concepto. Otros más conservadores usan un factor de 3,0, y algunas veces, usan el intervalo entre el máximo y el mínimo.

6 Principios generales de certificación

Hay un número de vías técnicamente válidas para la certificación de un material de referencia. Las vías válidas incluyen mediciones por uno o más métodos involucrando uno o varios laboratorios. Dependiendo del tipo de material de referencia, de sus requisitos para el uso final, de la calificación de los laboratorios involucrados y de la calidad del método o métodos, una de las vías puede ser escogida como la más apropiada. Dos de los más importantes aspectos del proceso de certificación involucra los conceptos de exactitud e incertidumbre de los valores determinados para las propiedades a certificar.

6.1 Certificación sobre la base de la exactitud

Donde sea técnicamente posible, los MR son generalmente certificados sobre la base de la exactitud.^[23] Por lo tanto, un valor certificado generalmente representa la mejor estimación real del valor "verdadero". En algunos casos, la medición no puede estar en términos de un valor verdadero, por lo tanto se adopta un valor propio asignado para el uso con un método específico. La certificación del MR, en ese caso, no requiere campaña de medición⁷, pero sí de un somero informe sobre el valor asignado y de la técnica de medición pertinente para la cual el MRC resulta ser un calibrador.

No se espera que los valores certificados se desvíen del valor "verdadero" en más de la incertidumbre de medición establecida. La incertidumbre establecida del valor de una propiedad debe tener en cuenta cualquier error sistemático y aleatorio inherente al proceso de medición, así como la variabilidad del material y debe describirlos conjunta o separadamente. Un número de diferentes vías de medición son comúnmente usadas por los certificadores e incluyen lo siguiente:

- a) Mediciones por un método único definitivo (como está definido en la sección 7) en un laboratorio único. Este método es generalmente utilizado por dos o más analistas trabajando independientemente. Frecuentemente se emplea un método soporte, exactamente caracterizado, para proporcionar seguridad adicional de que los datos son correctos.
- b) Mediciones por dos o más métodos de referencia independientes en un laboratorio. Los métodos deben tener inexactitudes estimadas pequeñas relacionadas con el requisito del uso final de la certificación.
- c) Mediciones hechas por una red de laboratorios certificados, usando uno o más métodos de exactitud demostrada.

⁷) Se asume aquí que el "método específico" es fácilmente transferible. Esto no es cierto en algunos casos.

En muchos casos, varias combinaciones de estas vías son utilizadas en la certificación. Existen numerosas ventajas y desventajas para cada procedimiento de certificación. La consideración más importante de todas estas vías es que los errores sistemáticos inherentes a los métodos usados, deberán estar bien caracterizados y, lo más posibles, minimizados. Los errores sistemáticos, la precisión del método de ensayo, la variabilidad del material y la estabilidad del material deben todos ser comprendidos y tomados en cuenta cuando se origine el informe de incertidumbre para una propiedad certificada de un MR.

6.2 Valores certificados y su significación. Incertidumbres

En el desarrollo de cada MRC, debe obtenerse la seguridad de que el material usado sea uniforme y estable, que los métodos de ensayo produzcan resultados coherentes y repetibles, y que las condiciones, bajo las cuales debe ser usado, estén cuidadosamente descritas. Eventualmente, estas informaciones cualitativas tendrán que ser llevadas a términos cuantitativos, usando datos generados a partir de los ensayos y condensados en un certificado que será comprensible y provechoso para el usuario.

Esta condensación de la información no es una tarea fácil!. Idealmente, las condiciones experimentales pueden ser descritas en detalle y todos los valores numéricos de las determinaciones individuales pueden ser presentados, de forma tal que el usuario pueda juzgar cómo usar mejor estos resultados.

Generalmente, el costo y el trabajo implicados en una presentación detallada no están justificados en relación con el número de veces que será provechosamente usada. En la mayoría de los casos, por consiguiente, los datos son procesados y condensados en la forma que están presentados en el certificado. Los valores numéricos están normalmente expresados en dos partes: el valor certificado de la propiedad y la incertidumbre de este valor.

La incertidumbre del valor certificado denota cuan bien es conocido este valor. Un número de expresiones diferentes están siendo usadas para la información de la incertidumbre, dependiendo de como el MRC fue desarrollado, del grupo de científicos involucrados y del uso para el cual fue concebido.

En muchos casos, los informes de incertidumbre están basados, en cierta forma, en el juicio subjetivo de los científicos involucrados, más que en una estricta interpretación de los datos. La adquisición de los datos es cara, algunas veces prohibitiva, y estos factores deben ser valorados contra el uso para el cual se concibió el MRC.

La NC ISO Guía 31^[3] está destinada para ayudar a los productores de MRC a preparar certificados claros y concisos. Los certificados comunicarán la información esencial sobre el material de referencia desde el productor hasta el usuario; en esencia, esta información es una declaración de los valores propios certificados, sus significados y sus límites de incertidumbre. El resto de la información gira alrededor de la información central y tiene dos propósitos: describir la naturaleza general y el uso del material, y asegurar al usuario de su integridad.

La NC ISO Guía 31 recomienda que los certificados tengan la siguiente información:

- 1) nombre y dirección de la organización certificadora;
- 2) título del documento;

- 3) legalidad del certificado;
- 4) nombre del material;
- 5) número de la muestra y/o número del lote;
- 6) fecha de certificación;
- 7) disponibilidad de otras formas/tamaño de los materiales de referencia;
- 8) origen del material de referencia;
- 9) suministrador del material de referencia;
- 10) preparador del material de referencia;
- 11) descripción del material de referencia;
- 12) información del uso que se pretenda dar;
- 13) estabilidad, instrucciones para la transportación y el almacenamiento;
- 14) instrucciones especiales para el uso correcto;
- 15) método de preparación;
- 16) información de la homogeneidad;
- 17) valores propios certificados y sus incertidumbres;
- 18) valores propios secundarios dados para información pero no certificados;
- 19) valores especiales obtenidos por laboratorios individuales o métodos;
- 20) significación de la incertidumbre estadística;
- 21) técnicas de medición utilizadas para la certificación;
- 22) nombres de analistas, investigadores y laboratorios participantes;
- 23) advertencias legales;
- 24) referencias (incluyendo relación adicional, si existe);
- 25) firmas o nombres de los funcionarios de certificación.

Más información y explicación sobre estos puntos, están contenidas en la NC ISO Guía 31. Otras referencias útiles sobre certificación de materiales de referencia están incluidas en la NBC Publicación Especial 408^[24] y en los Procedimientos del Simposio Internacional sobre Producción y Uso de Materiales de Referencia^[25].

7 Certificación por un método definitivo

7.1 Concepto de métodos definitivos aplicados a materiales de referencia

La certificación de un MR por un método de medición requiere que el método tenga una alta condición legal científica y un laboratorio o laboratorios de altísima calidad. El método deberá ser suficientemente exacto como para enfrentar él solo la determinación de la propiedad de interés. La exactitud real de este método deberá ser validada por intercomparaciones internacionales, siempre que sea posible. Tal método tendrá una validez, una fundamentación teórica bien descrita, a fin de que los resultados reportados tengan errores sistemáticos despreciables relativos a los requerimientos del uso final. La propiedad en cuestión es medida, o directamente en términos de unidades de medida básicas, o indirectamente relacionada con las unidades básicas a través de la teoría física o química expresada en ecuaciones matemáticas exactas. Si es posible, el laboratorio certificante asegurará que las unidades de medida básicas sean trazables con patrones nacionales y/o internacionales apropiados.

Para el propósito de esta Guía, un método de medición que tenga estas propiedades será denominado “método definitivo”^[16] Esta definición del término “definitivo” es algo diferente de la definición dada en el *Vocabulario Internacional de Términos Básicos y Generales en Metrología*^[5]. Sin embargo, la definición sugerida aquí se aplica más directamente a la certificación de los materiales de referencia. Los métodos definitivos no son generalmente prácticos para el trabajo científico en el terreno porque ellos requieren frecuentemente equipamiento especializado, ellos algunas veces resultan caros y dificultosos para su realización, y usualmente requieren de personal altamente experimentado. Así, la aceptación de un MR certificado de esta manera depende de que la comunidad de usuarios tenga confianza en la habilidad de los laboratorios certificantes para llevar a cabo el método definitivo. En la práctica, cuando se utiliza un único laboratorio certificante u organización, se requiere que el método definitivo sea realizado por dos o más analistas trabajando de forma independiente, usando preferentemente habilidades experimentales diferentes.

La finalidad para escribir un informe de incertidumbre es presentar un resumen claro, conciso y objetivo de qué se conoce acerca del material como resultado del programa de ensayo. La forma exacta del informe de incertidumbre dependerá de la naturaleza del material, de las necesidades de la comunidad de usuarios para el MRC en particular, y de los resultados del programa de ensayo. El buen éxito de la evaluación de la homogeneidad tiene especial importancia en la determinación de la vía en que las incertidumbres son expresadas. Dos importantes casos a considerar son:

- a) el material no homogéneo es despreciablemente pequeño, comparado con el error de medición (véase 7.2);
- b) La no homogeneidad es la mayor fuente de incertidumbre, comparada con el error de medición (véase 7.3).

7.2 Intervalo de confianza (material homogéneo)

7.2.1 Concepto

Cuando los resultados de la evaluación de la homogeneidad indican que la no homogeneidad del material es despreciable, todas las unidades del MRC pueden ser tratadas como que tienen idénticos valores para la propiedad certificada. Así, el valor de cada unidad es el mismo que el valor me-

dio para todas las unidades, y es necesario resumir la incertidumbre en el estimado de esa media. La única fuente de incertidumbre es el error de medición aleatorio, y un método apropiado y usado ampliamente para la expresión de esta incertidumbre es el intervalo de confianza para la media.

7.2.2 Modelo estadístico

Supongamos que los datos experimentales consisten de n mediciones independientes de la propiedad a certificar. Representamos éstas como X_1, X_2, \dots, X_n . El modelo matemático es:

$$X_i = \mu + \varepsilon_i \quad i = 1, 2, \dots, n$$

donde:

μ valor medio verdadero de la propiedad medida

ε_i error de medición asociado con la i -ésima medición.

En este análisis, los ε_i son tomados para hacer variables aleatorias normales independientes, con medias cero y varianza común (desconocida).

7.2.3 Transformación de datos

En algunos casos puede ser necesario transformar los datos (por ejemplo, por logaritmos) para satisfacer la suposición de una distribución normal de los errores. En este caso μ deberá ser interpretada como la media del valor transformado de la propiedad de interés, y el resumen final de los datos implica la retransformación del intervalo de confianza resultante hacia la escala original de medición. En otros casos, donde la carencia de una distribución normal es un problema, los procedimientos estadísticos fuertes o no paramétricos pueden ser usados para obtener un intervalo de confianza válido para la cantidad de interés^[26].

7.2.4 Descripción

Un intervalo de confianza del 95 % tiene la forma:

$$\bar{x} \pm ts/\sqrt{n}$$

donde:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

$$s = \left[\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \right]^{1/2}$$

$t = t_{0,975} (n - 1) =$ fractil 0,975 de la distribución de Student con $(n - 1)$ grados de libertad.

Para obtener un intervalo de confianza para un nivel de confianza diferente al 95 %, se utiliza un valor diferente de t . Por ejemplo, un intervalo de confianza de 99 % se obtiene tomando $t = t_{0,995}$ para los grados de libertad apropiados. Se observará que para cualquier grado de libertad dado, $t_{0,995}$ es un número mayor que $t_{0,975}$. Así, para un juego de datos dado, un intervalo de confianza de 99 % será siempre más amplio que uno calculado con un nivel de confianza del 95 %. Esto está de acuerdo con las noción intuitiva de que ampliando un intervalo dado alrededor de \bar{x} se incrementará la probabilidad de que la media verdadera μ estará contenida en cualquier lugar del intervalo. Explicaciones detalladas de los intervalos de confianza pueden ser encontradas en ISO 2602⁸, incluyendo las tablas de valores de t y t/\sqrt{n} .

7.2.5 Ejemplo de los intervalos de confianza

Un material de trióxido de arsénico de alta pureza (As_2O_3) fue certificado como un MRC por un método de ensayo coulombimétrico de gran exactitud. El trabajo experimental consistió en $n = 23$ determinaciones de la fracción másica, resultando $\bar{x} = 0,999\ 893$ g/g y $s = 0,000\ 104$ g/g. Los datos de este experimento varían simétricamente alrededor del valor medio, e incluyen algunos valores individuales por encima de 1,00 g/g. La técnica de medición real tiene la propiedad de que los resultados de medición válidos pueden conducir a valores calculados para la pureza por encima del 100 % aún cuando la pureza real, naturalmente, debe ser menor que 100 %. Por tanto, la naturaleza de los datos no contradice el uso de un intervalo de confianza estadístico basado en la distribución normal.

⁸) ISO 2602, Interpretación estadística de los resultados de ensayos - Estimación de la media - Intervalo de confianza.

El experimento consistió en el análisis por duplicado de 10 muestras del material y uno por triplicado de la muestra número 11. La variabilidad dentro de la muestra y entre muestras fueron de igual magnitud, confirmando que el material era homogéneo. La ISO 2602 da un valor $= t_{0,975} / \sqrt{n} = 0,432$ para $n = 23$. Por tanto, a un intervalo de confianza del 95 % (para la coulombimetría), la pureza de este material es:

$$\begin{aligned} \bar{x} \pm \left(\frac{t}{\sqrt{n}} \right) s &= 0,999\ 893 \pm (0,432) \times (0,000\ 104) \\ &= (0,999893 \pm 0,000\ 045) \text{ g/g} \end{aligned}$$

Los usuarios de este material pueden entender que la pureza del material está seguramente casi dentro de estos límites; esto es, la fracción másica está entre 0,999 848 y 0,999 938. Hablando más exactamente, si este experimento íntegro fuese realizado un número grande de veces (cada uno con $n = 23$) y si los intervalos de confianza de 95 % fuesen calculados a partir de los resultados de cada experimento, entonces solamente el 5 % de los intervalos generados fallaría al ocultar la pureza verdadera. Esta interpretación asume que el error sistemático del proceso de medición es despreciable.

7.3 Intervalo de tolerancia estadístico (material no homogéneo)

7.3.1 Concepto

Para algunos materiales, puede ser imposible o no práctico controlar la variación unidad por unidad a un nivel que es despreciable comparado con respecto a la incertidumbre de medición. En efecto, el reverso puede ser verdadero: la incertidumbre de medición puede ser despreciable comparada con la variación entre unidades. En esta situación, las unidades individuales de los MRC constituyen una población para la cual la propiedad de interés varía ligeramente de una a otra unidad. El informe de la incertidumbre en el certificado del MRC deberá entonces describir la extensión de la variación entre unidades en una población, de forma tal que el usuario pueda estar seguro que en cualquier unidad obtenida por él tendrá un valor que esté dentro de ese rango. El suministrador de un MRC debe determinar si la variabilidad inherente entre unidades es suficientemente pequeña como para que el material sea útil en el propósito para el cual se concibió, porque si no es así, no existe razón para proceder a la certificación.

7.3.2 Cálculo

A fin de generar un intervalo de tolerancia estadístico, el cual refleje propiamente la variabilidad de unidad en unidad del material, los datos X_1, X_2, \dots, X_n deben constar de n valores independientes, uno para cada una de las n unidades **diferentes**. El modelo matemático es:

$$X_i = \mu + \beta_i \quad y = 1, 2, \dots, n$$

donde:

μ valor medio de la propiedad, promediado sobre la población de todas las unidades

β_i diferencia entre el valor de la propiedad para la i -ésima unidad y la media poblacional.

Típicamente, la propiedad β_i reflejará también una inevitable contribución del error aleatorio de medición realizado en la medición de la i -ésima unidad. Las β_i son tomadas para hacer variables aleatorias independientes normales, con medias cero y varianza común (desconocidas). Como se planteó en 7.2.3, se asume que los datos originales han sido transformados, si es necesario, para alcanzar esta condición en las β_i . El cálculo del intervalo de tolerancia fue descrito en 5.5.

8 Certificación por ensayos interlaboratorios

8.1 Concepto general y práctico

El concepto de la certificación de un MR por ensayos interlaboratorios está basado al menos en dos suposiciones:

- a) existe una población de laboratorios que es igualmente competente en la determinación de las características del MR para proporcionar resultados con exactitud aceptable.
- b) la suposición a) implica que la diferencia entre resultados individuales, tanto internos como entre laboratorios, son estadísticos por naturaleza sin tener en cuenta las causas, (es decir, variación en los procedimientos de medición, personal, equipamiento, etc.).

Cada media de laboratorio es considerada ser un estimado insesgado de la característica del material. Usualmente, la media de las medias de los laboratorios se asume que es el mejor estimado de esta característica; sin embargo, en el caso de distribuciones muy irregulares, como las que pueden ser encontradas por ejemplo en análisis de trazas de elementos, el uso de un estadístico más fuerte, como la mediana o una media ajustada, puede resultar apropiado.

En la práctica, es limitado el tamaño de la población de laboratorios para un programa de análisis interlaboratorios. En muchos casos, por tanto, no puede ser llevado a cabo totalmente un modelo de diseño al azar.

8.1.1 Procedimiento general

El procedimiento general para la certificación de un MR por consenso interlaboratorios está descrito esquemáticamente en la Figura 3. Cada etapa puede ser tratada como si fuese distinta y poner criterios que deben ser satisfechos antes de proceder a la siguiente etapa.

8.1.2 Confirmación de homogeneidad como parte del programa interlaboratorios

Los resultados del programa interlaboratorios pueden servir como una confirmación final de la homogeneidad de un MR siempre que se haya seguido un diseño anidado de dos vías, en el cual han sido usadas unidades pq y donde p laboratorios determinen cada uno el valor de la característica de q unidades con n determinaciones replicadas por unidad. Es importante que los laboratorios no se desvíen del diseño.

8.2 Organización de programas interlaboratorios

Para ser exitoso, un programa interlaboratorios debe tener un objetivo bien definido, ser eficazmente diseñado y eficientemente organizado con lineamientos claros y concisos con los cuales puedan cumplir cabalmente los laboratorios participantes. La participación en un programa de este tipo lleva

implícito un acuerdo a anexar a estos lineamientos. Estos lineamientos consisten en tiempo objetivo, número de unidades, número de determinaciones replicadas por unidad, métodos de medición, tamaño de la porción de ensayo, donde sea aplicable, etc.

8.2.1 Tiempo objetivo

El organizador debe fijar el calendario de tiempo, es decir, las fechas en que las muestras van a ser distribuidas y en la que van a ser reportados los resultados.

8.2.2 Número de laboratorios participantes

El número mínimo de laboratorios participantes que comprende un programa interlaboratorios para la caracterización de un MR varía con la complejidad del procedimiento de medición necesario. Para procedimientos más complejos, mayor será la variación esperada entre laboratorios, de ahí la necesidad de un creciente número de laboratorios participantes para obtener un valor por consenso que tenga una exactitud predeterminada. En la práctica, desafortunadamente, mientras más complejo es el procedimiento, menor es el número de laboratorios capaces de realizar el procedimiento. En casos extremos, el órgano certificador puede ser precisado a renunciar totalmente a un programa interlaboratorios para ciertos candidatos especializados de MR.

Existe un acuerdo general entre laboratorios experimentados en ensayos interlaboratorios de que el número preferido de laboratorios participantes es 15 o más.

8.2.3 Número de unidades y de determinaciones replicadas

Si los resultados del programa interlaboratorios son para servir como una confirmación final de la homogeneidad de un MR, los valores de la característica para un mínimo de dos unidades del MR deberán ser determinados por cada laboratorio participante. De lo contrario, puede ser suficiente una unidad por cada laboratorio participante.

El número mínimo de determinaciones replicadas es dos por unidad del MR. Todas las determinaciones replicadas deben ser hechas en porciones de ensayo separadas.

8.2.4 Métodos de medición

El organizador de un programa interlaboratorios puede especificar el uso de un método único para los laboratorios participantes cuando tal procedimiento de medición "normalizado" y bien establecido se encuentre disponible. En caso contrario, el organizador debe dejar a cada laboratorio participante utilizar el método de su elección, siempre que él posea evidencias de la validez de dichos métodos.

8.2.5 Reporte de los resultados

Los laboratorios participantes reportarán los resultados individuales (no el promedio). El número de cifras significativas reportadas puede ser cumplirá con los lineamientos del programa. Se recomienda hacer un bosquejo con suficientes detalles del procedimiento de medición usado para permitir una comprensión de todas las etapas preliminares en el proceso de medición, por ejemplo, en análisis químico la descomposición de la muestra y la separación del (los) análisis de interés. Deben especificarse las referencias bibliográficas en los casos en que esto sea apropiado.

8.3 Procesamiento inicial de los resultados

Los resultados emitidos por los laboratorios participantes son evaluados en correspondencia con el procedimiento establecido en la Figura 4.

8.3.1 Presentación de los resultados

Por conveniencia del procesamiento y para futuras referencias, los resultados de un programa interlaboratorios serán agrupados en base a las características y tabulados sistemáticamente. Esta tabla incluirá la identificación del laboratorio y el método, los resultados individuales, la media del laboratorio y la desviación típica correspondiente. Sin embargo, si los laboratorios participantes determinaron el valor de la característica para más de una unidad del MR, se recomienda que la media en la unidad y la media general y las desviaciones típicas correspondientes, sean presentadas en una tabla separada de los resultados individuales. Cuando un laboratorio participante ha emitido más de una serie de resultados para una característica obtenidos por diferentes métodos de medición, cada serie debe ser tratada independientemente, es decir, como si fuera de otro laboratorio.

Se recomienda también que los resultados sean presentados en forma gráfica.

8.3.2 Valores fuera de límites explicables técnicamente

Los resultados deben ser revisados para explicar técnicamente los valores fuera de límites, los cuales son excluidos antes de realizar cualquier evaluación estadística. Si es posible, se recomienda que el laboratorio participante comprometido sea informado para su beneficio e invitado a presentar nuevos resultados como sustitución.

8.3.3 Número mínimo de laboratorios

Después de la exclusión de los valores fuera de límites, el número restante de series de resultados debe ser coherente con el principio del número mínimo de laboratorios necesarios para constituir un programa interlaboratorios.

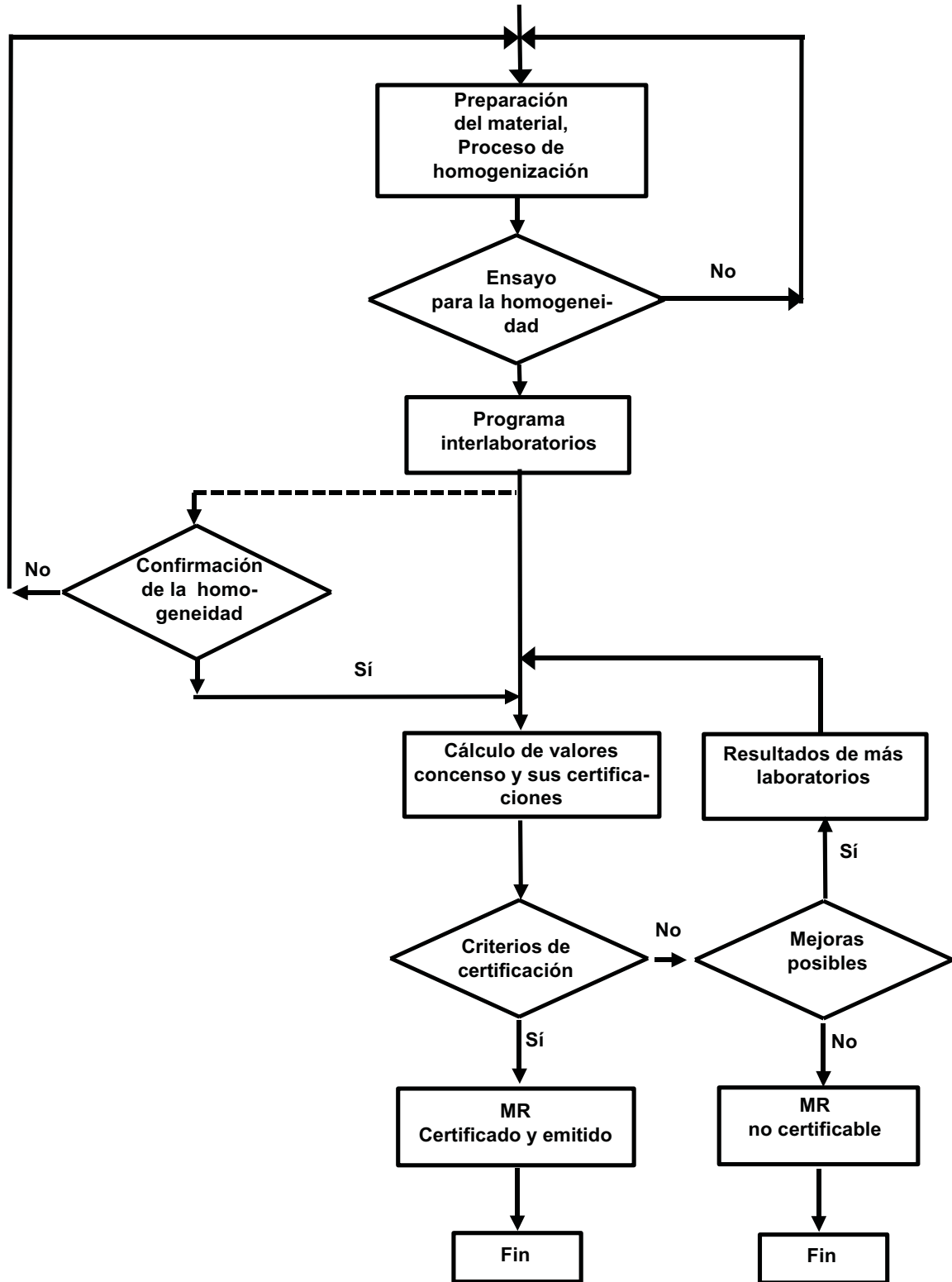


Figura 3. Diagrama esquemático del proceso de preparación y certificación de un MR mediante consenso interlaboratorios.

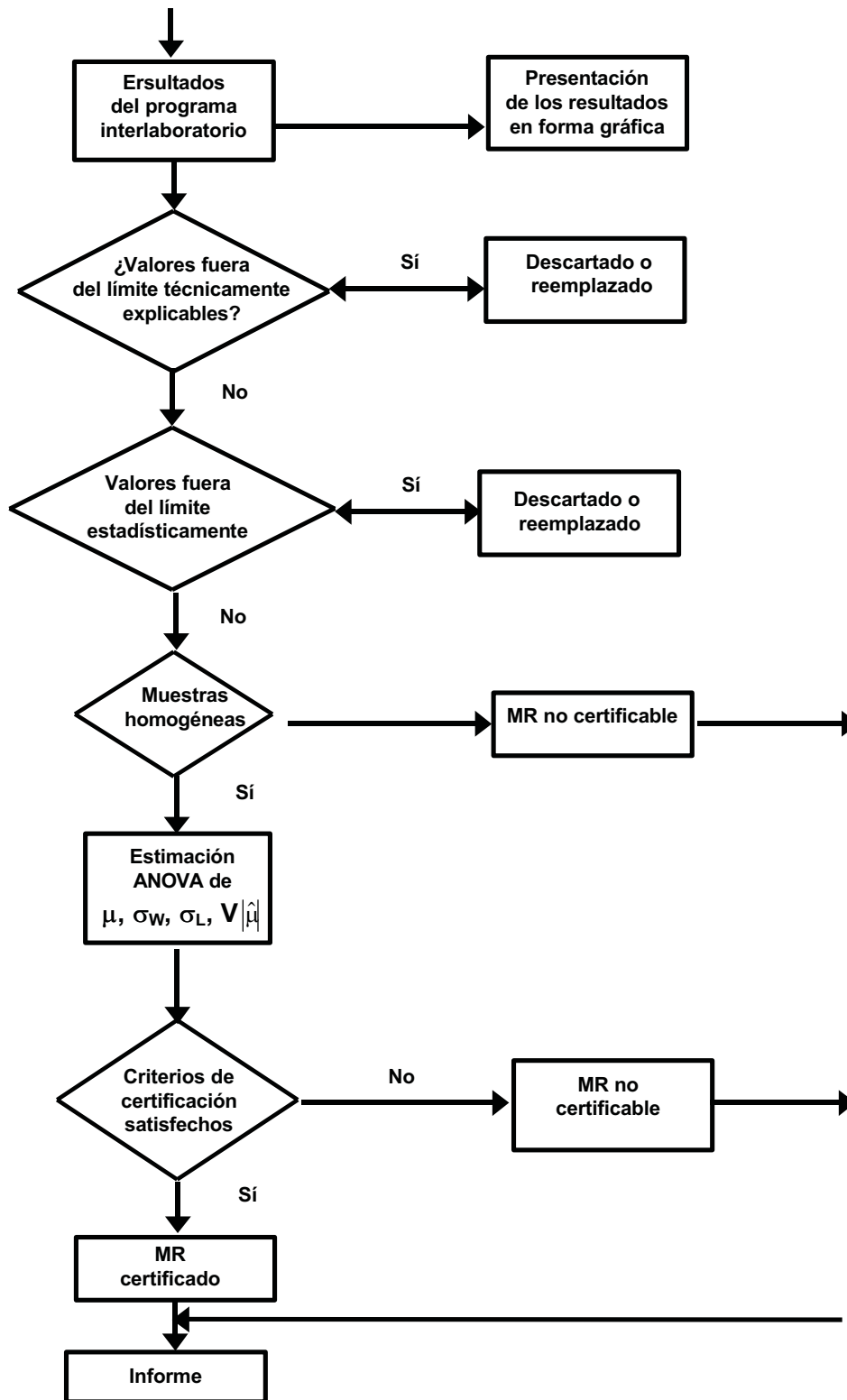


Fig. 4 Diagrama esquemático para la evaluación estadística de resultados interlaboratorios para la certificación de un MR

8.3.4 Distribución de la frecuencia de los resultados

Es esencial conocer cómo son distribuidos los resultados. En muchos casos, la distribución puede ser observada gráficamente.

8.3.4.1 Distribución multimodal

Si la mayoría de los resultados forman dos o más grupos, puede ser inferido un valor por consenso. Las siguientes posibilidades serán consideradas:

- a) si existe correlación de esos grupos con los procedimientos de los métodos de medición, y si la diferencia entre las medias de esos grupos es estadística y físicamente significativa, entonces no existe valor por consenso; en ese caso, es necesario perfeccionar los procedimientos de los métodos de medición para resolver el problema;
- b) si no existe correlación de esos grupos con los procedimientos de los métodos de medición, y si la diferencia entre esos grupos es estadística y físicamente significativa, puede ser necesario un mayor conjunto de resultados para superar los relativamente pobres métodos de medición disponibles.

8.3.4.2 Distribución unimodal

Si la mayoría de los resultados forman un grupo único, se puede inferir que existe un valor por consenso.

Si la distribución es unimodal, debe tomarse una decisión en cuanto a si la suposición de normalidad es razonable. Esta decisión puede basarse tanto en una observación visual del histograma, en la prueba de normalidad, o en experiencias pasadas con la naturaleza de las determinaciones.

Como se ha expuesto en 7.2.3, en algunos casos los resultados tienen que ser transformados a otra forma donde pueda asumirse que ellos siguen una distribución de frecuencia normal. Algunas transformaciones usualmente usadas incluyen logaritmos, raíz cuadrada y formas exponenciales.

8.3.5 Valores estadísticos fuera de límites

Un resultado único o un juego completo de resultados es sospechoso de ser un valor estadístico fuera de límite si su desviación, tanto en exactitud o precisión, con respecto a otros en el mismo juego o a otros juegos, respectivamente, es mayor que lo que pueda ser justificado por fluctuaciones estadísticas pertinentes para una distribución de frecuencia dada. Por eso, la efectividad para la detección de valores fuera de límites depende de la validez de la suposición de la distribución de frecuencia. La prueba para valores fuera de límites puede ser la prerrogativa del estadístico. Para un programa interlaboratorio, el status de valor fuera de límite puede ser conferido a resultados individuales, resultados para unidades individuales o para la serie completa de resultados de un laboratorio.

8.4 Análisis estadístico

8.4.1 Diseño anidado de dos etapas

Este modelo es usado cuando los resultados de un programa interlaboratorios se utilizan para confirmar la homogeneidad, así como para caracterizar el material. El esquema experimental se ilustra esquemáticamente en la Figura 5a). Los resultados pueden ser expresados por la ecuación

$$X_{ijk} = \mu + \alpha_i + \beta_{ij} + \varepsilon_{ijk} \quad \dots (5)$$

donde:

X_{ijk} es el k -ésimo resultado de la unidad de muestra j reportada por el laboratorio i ;

μ es la gran media;

α_i es el error causado por el laboratorio i ;

β_{ij} es el error debido a la j -ésima unidad experimental en el laboratorio i ;

ε_{ijk} error de medición.

8.4.2 Diseño anidado de una etapa

Este modelo es usado cuando el material es aceptado como homogéneo por los organizadores. El esquema experimental está ilustrado esquemáticamente en la Figura 5 b). La ecuación (5) puede entonces ser simplificada a:

$$X_{ik} = \mu + \alpha_i + \varepsilon_{ik}$$

8.4.3 Análisis del diseño anidado de dos etapas

Los parámetros a ser estimados son:

- μ , la gran media (la cual es usada como el valor por consenso).
- σ_L^2 , la varianza del error interlaboratorios (α_i)
- σ_U^2 , la varianza debida a la no homogeneidad entre unidades (β_{ij});
- σ_W^2 , la varianza del error de medición en el laboratorio (ε_{ijk}).

Todos esos parámetros pueden ser estimados simultáneamente por el método de análisis de varianza (ANOVA) (véase 8.4.3.1) si existen suficientes resultados de igual repetición (el mismo número de determinaciones por réplica de cada unidad y el mismo número de unidades por laboratorio) después de haber excluido los valores fuera de límites. Si ese requisito de ANOVA no se puede satisfacer debido al número de valores fuera de límites y/o de resultados perdidos, la significación de la varianza entre unidades (no homogeneidad) puede ser comprobada mediante el procedimiento sencillo para datos no balanceados dado en 8.4.3.2.

Los detalles teóricos y métodos adicionales para ANOVA balanceado y no balanceado están dados en libros de texto normalizados^[27, 28].

8.4.3.1 Cálculo de ANOVA de dos etapas

X_{ijk} es el k -ésimo resultado de la unidad de muestra j reportada por el laboratorio i ;

p es el número de laboratorios participantes;

q es el número de unidades por laboratorio;

n es el número de determinaciones por réplicas por unidad de muestra.

$$\bar{x}_{ij} = \frac{1}{n} \sum_{k=1}^n x_{ijk}$$

$$\bar{x}_i = \frac{1}{q} \sum_{j=1}^q \bar{x}_{ij}$$

$$\bar{x} = \frac{1}{p} \sum_{i=1}^p \bar{x}_i$$

Las sumas de los cuadrados SS_1 , SS_2 y SS_3 se calculan por las siguientes ecuaciones:

$$SS_1 = qn \sum_{i=1}^p (\bar{x}_i - \bar{x})^2$$

$$SS_2 = n \sum_{i=1}^p \sum_{j=1}^q (\bar{x}_{ij} - \bar{x}_i)^2$$

$$SS_3 = \sum_{i=1}^p \sum_{j=1}^q \sum_{k=1}^n (x_{ijk} - \bar{x}_{ij})^2$$

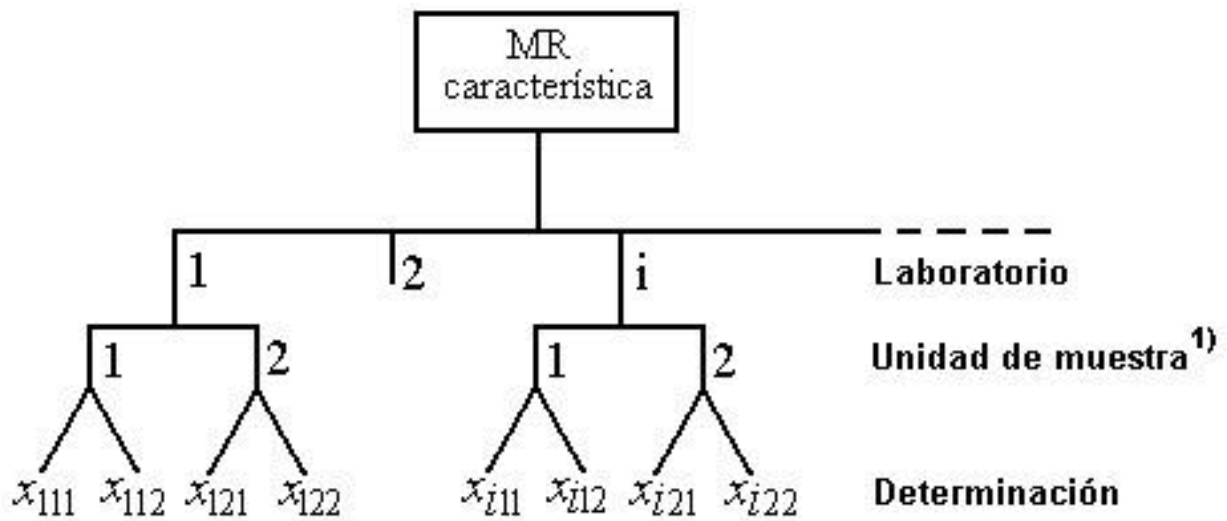
Los grados de libertad son:

$$f_1 = p - 1$$

$$f_2 = p(q - 1)$$

$$f_3 = pq(n - 1)$$

a) Diseño anidado de dos etapas:



1) Todas las unidades experimentales son diferentes. Sin embargo, en cada laboratorio ellas son numeradas 1, 2, ...

b) Diseño anidado de una etapa

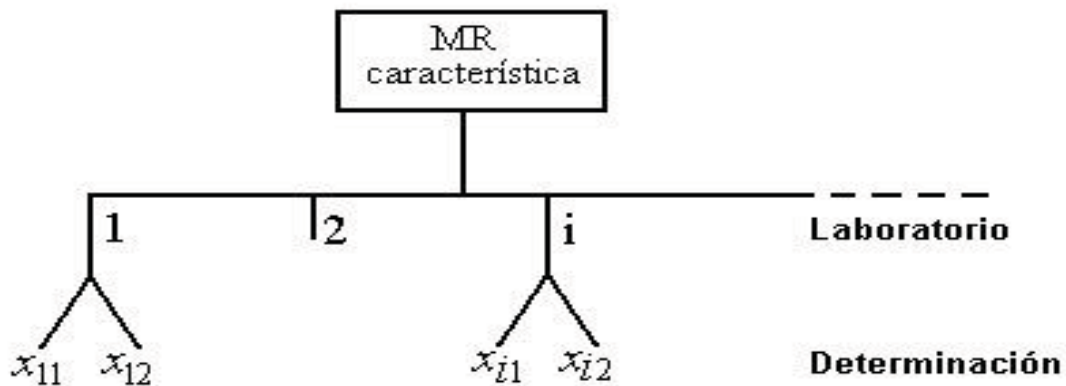


Figura 5. Esquema experimental para un programa interlaboratorios

y cada cuadrado medio está dado como:

$$MS_1 = SS_1/f_1$$

$$MS_2 = SS_2/f_2$$

$$MS_3 = SS_3/f_3$$

Estos resultados deben ser tabulados (véase Tabla 1).

Tabla 1 Tabla de ANOVA

Causa de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio	Valor esperado del cuadrado medio
Entre laboratorios	SS ₁	p-1	MS ₁	$\sigma_w^2 + n\sigma_U^2 + qn\sigma_L^2$
Entre unidades	SS ₂	p(q - 1)	MS ₂	$\sigma_w^2 + n\sigma_U^2$
Error de medición	SS ₃	pq(n - 1)	MS ₃	σ_w^2

Cada parámetro es estimado por las siguientes ecuaciones, donde el acento circunflejo denota el estimador:

$$\hat{\mu} = \bar{x}$$

$$\hat{\sigma}_L^2 = (MS_1 - MS_2)/qn$$

$$\hat{\sigma}_U^2 = (MS_2 - MS_3)/n$$

$$\hat{\sigma}_W^2 = MS_3$$

Si el valor numérico de $\hat{\sigma}_L^2$ ó $\hat{\sigma}_U^2$ es negativo, el cero debe ser usado en su lugar.

Las pruebas para la significación estadística son:

a) varianza entre unidades (no homogeneidad)

$$F_{2/3} = MS_2/MS_3$$

la que será comparada con el valor crítico de la distribución F para p(q - 1) y pq(n - 1) grados de libertad;

b) varianza entre laboratorios

$$F_{1/2} = MS_1/MS_2$$

la que será comparada con el valor crítico de la distribución F para (p - 1) y p(q - 1). grados de libertad.

La varianza del valor por consenso $\bar{\bar{x}}$ es estimada por:

$$\hat{V}(\bar{\bar{x}}) = \frac{MS_1}{pqn}$$

El intervalo de confianza para μ basado en $\bar{\bar{x}}$ es desde A hasta B, donde:

$$A = \bar{\bar{x}} - t_{1-\alpha/2}(p-1) \sqrt{\frac{MS_1}{pqn}}$$

$$B = \bar{\bar{x}} + t_{1-\alpha/2}(p-1) \sqrt{\frac{MS_1}{pqn}}$$

donde: $t_{1-\alpha/2}(p-1)$ es el fractil $1 - \alpha/2$ de la distribución t con $(p-1)$ grados de libertad.

8.4.3.2 ANOVA modificado para datos no balanceados

X_{ijk} es el k -ésimo resultado de la unidad experimental j reportada por el laboratorio i ;

p es el número de laboratorios participantes;

q_i es el número de unidades en el laboratorio i ;

n_{ij} es el número de determinaciones por réplicas de la unidad experimental ij .

$$\bar{x}_{ij} = \frac{1}{n_{ij}} \sum_{k=1}^{n_{ij}} x_{ijk}$$

$$\bar{\bar{x}}_i = \frac{\sum_{j=1}^{q_i} n_{ij} \bar{x}_{ij}}{\sum_{j=1}^{q_i} n_{ij}}$$

La suma de los cuadrados SS_2 y SS_3 se calculan por las siguientes ecuaciones:

$$SS_2 = \sum_{i=1}^p \sum_{j=1}^{q_i} n_{ij} (\bar{x}_{ij} - \bar{\bar{x}}_i)^2$$

$$SS_3 = \sum_{i=1}^p \sum_{j=1}^{q_i} \sum_{k=1}^{n_{ij}} \left(x_{ijk} - \bar{x}_{ij} \right)^2$$

Los grados de libertad son:

$$f_2 = \sum_{i=1}^p (q_i - 1)$$

$$f_3 = \sum_{i=1}^p \sum_{j=1}^{q_i} (n_{ij} - 1)$$

y los cuadrados medios están dados como:

$$MS_2 = SS_2 / f_2$$

$$MS_3 = SS_3 / f_3$$

Estos resultados deben ser tabulados (véase tabla 2)

Tabla 2 - Tabla ANOVA

Causa de variación	Suma de cuadrados	Grados de libertad	Cuadrado medio
Entre unidades	SS ₂	f ₂	MS ₂
Error de medición	SS ₃	f ₃	MS ₃

La prueba para la significación estadística de la varianza entre unidades (no homogeneidad) es:

$$F_{2/3} = MS_2 / MS_3$$

la cual debe ser comparada con el valor crítico de la distribución F para grados de libertad

$$\left\{ \sum_i (q_i - 1) \right\} \text{ y } \left\{ \sum_i \sum_j (n_{ij} - 1) \right\}$$

8.4.4 Análisis del diseño anidado de una etapa

Para casos donde el material es considerado como homogéneo, es decir, que todas las unidades son idénticas, todos los resultados reportados por un laboratorio son considerados como réplicas.

x_{ij} es el j -ésimo resultado reportado del laboratorio i ;

p es el número de laboratorios participantes;

n_i es el número de resultados reportados por el laboratorio i .

$$\bar{x}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ijk}$$

$$\bar{x} = \frac{1}{p} \sum_{i=1}^p \bar{x}_i$$

La varianza del valor por consenso, \bar{x} , es simplemente estimada por

$$\hat{V}(\bar{x}) = \frac{1}{p(p-1)} \sum_{i=1}^p (\bar{x}_i - \bar{x})^2$$

con grados de libertad $(p - 1)$.

El intervalo de confianza para el valor por consenso (media de las medias) es el intervalo desde A hasta B, donde

$$A = \bar{x} - t_{1-\alpha/2}(p-1) \left(\hat{V}(\bar{x}) \right)^{1/2}$$

$$B = \bar{x} + t_{1-\alpha/2}(p-1) \left(\hat{V}(\bar{x}) \right)^{1/2}$$

y $t_{1-\alpha/2}(p-1)$ es como se describe en 8.4.3.1

9 Certificación basada en un enfoque metrológico

9.1 Conceptos

El objetivo de este enfoque es producir valores certificados cuya exactitud e incertidumbre son demostradas mediante evidencias experimentales.

El primer concepto básico de este enfoque es que cuando la propiedad, física o química, de un material puede ser definida desde su primer principio, su valor no depende de un método particular

usado para la medición⁹. Cuando tal valor de una propiedad va a ser certificado, es por lo tanto importante para el órgano de certificación mostrar que el valor no incluye un error sistemático específico para un método o para un laboratorio. El procedimiento consiste en medir la propiedad teniendo en cuenta diferentes métodos, los cuales son considerados como los más exactos para ese fin y que son aplicados por los laboratorios más experimentados en esos respectivos métodos. Este enfoque también se adopta en entidades que trabajan solas: ellas usan diversos métodos, posiblemente con operarios que trabajan independientemente, y comparan los resultados.

El segundo concepto es que el establecimiento de la incertidumbre, la cual es una parte importante del valor asignado a un patrón de medición, puede dejar de ser confiable cuando no está basado en una muy cuidadosa comparación entre los resultados de diferentes laboratorios (de alto nivel) y diferentes métodos. Esto se ilustra con ejemplos en 9.2 y 9.3.

La medición de las magnitudes referidas anteriormente es trazable o debe ser trazable a escalas de medición, trazables ellas a su vez al SI. Por definición, la trazabilidad es la propiedad de un resultado de una medición, mediante la cual puede ser relacionada con patrones apropiados a través de una cadena ininterrumpida de comparaciones.

La trazabilidad es necesaria para sustentar el concepto de exactitud. La trazabilidad de los procesos analíticos es más difícil de establecer que en las mediciones físicas. Los problemas involucrados en esta trazabilidad se exponen en detalle en 9.3.

En 9.4, se establecen ejemplos de propiedades, las cuales están definidas solamente por un método y pueden ser trazables solamente a una escala de medición convencional.

9.2 Certificación de propiedades físicas

Las mayoría de las mediciones exactas para las unidades básicas, sus múltiplos más comunes y sus submúltiplos, se llevan a cabo en los laboratorios metrológicos primarios. Aquí, todas las fuentes de errores e incertidumbres son investigadas en gran detalle; los métodos de medición han sido perfeccionados durante muchos años para reducir las incertidumbres. La exactitud de esas mediciones está claramente bien establecida, especialmente cuando ellas han sido objeto de las comparaciones interlaboratorios. Deben hacerse con reserva las mediciones donde no han existido intercomparaciones. Además, cualquier nuevo laboratorio que se establezca necesita muchas intercomparaciones para asegurar que sus propios estimados de exactitud son correctos y que ningún error ha escapado a su atención.

Las intercomparaciones adicionan confianza a la incertidumbre calculada por los laboratorios de metrología individualmente. Algunas veces ellos utilizan factores de seguridad que no son necesarios; a veces ellos subestiman sus propias incertidumbres.

La práctica actual, mediante la cual cada laboratorio metrológico evalúa la incertidumbre de una medición en particular por sí mismo, es intrínsecamente peligrosa. No es posible para un laboratorio

⁹) Existen propiedades que están definidas solamente como una función de un método; esta cuestión se trata en 9.4.

solo evitar todos los errores en todas las circunstancias, en particular para las unidades derivadas. Las intercomparaciones detectan errores que no han sido tomados en cuenta y situaciones donde todos los parámetros que influyen en la medición no son suficientemente bien controlados.

Desafortunadamente, no existe un requisito general en Metrología que establezca que la incertidumbre esté basada en adecuadas intercomparaciones. La certificación de un material de referencia sobre la base de los resultados de un laboratorio metrológico único puede por tanto implicar un riesgo, el cual no debe ser desestimado.

Cuando se emprende la certificación de una propiedad o magnitud física, es por tanto importante tener una intercomparación entre los mejores laboratorios metrológicos, seguido de una amplia exposición de los resultados con todos los participantes para resolver cualquier posible discrepancia. Si los laboratorios metrológicos primarios no están involucrados ellos mismos en la medición, debe ser establecida previamente la trazabilidad completa de los laboratorios participantes con los respectivos laboratorios nacionales.

Los participantes deben entonces comparar sus mediciones y exponer todos los posibles errores responsables de discrepancias y eliminarlos, aún permaneciendo independientes. Esto se describe más detalladamente para mediciones químicas en 9.3.2.

Si fuese posible más de un método, y si esos métodos tuviesen igual validez, es importante compararlos. No obstante, es provechoso recordar que el método con la menor ruta de trazabilidad, es decir, con la relación más directa con las unidades básicas, tiene una mayor probabilidad de ser más exacto.

En casos extremos, pueden existir naturalmente situaciones en las que un laboratorio único, que ha comparado su método con todos los demás métodos posibles y que ha eliminado la mayor parte de las causas de errores, es capaz de perfeccionar su método para reducir la incertidumbre aunque tomando considerables precauciones para evitar cualquier fuente accidental de errores.

Algunos problemas de medición en el campo de las propiedades físicas pueden ser brevemente ilustrados por la conductividad térmica de los materiales aislantes y refractarios. Hasta hace algunos años, los laboratorios no eran capaces de llevar a cabo este tipo de mediciones con la exactitud requerida aunque la calibración de la instrumentación mostrara satisfacción. La placa de calentamiento protegida utilizada para la medición fue construída y operada en correspondencia con las normas nacionales e internacionales existentes. La conformidad parecía ser satisfactoria para aplicaciones técnicas simples. Sin embargo, en la mayoría de los laboratorios existía un error sistemático. Ocurría una pérdida de calor por encima de la temperatura ambiental debido a que el aro protector no era suficiente. Cualquier material de referencia certificado sobre esa base podía tener una trazabilidad errónea. El método y el equipamiento fueron por tanto modificados hasta que la pérdida de calor resultó ser despreciable.

La determinación exacta de la conductividad térmica de materiales refractarios resulta muy dificultosa por el método directo cuando se usa el aparato de placa de calentamiento protegida, principalmente debido a la pérdida del calor y a dificultades experimentales. Los métodos tales como el método del alambre caliente o el método de destello no presentan tales dificultades, pero la trazabilidad de los mismos no es fácil de establecer y por tanto, esos métodos no son los mejores para la certificación. Sin embargo, los resultados de esos métodos son importantes como una verificación de los resultados de la placa de calentamiento protegida.

9.3 Certificación de una composición química

9.3.1 Trazabilidad

En el campo de la química analítica, no hay establecido sistema de medición organizado como en el campo de la metrología, con laboratorios primarios y de calibración y patrones de medición disponibles en circulación. El concepto de exactitud es por consiguiente más difícil de alcanzar y la trazabilidad es más difícil de realizar.

En química, las calibraciones en el sentido acostumbrado no son la mayor fuente de dificultades aun cuando la tarea del químico es más ardua que la del metrólogo. El necesita no solamente patrones físicos de masa, volumen, temperatura, etc. sino también patrones de todas las especies químicas que él tiene que determinar: elementos, compuestos orgánicos, etc. Cada uno de esos patrones químicos tiene una incertidumbre (por ejemplo, impurezas), la cual es a veces subestimada.

El problema mayor es sin embargo la trazabilidad del proceso analítico completo: la cadena de trazabilidad se ve interrumpida cada vez que la muestra es modificada física o químicamente durante el proceso analítico.

Como la variedad de procedimientos para procesar la muestra es amplia, no es posible tratar la trazabilidad en general. Los siguientes apartados deben ser considerados solamente como ejemplos.

9.3.1.1 Pesaje de la muestra

El primer paso del proceso analítico es el pesaje de la muestra. Este no presenta problemas de trazabilidad si la balanza es calibrada periódicamente. Los errores humanos no están excluidos, pero los mismos no son frecuentes.

9.3.1.2 Tratamiento de la muestra

Cada vez que la muestra es disuelta o sometida a tratamiento similar, la cadena de trazabilidad se interrumpe y cualquier evaluación de incertidumbre debe tener esto en cuenta. Para establecer la trazabilidad para esa parte del procedimiento de medición, un laboratorio debe demostrar la relación existente entre la muestra inicial y la solución preparada a partir de ella. Las principales cuestiones a responder serán, ¿fue la muestra totalmente disuelta?, ¿cuáles fueron las pérdidas?, ¿hubo contaminación?. Si el análisis no es para determinar un elemento, sino un compuesto, ¿cambió el compuesto durante la etapa de disolución?. En el caso de compuestos orgánicos, la eficiencia de extracción es una de las causas principales de dificultades.

Si después del tratamiento de la muestra, la solución es sometida a ulteriores manipulaciones (preconcentración, precipitación, etc.) cada paso complica la ruta de trazabilidad y añade nuevas posibilidades de pérdidas o contaminaciones que deben ser investigadas.

Es bien conocido que algunos de los parámetros relacionados aquí dependen más de la matriz que del elemento o compuesto a determinar.

9.3.1.3 Determinación final

El tercer paso en un proceso analítico es la determinación final. Además de la gravimetría, la titrimetría y la coulombimetría, la mayoría de los métodos son indirectos, por ejemplo, la espectrometría y la absorción atómica. La instrumentación usada para esas mediciones proporciona una señal que debe estar en correlación con la concentración de la sustancia de interés en la muestra desconocida. Esta correlación se establece por medio de una curva de calibración.

En este caso hay dos grupos de problemas a considerar:

- ¿hay algún error introducido al producir la curva de calibración y cuál es la exactitud?
- ¿es correcto usar esta curva de calibración en particular?

Si suponemos que la calibración puede ser ejecutada por medio de soluciones, entonces los parámetros más importantes a tener en cuenta son:

- la exactitud de las mediciones (masa, volumen) realizadas para la preparación de la solución;
- la pureza de los elementos o sustancias, la estequiometría de los compuestos, etc.;
- la pureza del agua o del solvente.

Los errores debidos a la curva de calibración no son raros, aún en los buenos laboratorios.

No obstante, tal y como se expone en 9.3.1.4 también los errores mayores se deben al hecho de que los usuarios a veces elaboran curvas de calibración que no son apropiadas para las soluciones que ellos tienen que analizar; esos se denominan efectos matrices, interferencias, etc.

En términos metrológicos, esto puede ser expresado como sigue: cada laboratorio elabora para sí mismo una escala de medición que no es totalmente apropiada para las mediciones a realizar, y cada uno elabora una escala de medición diferente.

9.3.1.4 Efecto matriz

La respuesta de un elemento particular de un proceso de medición (por ejemplo, espectrometría, absorción atómica) puede depender de la solución (viscosidad, conductividad, potencia iónica) o de los iones presentes en el mismo (interferencias).

Además de un gran número de dichos casos en los análisis inorgánicos, en la química clínica se presentan efectos matrices serios, donde algunos métodos designados para analizar un suero pueden ser inadecuados para soluciones acuosas. Para tales métodos la calibración debe ser hecha con suero humano; si esto no es posible, debe ser demostrada la validez de cualquiera otra matriz.

A este respecto, el término "calibrante" usado por bioquímicos, puede ser falso. Análogamente, en la química inorgánica, una solución de calibración debe simular igualmente la solución que se analiza.

9.3.2 Trabajo de certificación

La tarea de cualquier laboratorio participante en un ejercicio de certificación de un nuevo material de referencia incluye el estudio de los parámetros mencionados en 9.3.1. Un estudio completo requiere la comparación de diferentes métodos de tratamiento de las muestras y diferentes métodos de determinación. No obstante, esto puede ejecutarse mejor colectivamente con el objetivo de tener la colaboración de especialistas experimentados en cada método. Además, para cada método existirá más de un laboratorio a fin de evitar errores sistemáticos causados por efectos de laboratorio o efectos del operario. Puede puntualizarse, que los errores (por, ejemplo, aquellos causados por contaminaciones) sólo pueden ser detectados comparando los resultados de diferentes laboratorios.

La necesidad de examinar cuidadosamente los resultados de los diferentes participantes puede ser ilustrada con los ejemplos dados en las tablas 3 y 4, los cuales son más bien típicos de los análisis de traza de elementos, en muy bajos niveles. Los laboratorios encuentran frecuentemente valores que son demasiado altos, porque todos ellos producen alguna contaminación. Si uno también ha adoptado apresuradamente el valor medio de sus resultados, tendría un error sistemático por exceso, y un material de referencia totalmente no confiable desde el punto de vista de la trazabilidad. Esto explica por qué el procedimiento propuesto para enfocar la exactitud está compuesto de varios pasos en los cuales los participantes exponen todas las fuentes de errores en todas las partes del procedimiento analítico y tratan de reducirlos. Entonces los análisis son repetidos (posiblemente no en las mismas muestras exactamente), y los resultados expuestos nuevamente tantas veces como sea necesario hasta obtener suficiente convergencia.

En el caso de los denominados métodos “definitivos” como IDMS también existe la necesidad de diversos laboratorios. Para una determinación en particular pueden existir más de un método “definitivo”, o diferentes variaciones de un método “definitivo”; esto es esencial naturalmente para verificar que ellos ofrecen el mismo resultado y este no es necesariamente el caso. Si después de comparar detalladamente los resultados de diferentes laboratorios no fuese posible identificar los errores, la variación de los resultados (entre laboratorios) representa la incertidumbre de la técnica en ese proceso. Trabajar con un solo laboratorio puede quizás conducir a una menor dispersión de los resultados, pero ésta no tiene necesariamente que representar la incertidumbre real.

En resumen, el trabajo de certificación de acuerdo con el enfoque propuesto aquí para un material homogéneo y estable debe incluir los siguientes pasos:

- examen, con laboratorios experimentados, de las metodologías más confiables (exactas) para el análisis del elemento o sustancia en la matriz particular considerada;
- una primera ronda de análisis;
- una detallada exposición de los resultados con todos los participantes para tratar de hallar las explicaciones de las diferencias; particular atención debe concederse a:
 - tratamiento de la muestra,
 - posibles pérdidas, contaminaciones,
 - tratamientos de la solución,
 - errores incluidos en la curva de calibración,

- combinar la calibración al producto para analizar efectos matrices, interferencias;
- una segunda ronda de análisis con los mismos laboratorios, pero tal vez con un material de composición ligeramente diferente;
- discusión;
- rondas adicionales de análisis que sean necesarias.

El procedimiento descrito conduce a menudo a rechazar algún método(s) o a excluir algunos laboratorios que no pueden perfeccionar su funcionamiento. Al final de este largo procedimiento, uno tiene un conjunto de resultados técnicamente consistentes para los cuales uno calcula el valor medio, y su intervalo de confianza de 95 % (adoptado como incertidumbre). Ejemplos de etapas sucesivas se brindan en las Figuras 6 y 7. Las estadísticas no son usadas para otros propósitos que no sea para verificar que las condiciones han sido cumplimentadas para calcular un intervalo de confianza de 95 % .

Las estadísticas para el cálculo son las mismas que las mostradas en la NC ISO Guía 33^[29].

Cuando los resultados no son coherentes, se debe concluir que el trabajo técnico no ha terminado y que la certificación no es posible.

Debe señalarse que para elementos-traza o para la certificación de niveles de impureza, la distribución de los resultados puede ser logarítmica normal. El intervalo de confianza puede ser no simétrico.

NOTAS

1) El(los) método(s) utilizado(s) para certificar un material de referencia resultan a veces muy diferentes de los métodos utilizados en las prácticas de rutina (por ejemplo, para certificar cortisol en suero, se tiene que usar GCMS, mientras que en la práctica el método comúnmente usado es de radio-inmunoensayo). En esos casos, es importante verificar que el MR es apropiado para ser usado con el método de rutina.

En la Figura 9, puede notarse que solamente se tuvieron en cuenta los resultados GCMS para la certificación. Los otros métodos fueron usados para verificar la idoneidad del MR.

2) Para la preparación de un material de referencia en el campo de la biomedicina en particular, el suero sanguíneo es tratado con agentes estabilizadores o es liofilizado. Esto es esencial para verificar la aptitud del material de referencia después de esos tratamientos.

Tabla 3 - Trazas de elementos en la leche

Valores en nanogramos por gramos

Elemento	Primera intercomparación (rango de los resultados)	Campaña de certificación (rango de los resultados)	Certificado
Cd	0,4 a 4 500	1 a 5,6	2,9
Hg	0,6 a 42	0,73 a 1,27	1,0
Pb	68 a 5 500	92,4 a 112,5	104,5
Cu	470 a 9257	475 a 700	545

Tabla 4 - Resultados de análisis de hojas de olivo

Elemento	Resultados de 1979 μ/g	Relación	Resultados de 1981 μ/g	Relación
Cd	0,050 a 6,654	133	0,054 3 a 0,121	2,2
Pb	17,6 a 33,3	1,9	20,2 a 26,4	1,3
Hg	0,005 a 0,702	140	0,247 a 0,336	1,4
Cu	0,5 a 131,9	264	43,2 a 50,8	1,2
Zn	12,3 a 31,6	2,6	14,5 a 17,7	1,2
Mn	0,4 a 4,6	11,5	51 a 61,8	1,2

Tabla 5. Determinación de pesticidas en leche en polvo con ciertos componentes añadidos

Compuesto	Resultado mg/kg	Relación	Cantidad añadida mg/kg
HCH	0,001 a 0,22	220	0,28
α -HCH	0,009 a 0,60	67	0,11
γ -HCH	0,001 14 a 0,18	158	0,20
DDE	0,004 3 a 0,47	109	0,54
<i>op'</i> DDT	0,003 a 0,24	80	-
β -HCH	0,01 a 0,13	13	0,08
β -HEPO	0,001 a 0,13	130	0,12
Dieldrin	0,01 a 0,104	10	0,10
<i>pp</i> DDT	0,005 a 0,36	72	-

9.4 Certificación de propiedades convencionales

En química, bioquímica y otras tecnologías, muchas propiedades están definidas solamente por un método, un procedimiento de ensayo o un equipo en particular. Ejemplos son las propiedades mecánicas de materiales, actividad de enzimas, etc. Los resultados de esas mediciones o ensayos pueden estar sujetos a gran variabilidad con graves consecuencias económicas.

Como en cualquier otra medición, los resultados dependen de la forma en que es aplicado el procedimiento. Sin embargo, no siempre se describe el procedimiento en las normas escritas en todos sus detalles necesarios y el operario no tiene medios de verificar si es correcta la forma en que él ha interpretado y aplicado el procedimiento. Por eso la necesidad del material de referencia.

Los diagramas de la Figura 10 muestran los resultados de la determinación de la actividad de una enzima (γ -glutamil transferasa) en una matriz de albúmina con el mismo método IFCC. Los laboratorios mostrados en la parte derecha, han sido previamente entrenados con el método. Los laboratorios mostrados en la parte izquierda, eran de alto nivel científico, pero no poseían experiencia previa en el método. Mientras que los dos diagramas superiores de la Figura 10 se refieren a un material, el diagrama inferior concierne a un material diferente.

Similarmente, donde un ensayo depende del uso de una máquina o equipamiento en particular, es posible, pero extremadamente costoso y extenso, verificar que la máquina satisface todas las especificaciones. Una forma simple de evitar esto, es medir o comprobar una muestra de referencia. Si los resultados son satisfactorios, esto significa que la máquina está en buenas condiciones y que por tanto, los resultados pueden ser considerados trazables a la escala de medición establecida por la norma pertinente escrita. Naturalmente, el trabajo de certificación para establecer materiales de referencia para tales propiedades o escalas de medición requiere la aplicación de los mismos prin-

cipios explicados anteriormente. La medición de esos parámetros, los cuales pueden ser la masa, el volumen, la longitud o la temperatura, deben ser de por sí exactas y trazables y por tanto, pueden requerir una calibración extensa. Con frecuencia es necesario un considerable esfuerzo para investigar la influencia de varios parámetros de los procedimientos y del equipamiento, sobre los resultados de medición. Las verificaciones y calibraciones deben ser hechas independientemente en unos pocos laboratorios, por no decir en varios, con el fin de evitar un sesgo uniforme que pudiera aparecer como un buen consenso y dar una ilusión de exactitud.

9.5 Uso de materiales de referencia para establecer la trazabilidad

En 9.3.1, se da una visión de un número de parámetros que un laboratorio debe controlar y verificar para asegurar la trazabilidad de las determinaciones. Hacer esto con todos los detalles necesarios es un trabajo muy arduo.

Esto puede ser considerablemente simplificado, usando un material de referencia certificado de trazabilidad establecida. El material de referencia tiene que ser suficientemente similar (en la matriz) a la muestra real que se analiza, a fin de incluir todos los problemas analíticos que pudieran ser causas de errores en las determinaciones. Naturalmente, el usuario debe aplicar al material de referencia el mismo procedimiento analítico que aplica a su muestra desconocida.

Cuando el laboratorio utiliza uno de esos materiales de referencia halla solamente una diferencia despreciable con respecto al valor de certificación, que indica al mismo tiempo que el resultado es exacto y que es trazable a la escala de medición fundamental. Si la diferencia no es permisible, esto indica que el procedimiento de medición incluye errores, los cuales deben ser identificados y eliminados. Se sugiere que los pasos más críticos sujetos a errores sean el tratamiento de la muestra y la coordinación de la calibración.

Por ello, el papel del material de referencia es comparable al de los patrones de transmisión usados en los laboratorios de metrología en la industria, donde se permite trabajar con un margen específico de incertidumbre.

Los materiales de referencia también hacen posible el establecimiento de la incertidumbre de una medición para determinaciones analíticas o pruebas tecnológicas.

La importancia de un material de referencia certificado va por tanto más allá que la definición de material de referencia dada en NC ISO Guía 30^[2].

Un material de referencia es usado no solamente

- para la calibración de un aparato,
- para la verificación de un procedimiento de medición,
- sino también
- para el establecimiento de la trazabilidad de los resultados de medición,
- para la determinación de la incertidumbre de esos resultados.

- Finalmente, no podemos olvidar que el uso de un material de referencia no elimina totalmente la importancia de las auditorías, cuyo propósito es verificar que no se produzcan errores en el uso del MR.

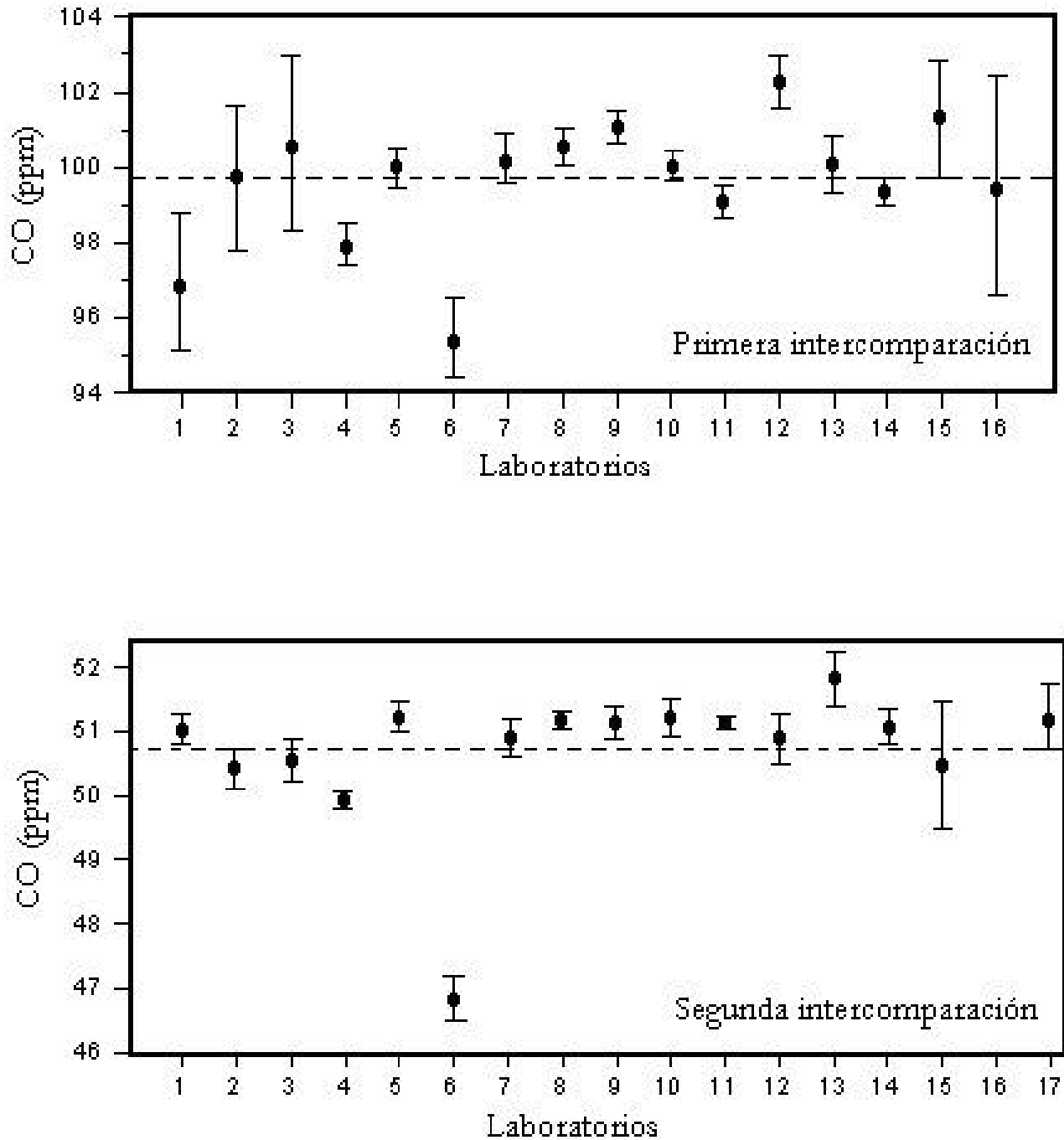


Figura 6. Resultados de la primera y segunda intercomparación de análisis de monóxido de carbono en nitrógeno

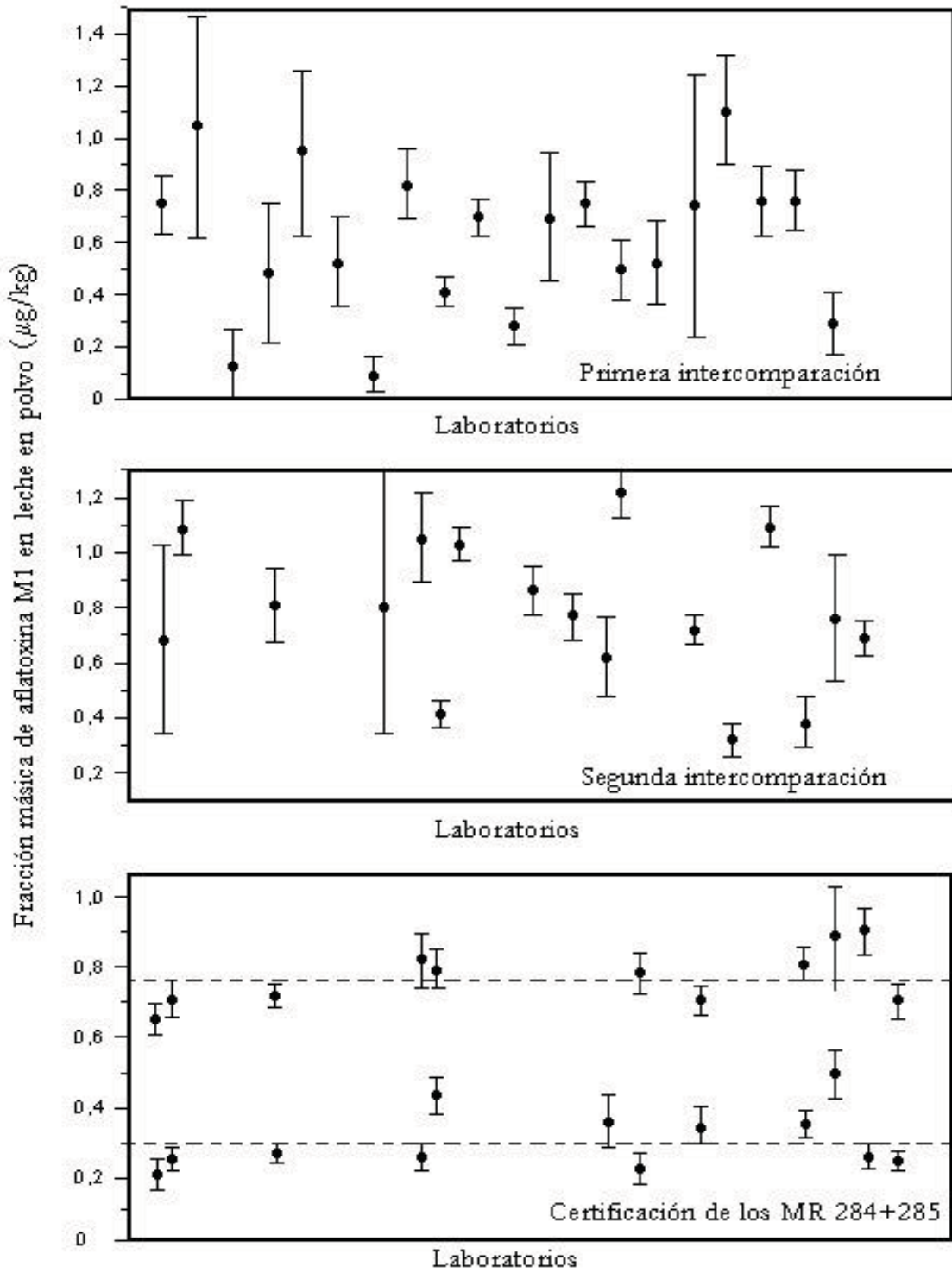


Figura 7. Evolución de los resultados en sucesivas intercomparaciones para la determinación de aflatoxina en la leche en polvo

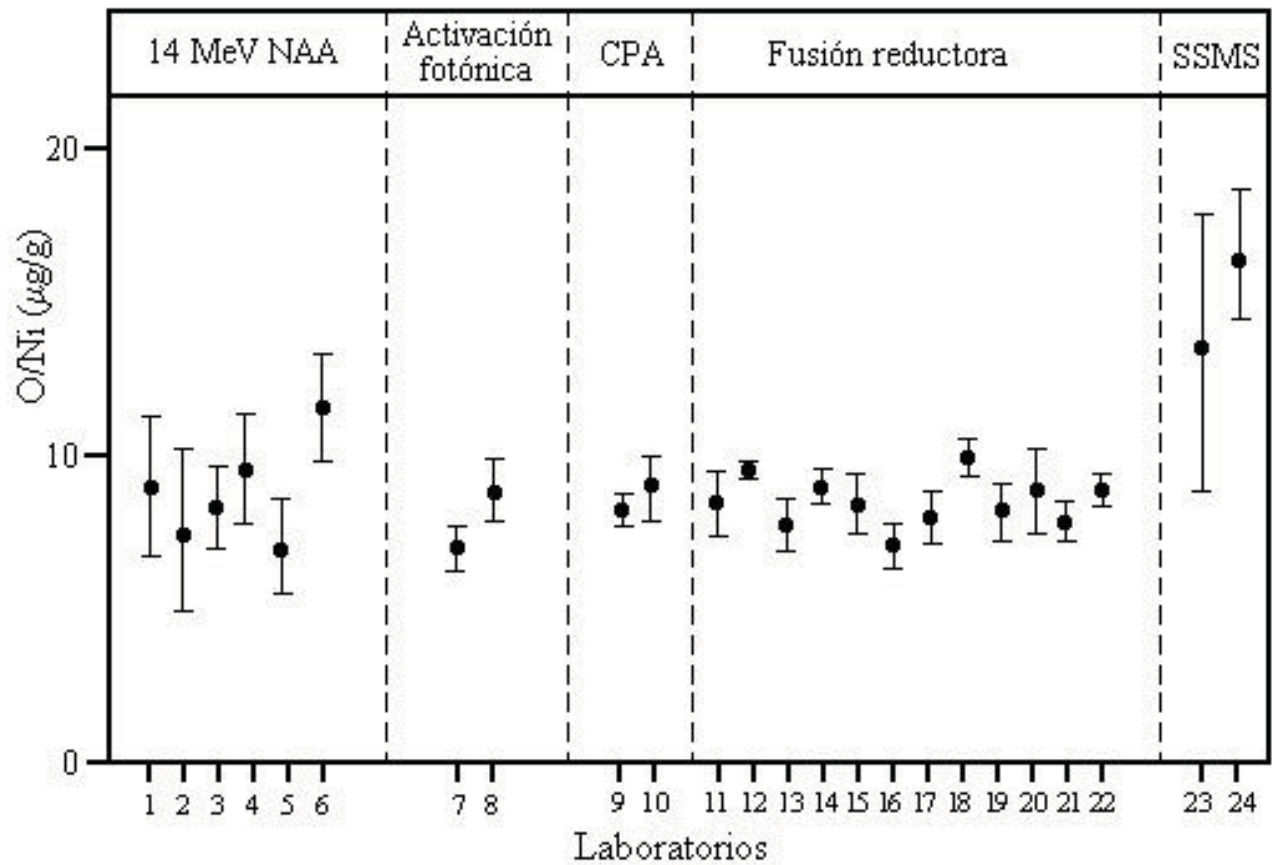
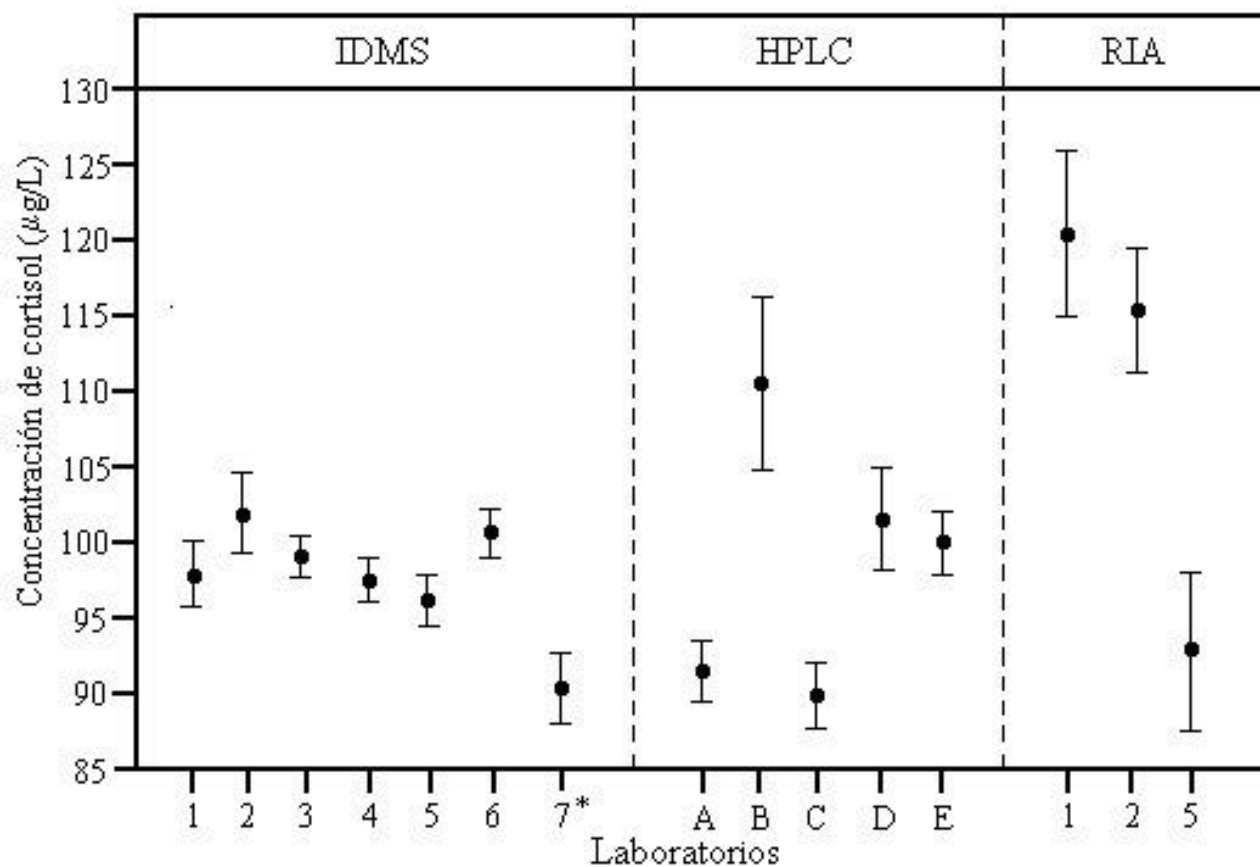


Figura 8. Resultados de laboratorios individuales para la relación oxígeno/níquel



* Excluido para la certificación

Figura 9. Determinación de cortisol en suero humano reconstituido

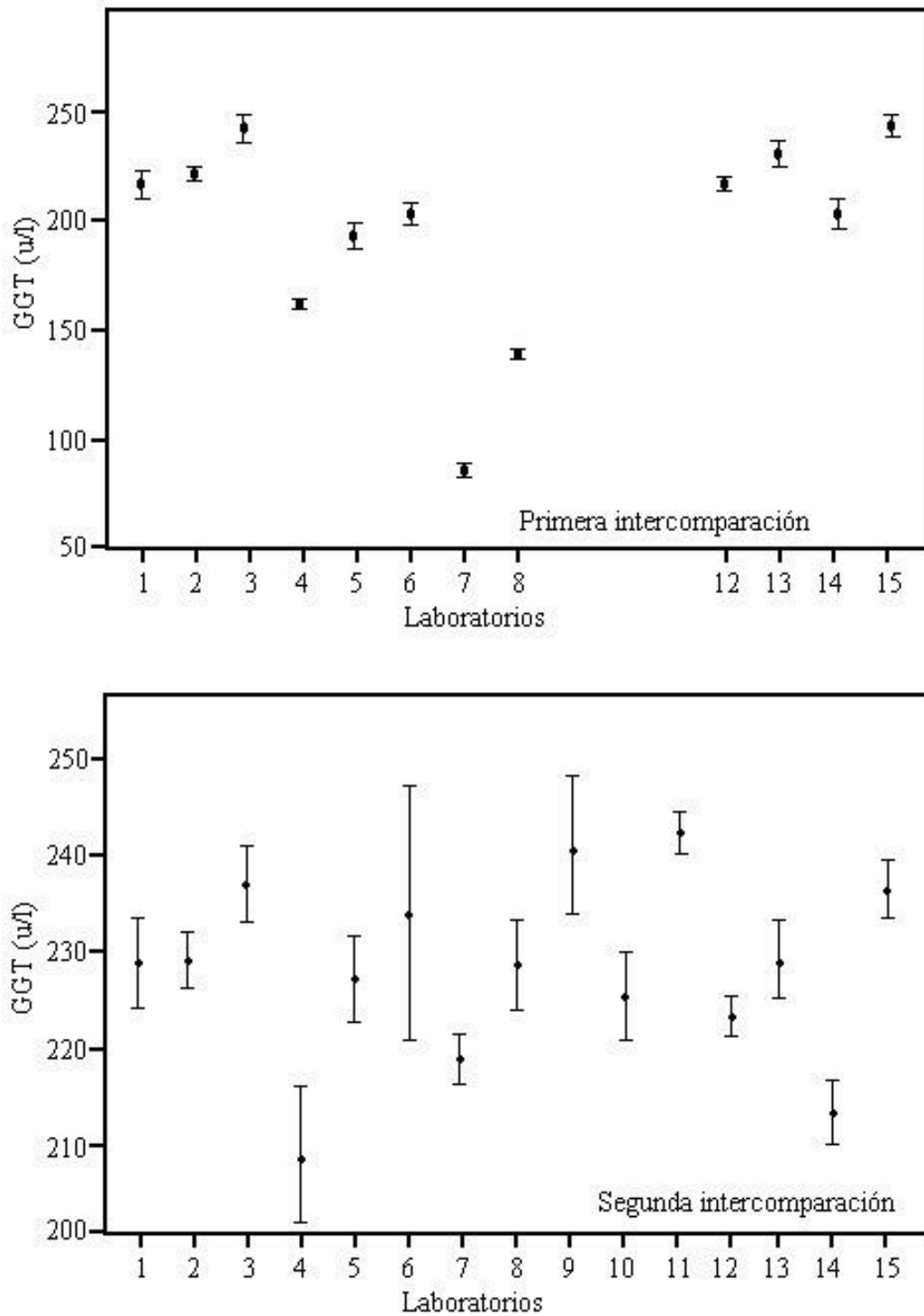


Figura 10a). Resultados sucesivos para la determinación de γ -glutamilttransferasa en albúmina: Primera y segunda intercomparaciones

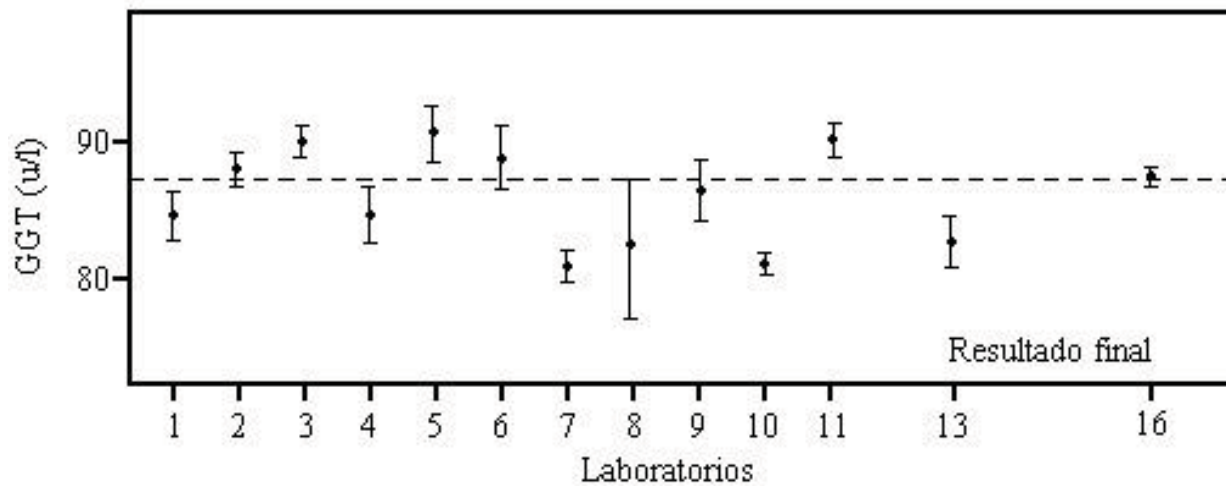


Figura 10b). Resultados sucesivos para la determinación de γ -glutamyltransferasa en albúmina: Campaña final

ANEXO A BIBLIOGRAFÍA

- [1] ISO Guía 6, *Mención de materiales de referencia en Normas Internacionales* (Sustituye: texto incorporado en las *Directivas para el trabajo técnico de la ISO*, catorce edición, 1985).
- [2] NC ISO Guía 30, *Términos y definiciones usadas en relación con materiales de referencia*. Primera Edición, 1981.
- [3] NC ISO Guía 31, *Contenido de los certificados de materiales de referencia*. Primera Edición, 1981.
- [4] ISO. *Directorio de materiales de referencia certificados*. Primera Edición, 1982.
- [5] BIPM, IEC, ISO, OIML. *Vocabulario Internacional de Términos Generales y Básicos en Metrología*. Ginebra, Organización Internacional de Normalización, 1984.
- [6] OIML. *Vocabulario de Metrología Legal*, París. Organización Internacional de Metrología Legal, 1978.
- [7] BIPM. *El Sistema Internacional de Unidades (SI)*. París, Buró Internacional de Pesos y Medidas, Offilib, 4. edición, 1981.
- [8] Marschal, A. *El mole: de la definición al usuario*. Boletín BNM (París) **39** 1980: 33-39.
- [9] Plebanski, T. Materiales de referencia recomendados para la materialización de las propiedades físico-químicas. Sección: viscosidad, *Pure Appl. Chem.* **52** 1980: 2393-2904.
- [10] Ditmars, D.A., Ishihara, S., Chang, S.S., Bernstein, G y West, E.D. Material de referencia patrón de entalpía y capacidad calorífica: zafiro sintético ($\alpha\text{-Al}_2\text{O}_3$) desde 10 hasta 2 250 K. *J. Res. Nat. Bur. Stand.* **87** 1982: 159-163.
- [11] Holden, N.E. Pesos atómicos de los elementos 1979. *Pure Appl. Chem.* **52** 1980: 2349-2384.
- [12] Escala Práctica Internacional de Temperatura. *Rendición de Cuentas de Ciencias de la 15. Conferencia General de Pesos y Medidas*, 1975.
- [13] Diehl, H. *Anal. Chem.* **51** 1979: 318A-330A.
- [14] Cali, J.P. y cols. El papel de los materiales de referencia patrones en los sistemas de medición. *NBS Monograph 148*, Washington, DC, National Bureau of Standards, 1975.
- [15] Marschal, A. *Materiales de referencia*. Buró Nacional de Metrología, Laboratorio Nacional de Ensayos, París.
- [16] Ku, H.H. (Ed.) Medición de precisión y calibración -conceptos estadísticos y procedimientos. *Spec. Publ.* 300, Vol 1, Washington, DC, National Bureau of Standards, 1970, p. 346.

- [17] Cali, J.P. y cols. Un método de arbitraje para la determinación de calcio en suero. *Spec. Publ.* 260-36, Washington, DC, National Bureau of Standards, 1972, p. 136.
- [18] Natrella, M.G. Estadística experimental. *Manual NBS 91*. Washington, DC, National Bureau of Standards, 1966.
- [19] Eisenhart, C.E. Contribución a la discusión de panel sobre el ajuste de las constantes fundamentales. *Spec. Publ.* 343, Washington, DC, National Bureau of Standards, 1970.
- [20] Boroviczény, G. de, (Ed.) Métodos eritrocitométricos y su normalización. pp 74-76; 108-116 (*Bibl. Haemat. FAsc.* 18, Karger Basel, 1964).
- [21] Van Kampen, E.J. y Zilstra, W.G. Espectrofotometría de la hemoglobina y derivados de la hemoglobina. *Adv. Clin. Chem.* **23** 1983: 199.
- [22] Catanzaro, E.J. y cols. Relación entre la abundancia isotópica absoluta y el peso atómico del rubidio terrestre. *J. Res. Nat. Bur. Stand.* **73A** (5) 1969: 511.
- [23] Uriano, G.A. y Gravatt, C.C. El papel de los materiales de referencia y métodos de referencia en análisis químico. *Crit. Rev. in Anal. Chem.* **6** 1977: 361.
- [24] Seward, R.W. (Ed.) Materiales de referencia patrones y mediciones objetivas. *Procedimientos del 6. Simposio de Recursos Materiales, Spec. Publ.* 408, Washington, DC, US Govt. Printing Office, 1975.
- [25] Schmitt, B.F. (Ed.) Producción y uso de materiales de referencia. *Procedimientos del Simposio Internacional sobre Producción y Uso de Materiales de Referencia* efectuado en el Bundesanstalt für Materialprüfung, Berlín (1980).
- [26] Gross, A.M. Robustez del intervalo de confianza con distribuciones simétricas de larga cola. *J. Am. Stat. Assoc.* **71** 1976: 409.
- [27] John, P.W.M. *Diseño estadístico y análisis de experimentos*, New York, NY, Macmillan, 1971, p. 76.
- [28] Searle, S.R. *Modelos lineales*, New York, NY, John Wiley, 1971.
- [29] NC ISO Guía 33, Uso de los materiales de referencia certificados. Primera Edición, 1989.