

Variabilidad intra e interlaboratorios en la determinación de la glucosa plasmática. Implicaciones para los estudios epidemiológicos y la práctica clínica

Salvador Peiró^a, Eugeni Berlanga^b, Francesc Prats^c, M. Àngels Ruiz^d, Carme Mora^e, Pilar Soriano^e, Marisa Urcola^f, M. Carme Villá^g y grupo EM-CCDD

^aFundación Instituto de Investigación en Servicios de Salud. ^bServicio de Laboratorio y Análisis Clínicos. Corporació Sanitària Parc Taulí. Sabadell. ^cCentro Inmunológico de Cataluña [CIC]. ^dServicio de Laboratorio y Análisis Clínicos. Hospital Esperit Sant. ^eServicio de Laboratorio y Análisis Clínicos. Hospital de Figueres. ^fServicio de Laboratorio y Análisis Clínicos. Institut d'Assistència Sanitària de Girona-Hospital Santa Caterina. ^gServicio de Laboratorio y Análisis Clínicos. Fundació Hospital Asil de Granollers.

Grupo EM-CCDD (Gestión Clínica y Utilización de la Analítica Diagnóstica Basadas en Evidencia Autogenerada-Estudio Multicéntrico Impacto del Cambio en los Criterios Diagnósticos de la Diabetes): Salvador Peiró, Julián Libero, Sergio García (Fundación Instituto de Investigación en Servicios de Salud), Francesc Prats (Centro Inmunológico de Cataluña [CIC]), M. Carme Villá, J.M. Rodriguez (Fundació Hospital Asil de Granollers), M. Àngels Ruiz (Hospital Esperit Sant), Carme Mora, Pilar Soriano (Hospital de Figueres), Marisa Urcola (Institut d'Assistència Sanitària de Girona - Hospital Sta. Caterina), Eugeni Berlanga, Montserrat Torra (Corporació Sanitària Parc Taulí) y Joan Josep Artells (Fundación Salud, Innovación y Sociedad).

Este estudio fue promovido y financiado por la Fundación Salud, Innovación y Sociedad.

Correspondencia: Salvador Peiró.

Fundación IISS. San Vicente, 112, 3. 46008 Valencia. España.

Correo electrónico: speiro@comv.es

Resumen

Fundamentos: La precisión de las determinaciones de glucemia es esencial para el correcto diagnóstico de la diabetes, y la variabilidad intra e interlaboratorios podría tener implicaciones relevantes desde el punto de vista clínico y epidemiológico.

Métodos: Evaluación de la exactitud y la precisión intra e interlaboratorios de las determinaciones de glucemia en cinco laboratorios hospitalarios de la red pública de Cataluña. Cada laboratorio analizó 20 diluciones duplicadas (40 muestras) de suero control líquido valorado, ciego a las titulaciones, a la existencia de muestras duplicadas y manteniendo las condiciones usuales de práctica. Se analizó la existencia de diferencias entre laboratorios y con los valores de referencia, entre muestras duplicadas del mismo laboratorio y las diferencias en el porcentaje de titulaciones que serían clasificadas en las diferentes categorías diagnósticas de diabetes.

Resultados: La concordancia de categorías diagnósticas entre los hospitales y las titulaciones valoradas fue muy elevada (estadísticos $\kappa > 0,80$). Los valores medios de cada lote fueron muy próximos entre sí y muy próximos a los de la titulación original estimada, con variaciones, entre 1 y 3 mg/dl, carentes de relevancia clínica. Tampoco existieron diferencias relevantes en el análisis intralaboratorio de las muestras duplicadas. El porcentaje de muestras de cada hospital en las diferentes categorías diagnósticas no demostró diferencias significativas.

Conclusion: La excelente exactitud y la precisión en la determinación de glucemia en los laboratorios de los centros analizados permite las comparaciones en los estudios epidemiológicos y reasegura a los clínicos en la calidad y exactitud de los datos ofrecidos por estos laboratorios.

Palabras clave: Glucemia. Control de calidad. Reproducibilidad de resultados. Diabetes mellitus/diagnóstico.

Summary

Background: Precision in blood glucose determinations is crucial for the correct diagnosis of diabetes. Intra- and inter-laboratory variability may have important implications in clinical practice and epidemiological studies.

Material and methods: We evaluated intra- and inter-laboratory precision and accuracy of blood glucose determinations in five hospital laboratories in the public health system of Catalonia, Spain. Each laboratory analyzed 20 duplicate dilutions of pre-evaluated control liquid serum (40 samples) under normal conditions. The laboratories were blind to the titration and the existence of duplicate samples. Inter-laboratory variations and differences with respect to reference values, between the duplicate samples in the same laboratory and in the percentage of titrations classified in each diagnostic category for diabetes were analyzed.

Results: The agreement in diagnostic categories between the hospitals and the pre-evaluated titrations was very high (κ -statistics > 0.80). The mean values for each lot were very close, and were very close to the original titration estimates, with variations of between 1 and 3 mg/dL, which were not clinically relevant. No significant differences were detected in the intra-laboratory analysis of the duplicate samples. The percentage of samples for each hospital in the different diagnostic categories showed no significant differences.

Conclusions: The laboratories of the hospitals analyzed showed excellent precision and accuracy in blood glucose determinations, thus allowing comparative epidemiological studies to be performed and reaffirming the quality and accuracy of the data supplied by these laboratories.

Key words: Blood glucose. Quality control. Reproducibility of results. Diabetes mellitus/diagnosis.

Introducción

Los criterios tradicionales para el diagnóstico de la diabetes mellitus fueron establecidos por la Asociación Americana de Diabetes (ADA) en 1979 y adoptados por la Organización Mundial de la Salud (OMS) entre 1980 y 1985¹. En 1997, la ADA propuso un cambio de estos criterios (tabla 1) rebajando desde 140 a 126 mg/dl el dintel de glucemia basal (GB) para el diagnóstico de diabetes y recomendando prescindir de la prueba de sobrecarga oral de glucosa (SOG) en la práctica clínica habitual^{2,3}. Estos nuevos criterios son objeto de polémica, no tanto por el descenso del dintel diagnóstico de la GB (que supone un incremento de la sensibilidad de la prueba para el diagnóstico de diabetes), sino por el abandono de la SOG, que tiene el efecto contrario y conllevaría un significativo incremento de la diabetes no diagnosticada, aspecto evidenciado en diversos estudios⁴⁻⁷, incluyendo estudios en diversos lugares de España⁸⁻¹¹. De forma reciente un comité asesor de la OMS se ha sumado a la recomendación de rebajar el dintel de la prueba de glucemia basal¹², pero manteniendo el uso de la SOG cuando los valores detectados se encuentran en el rango de valores intermedios (111 a 125 mg/dl).

En la actualidad, los estudios empíricos combinan los criterios ADA-1997 y OMS-1985, para acabar manejando cinco grupos diagnósticos de diabetes en relación con diversos valores de referencia: glucemia normal (glucemia en ayunas < 110 mg/dl y glucemia a las 2 h post-SOG < 140 mg/dl), glucemia en ayunas alterada (glucemia > 110 y < 126 mg/dl), intolerancia a la glucosa (glucemia a las 2 h post-SOG > 140 mg/dl y < 200 mg/dl), diabetes establecida según criterios ADA-1997 (glucemia al azar > 200 mg/dl,

acompañada de clínica, o dos muestras de glucemia en ayunas > 126 mg/dl) y diabetes establecida según criterios OMS-1985 (glucemia al azar > 200 mg/dl acompañada de clínica, o dos muestras de glucemia basal > 140 mg/dl o glucemia a las 2 h post-SOG > 200 mg/dl).

Dado que los datos de laboratorio son esenciales para establecer estas categorías diagnósticas, la precisión de las determinaciones de glucemia se convierte en un aspecto crítico para la correcta evaluación de los pacientes, y la variabilidad intra e interlaboratorios podría tener implicaciones relevantes desde el punto de vista clínico y epidemiológico. El objetivo de este estudio es analizar la exactitud y la precisión de las determinaciones de glucosa sérica intra y interlaboratorio, así como sus implicaciones clínicas y epidemiológicas.

Material y métodos

Diseño

Estudio de la exactitud y precisión intra e interlaboratorios de los resultados de las determinaciones de glucemia sérica en cinco servicios hospitalarios de análisis clínicos.

Entorno

Laboratorios clínicos de 5 hospitales comarcas (Hospital Esperit Sant, Hospital de Figueres, Corporació Sanitaria Parc Taulí de Sabadell, Hospital de Santa Caterina y Hospital General de Granollers) pertenecientes a la Xarxa Hospitalària d'Utilització Pública del Servei Català de la Salut.

Material

A partir de diluciones cruzadas de tres lotes primarios de suero control líquido valorado chemTRAK PLATINUM (Medical Analysis Systems, Inc.) se obtuvieron 10 lotes de suero con titulaciones de glucemia entre 103 y 172 mg/dl. Cada uno de estos lotes se distribuyó en 20 muestras de 2 ml, para obtener un total de 200 muestras. A cada laboratorio participante se le suministraron 40 de estas muestras, cuatro de cada una de las 10 titulaciones. Los tubos, facilitados por cada centro para mantener las condiciones habituales de laboratorio, fueron ordenados con valores de titulación no consecutivos y de forma diferente para cada laboratorio, e identificados con rotulador indeleble (una letra identificativa de cada centro y dos números consecutivos, del 1 al 40). Los laboratorios fueron ciegos a este orden, a los valores reales de las titulaciones y a la existencia de cuatro muestras repetidas de cada titulación. Los laboratorios analizaron las mismas muestras dos veces, en días diferentes, para un número final previsto de 400 determinaciones (10 titulaciones × 4 muestras de cada titulación × 5 laboratorios × dos determinaciones de cada muestra = 400 resultados analíticos), pero un tubo se destruyó accidentalmente antes de realizar la primera determinación (lote 5, hospital 1) por lo que el análisis final se realizó sobre 398 determinaciones.

Tabla 1. Criterios diagnósticos y categorías diagnósticas de diabetes según la OMS-1985 y la ADA-1997

	OMS-1985	ADA-1997
Síntomas	Síntomas de diabetes (poliuria, polidipsia, pérdida de peso sin otra causa) y glucemia plasmática al azar > 200 mg/dl (11,1 mmol/l)	Igual que OMS (1985)
Glucemia basal	Glucemia plasmática en ayunas > 140 mg/dl (7,7 mmol/l) en dos ocasiones*	Glucemia plasmática en ayunas > 126 mg/dl (7,0 mmol/l) en dos ocasiones*
Sobrecarga oral de glucosa	Glucemia plasmática a las 2 h de la sobrecarga oral de glucosa > 200 mg/dl (11,1 mmol/l).	Igual que OMS (1985), aunque se desaconseja el uso de la prueba de sobrecarga en la práctica clínica habitual

*Los resultados deben ser confirmados con la realización de una segunda prueba en otro día.

Métodos

Cada laboratorio utilizó sus métodos usuales para la preparación y análisis de las muestras, siempre con determinación de la glucosa en el tubo primario, intercalando los tubos del estudio en las series analíticas de la tarea diaria de pacientes y replicando las condiciones habituales de práctica. Cuatro laboratorios utilizaron el método de la glucosa-oxidasa, mientras que uno utilizó el de la hexocinasa. La segunda tanda de análisis se realizó en 24-48 h tras la primera, manteniendo las muestras en la nevera (+4 °C) durante el intervalo. Los resultados fueron introducidos directamente por cada centro en formato Excel™ a un sistema informático que disponía de un programa preparado específicamente para el estudio, remitidos por vía Internet y consolidados en una base datos común.

Análisis

En primer lugar, se describieron gráficamente los valores de glucosa obtenidos para cada una de las 10 titulaciones por los 5 centros participantes, y se valoró la presencia de diferencias estadísticamente significativas entre los valores medios de cada hospital (los valores de glucosa cumplían el supuesto de normalidad), y entre éstos y los valores medios de las titulaciones originales, mediante la verificación del no solapamiento de sus intervalos de confianza (IC) del 95%. A continuación se analizó la concordancia entre la proporción de pacientes incluidos en las cuatro categorías diagnósticas pre establecidas (normal: < 111 mg/dl; glucemia basal alterada: 111-125 mg/dl; diabético según criterios ADA-1997: 126-140 mg/dl; diabético según criterios OMS-1985: > 140 mg/dl), según las titulaciones originales y los valores obtenidos por cada hospital. Para este análisis se utilizó el coeficiente de correlación de Spearman y el estadístico κ de Cohen en sus versiones simple (κ_s) y ponderado (κ_w), ya que este último es más adecuado para analizar la concordancia entre más de dos categorías ordinales. El estadístico κ valora la diferencia entre el grado de acuerdo esperable por azar y el grado de acuerdo realmente obtenido; el primero recibe el valor 0, mientras el acuerdo perfecto no debido al azar recibe el valor 1. Aunque no existe un criterio estandarizado de valoración de los resultados de κ , se adoptaron las interpretaciones sugeridas por Landis y Fleiss ($\kappa < 0,0$: acuerdo pobre; $\kappa = 0,00-0,20$: acuerdo mínimo; $\kappa = 0,21-0,40$: acuerdo débil; $\kappa = 0,41-0,60$: acuerdo moderado; $\kappa = 0,61-0,80$: acuerdo sustancial; $\kappa = 0,81-1,00$: acuerdo casi perfecto)^{13,14}. A continuación, se analizaron las diferencias de medias entre laboratorios para cada uno de los 10 lotes utilizando el análisis de la varianza (*anova oneway*) y el análisis de varianza con efectos aleatorios (*oneway random effects*), que ofrece un estimador de la varianza debida a la muestra (desviación estándar muestra) y debida al hospital (desviación estándar hospital), permitiendo valorar el impacto de la variabilidad intra y interhospital en la varianza total. Se utilizó también el coeficiente de correlación intraclass que, en este contexto, indica la parte de varianza explicada por los factores muestra y hospital, y es interpretable de forma similar al coeficiente de determinación. Seguidamente se compararon, para cada hos-

tal, las medias de los cuatro valores de cada titulación analizados en la primera tanda, frente a los mismos valores de la segunda, utilizando el análisis de la varianza de medidas repetidas. Finalmente, se valoró mediante la prueba de la χ^2 si la proporción de muestras en cada uno de los cuatro grupos diagnósticos variaba entre hospitales. Todos los análisis se realizaron con el paquete estadístico STATA™ versión 6.0 (Stata Corporation, University Drive East, College Station, Texas).

Resultados

En la figura 1 se expone la distribución de los valores obtenidos por cada laboratorio en las ocho determinaciones realizadas de cada una de las 10 titulaciones analizadas, apreciándose una gran agrupación de los valores en todos los hospitales (excepto un valor extremo del hospital 1 en el lote 4) y el centrado en torno a la titulación original estimada, aunque los hospitales 3 y 5 tienden a ofrecer sistemáticamente valores algo por debajo de los estimados. En la tabla 2 se exponen los valores medios y el IC del 95% del total de determinaciones realizadas en cada uno de los hospitales y los que corresponderían a los valores de las titulaciones estimadas, sin que existan diferencias significativas entre hospitales ni de ninguno de los hospitales con los valores medios estimados. En la tabla 3 se presentan los estadísticos de concordancia entre la clasificación diagnóstica en cuatro grupos (normal, glucemia basal alterada, diabetes ADA-1997 y diabetes OMS-1985) derivada de las titulaciones originales y la derivada de los valores obtenidos por cada laboratorio. En todos los casos se obtuvieron coeficientes de correlación significativos y muy cercanos o superiores al 90%, y los estadísticos κ se situaron siempre en valores de acuerdo excelente, y significativamente diferentes de los esperables por azar.

En la tabla 4 se exponen los resultados del análisis de la varianza para cada uno de los 10 lotes analizados. En términos generales, los valores medios de cada lote eran muy próximos entre sí (apenas existen diferencias entre los valores máximos y mínimos) y muy próximos a los de la titulación original estimada, aunque con tendencia a situarse discretamente por debajo de ésta. Pese a que las diferencias entre hospitales fueron significativas en todos los casos (datos no expuestos) apenas variaban en un 1-3 mg/dl, por lo que carecían de relevancia clínica y se deben precisamente a que la varianza intrahospital es mínima. Los coeficientes de variación intrahospital (no expuestos en la tabla) eran inferiores al 2% con una excepción (debida al valor extremo del hospital 1 en la titulación 4). Los coeficientes de variación para cada titulación se situaron entre el 2 y el 3% (con excepción del lote donde se produjo el valor extremo, que tuvo un coeficiente de variación del 4,6%). En el análisis de efectos aleatorios la varianza debida al hospital fue siempre mayor que la debida a la muestra (otra vez, con excepción del lote con el valor extremo). En la tabla 5 se presentan los resultados del análisis de medidas repetidas entre la primera y la segunda tanda de cuatro muestras del mismo lote. Sólo en cinco de las 50 comparaciones realizadas se encontraron diferencias significativas entre las dos tandas de análisis, e

Figura 1. Distribución de los valores obtenidos por cada hospital para cada titulación original. H: hospital; la línea corresponde al valor de la titulación original de cada lote y los puntos a los valores de cada una de las determinaciones realizadas por cada hospital (n = 8 por hospital y lote, excepto para el lote 5 del hospital 1, con n = 6).

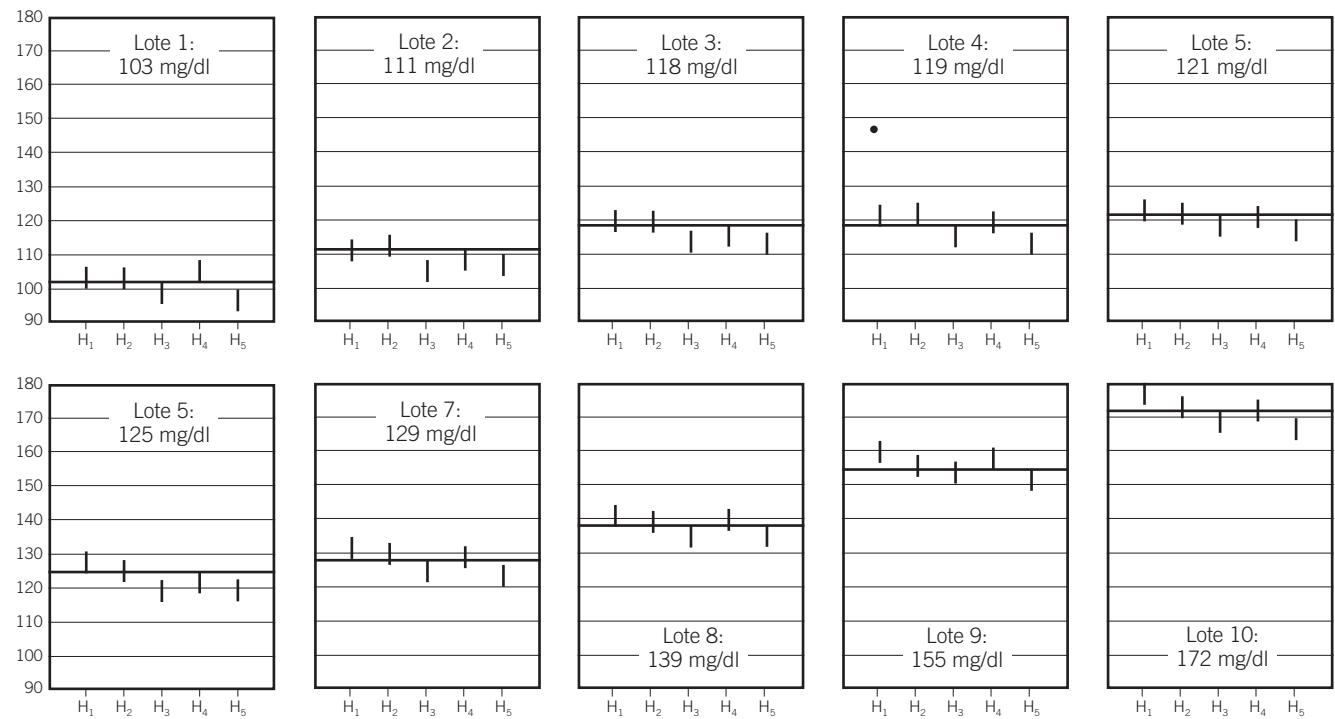


Tabla 2. Media e intervalos de confianza de los valores de glucemia en los diferentes hospitales y en las titulaciones de origen

	n	Titulación estimada (mg/dL)	IC del 95%	Mínimo	Máximo
Hospital	1	78	131,36	126,43	136,29
	2	80	129,59	125,07	134,10
	3	80	125,13	120,61	129,64
	4	80	129,00	124,41	133,59
	5	80	123,96	119,49	128,44
Total	398	127,79	125,75	129,83	96 179
Original	80	129,20	124,76	133,63	103 172

IC: intervalo de confianza del 95%.

incluso en estos casos se trataba de diferencias mínimas sin relevancia clínica.

En la tabla 6 se expone la distribución de los valores en las categorías diagnósticas de diabetes en los diferentes hospitales. La prevalencia de valores normales en la muestra variaría desde el 12,8% en el hospital 1 al 21,3% en el hospital 3, mientras que la prevalencia de diabetes ADA en la muestra apenas superaría el 30% para el hospital 5, y se situaría en torno al 48% para el hospital 1, aunque estas diferencias no fueron significativas.

Tabla 3. Concordancia entre la proporción de valores en cada categoría diagnóstica según las titulaciones originales y los valores obtenidos por cada hospital

	r _{Spearman}	K _s	K _w	IC del 95%	K _w
Hospital 1	0,91*	0,78*	0,90*	0,83	0,97
Hospital 2	0,91*	0,76*	0,93*	0,89	0,97
Hospital 3	0,93*	0,80*	0,92*	0,87	0,96
Hospital 4	0,95*	0,82*	0,94*	0,90	0,97
Hospital 5	0,89*	0,72*	0,90*	0,85	0,95

*p < 0,001; K_s: kappa simple; K_w: kappa ponderado; IC: intervalo de confianza del 95% del estadístico kappa ponderado.

Categorías diagnósticas: normalidad, glucemia basal alterada, diabetes ADA-1997 y diabetes OMS-1985.

Discusión

Los resultados obtenidos ponen de manifiesto una excelente precisión intralaboratorio, con una buena agrupación de valores, siempre muy próximos al valor estimado del control, coeficientes de variación mínimos y una excelente reproducibilidad entre la primera y la segunda tanda de análisis del mismo lote. La concordancia entre laboratorios fue también muy elevada, aunque dos hospitales tienden a ofrecer valores

Tabla 4. Análisis de varianza de una vía y análisis de varianza con efectos aleatorios

Lote	Titulación original, mg/dl	Valor mínimo, mg/dl	Valor máximo, mg/dl	Valor promedio, mg/dl	Coeficiente de variación	Coeficiente de correlación intraclace	Desviación típica, muestra	Desviación típica, hospital
Lote 1	103	97,50	102,93	100,91	2,36	0,75	1,24	2,24
Lote 2	111	105,75	111,50	108,85	2,77	0,64	1,86	2,61
Lote 3	118	112,87	118,37	115,76	2,38	0,76	1,41	2,61
Lote 4*	119	113,75	123,62	118,33	4,62	0,37	4,43	3,54
Lote 5	121	115,75	122,00	118,92	2,35	0,64	1,73	2,43
Lote 6	125	119,25	125,87	122,92	2,54	0,79	1,50	3,01
Lote 7	129	123,25	130,75	127,65	2,53	0,76	1,66	3,05
Lote 8	139	133,70	141,00	137,85	2,68	0,77	1,83	3,55
Lote 9*	155	150,37	159,00	154,99	2,46	0,74	2,04	3,56
Lote 10	172	166,87	177,00	171,24	2,32	0,79	1,91	3,86

p < 0,05 en todos los casos. *No cumplen el supuesto de homogeneidad de varianzas; los valores mínimo y máximo se refieren a los valores medios obtenidos por cada hospital para la correspondiente muestra y el valor promedio al conjunto de hospitales.

Tabla 5. Comparación de resultados entre los análisis realizados de las mismas muestras. Análisis de varianza de medidas repetidas

Hospital 1		Hospital 2		Hospital 3		Hospital 4		Hospital 5		
Lote	Tanda1	Tanda2	Tanda1	Tanda 2	Tanda 1	Tanda 2	Tanda 1	Tanda 2	Tanda 1	Tanda 2
Lote 1	102,00	102,25	103,16	102,71	98,75	100,50	102,00	102,75	96,75	98,25
Lote 2	112,75	110,20	110,82	111,27	104,75	106,75	109,50	110,00	105,25	107,25
Lote 3	118,75	118,00	118,03	118,03	*111,75	114,25	115,50	117,25	*112,00	113,75
Lote 4	127,25	120,00	120,73	119,83	115,50	115,75	118,25	118,50	113,00	114,50
Lote 5	122,66	121,33	119,83	120,73	117,25	117,00	119,50	121,00	115,00	116,50
Lote 6	*126,50	125,25	125,23	124,78	119,00	121,00	124,25	124,75	118,25	120,25
Lote 7	131,50	130,00	129,29	129,74	125,50	125,75	127,75	130,50	*122,00	124,50
Lote 8	141,50	140,50	140,55	138,75	133,50	134,75	140,00	141,50	133,00	134,50
Lote 9	158,75	159,25	*159,02	154,97	151,50	152,50	156,50	156,75	149,25	151,50
Lote 10	177,00	177,00	171,64	172,54	167,75	169,00	171,75	172,00	166,00	167,75

*p < 0,05 entre primera y segunda tanda de análisis del mismo lote. N tanda₁ = 4 y n tanda₂ = 4, excepto para el lote 4 del hospital 1 con n = 2.

Tabla 6. Distribución de los valores obtenidos por cada hospital según grupos diagnósticos

	< 111		111-125		126-140		> 140		Total		
	n	%	n	%	n	%	n	%	n	%	
Hospital	1*	10	12,8	30	38,5	17	21,8	21	26,9	78	100,0
	2	13	16,3	32	40,0	16	20,0	19	23,8	80	100,0
	3	17	21,3	35	43,8	12	15,0	16	20,0	80	100,0
	4	13	16,3	36	45,0	11	13,8	20	25,0	80	100,0
	5	16	20,0	39	48,8	9	11,3	16	20,0	80	100,0
	Total	69	17,3	172	43,2	65	16,3	92	23,1	398	100,0
Original/N = 80	8	10,0	40	50,0	16	20,0	16	20,0	80	100,0	

*Dos casos sin datos; p (χ^2) entre hospitales: no significativa.

algo menores (entre 1 y 5 mg/dl) respecto a los resultados de los otros tres hospitales. La prevalencia de cada una de las cuatro categorías diagnósticas en la muestra analizada no presentó diferencias significativas entre hospitales; los laboratorios 3 y 5 evidenciaron una no significativa menor proporción de diabetes ADA y OMS (asociada con los valores más bajos de sus determinaciones) que probablemente alcanzaría significación estadística en una muestra mayor. No obstante, esta situación deriva de la presencia en la muestra de tres titulaciones estimadas en los valores frontera (111, 125 y 139 mg/dl), en las que mínimas variaciones en el resultado de la determinación, producían la inclusión en categorías diagnósticas diferentes. De hecho, la mayor parte de las discordancias se deben a la clasificación de muestras de estas titulaciones en categorías diagnósticas contiguas. Aunque teóricamente este efecto podría afectar la prevalencia de diabetes en áreas sanitarias servidas por diferentes hospitales, su relevancia real dependerá de la prevalencia poblacional de los valores frontera entre grupos diagnósticos.

Los estudios que evalúan la precisión intra o interlaboratorios clínicos ponen de manifiesto resultados muy dispares¹⁴⁻²⁶. Por ejemplo, de los 1.500 laboratorios estadounidenses que en 1990 participaron en el Quality Assurance Service del College of American Pathologists sólo el 61% cumplía los requisitos de precisión establecidos reglamentariamente por la Clinical Laboratory Improvement Amendments de 1988²⁵, aunque en 1994 el 97% de los laboratorios hospitalarios y el 91% de los de otro tipo cumplían con las especificaciones de esta norma²⁶. Esta variabilidad hace necesaria su monitorización en los estudios clínicos y epidemiológicos multicéntricos²⁷; de hecho, el objetivo primario de este estudio era valorar si la prevalencia en cada una de las cuatro categorías diagnósticas estaba o no afectada por la variabilidad en las determinaciones realizadas por cada laboratorio como paso previo a un estudio de prevalencia de diabetes. En todo caso, y más allá de los estudios, se trata de un aspecto esencial para la práctica clínica^{28,29} y la transmisión de esta información a los clínicos debería ser valorada.

Entre las limitaciones de este estudio hay que citar, en primer lugar, la ya comentada presencia de valores frontera, con el efecto de incrementar las diferencias en el análisis por categorías diagnósticas, pese a la gran similitud en los valores obtenidos. Igualmente, hay que señalar que, pese al cegado al orden y valores de las titulaciones estimadas, los laboratorios participantes conocían que el estudio se estaba llevando a cabo y las muestras eran fácilmente identificables por su aspecto, por lo que pudieron desarrollar comportamientos reactivos, tipo efecto Hawthorne u otros, modificando su comportamiento habitual y tratando las muestras con mayores cuidados. En este sentido, algunos estudios han demostrado que la variabilidad es mayor cuando los laboratorios no están informados de que se está realizando una evaluación de la precisión diagnóstica³⁰. Respecto al análisis, hay que señalar que el estadístico κ , que sólo puede alcanzar el valor 1 cuando las celdas fuera de la diagonal están vacías, infraestima en alguna medida el grado de acuerdo³¹, que sería todavía algo mayor que el hallado.

Las pruebas de laboratorio son una extensión de la evaluación física de los pacientes que pueden orientar o confirmar el diagnóstico y guiar diversas decisiones clínicas, aspecto especialmente relevante en el caso de la diabetes. El estudio realizado evidencia excelentes exactitud y precisión en la medición de este parámetro en los laboratorios de los cinco hospitales analizados, permitiendo las comparaciones en los estudios epidemiológicos y, sobre todo, reasegurando a los clínicos en la calidad y exactitud de los datos ofrecidos por estos centros.

Bibliografía

1. World Health Organization. WHO Study Group. Diabetes mellitus. Technical report series 727. Geneva: OMS, 1985.
2. American Diabetes Association. Report of the Expert Committee on the diagnosis and classification of diabetes mellitus. Diabetes Care 1997;20:1183-97.
3. American Diabetes Association. Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Clinical Practice Recommendations 2000. Diabetes Care 2000; 23(Suppl 1).
4. The DECODE Study Group on behalf of the European Diabetes Epidemiology Group. Will new diagnostic criteria for diabetes mellitus change phenotype of patients with diabetes? Reanalysis of European epidemiological data. BMJ 1998; 317: 371-5.
5. The DECODE Study Group on behalf of the European Diabetes Epidemiology Group. Is fasting glucose sufficient to define diabetes? Epidemiological data from 20 European centers. Diabetología 1999;42:647-54.
6. Wahl PW, Savage PJ, Psaty BM, Orchard TJ, Robbins JA, Tracy RP. Diabetes in older adults: comparison of 1997 American Diabetes Association classification of diabetes mellitus with 1985 WHO classification. Lancet 1998;352:1012-5.
7. De Vegt F, Dekker JM, Stehouwer CDA, Nijpels G, Bouter LM, Heine RJ. The 1997 American Diabetes Association criteria versus the 1985 World Health Organization criteria for the diagnosis of abnormal glucose tolerance: poor agreement in the Hoorn Study. Diabetes Care 1998;21:1686-90.
8. Costa B, Franch J, Martín F, Morató J, Donado A, Basora J, IGT Research Group. Impact of the American Diabetes Association Diagnosis Criteria on High-Risk Spanish Population. Diab Res Clin Pract 1999;46:75-81.
9. Martín Luján F, Costa Pinel B, Donado-Mazarrón Romero A, Basora Gallissà T, Basora Gallissà J, Piñol Moreso JL, et al. Los criterios de la ADA infravaloran el impacto de la diabetes en la población española de más alto riesgo. Aten Primaria 2000; 26:517-24.
10. Alberti KGMM, Zimmet PZ, for the WHO Consultation. Definition, diagnosis and classification of diabetes mellitus and its complications (I). Diagnosis and classification of diabetes mellitus: provisional report of a WHO consultation. Diabet Med 1998;15:539-53.
11. Pulgar Suárez M, Gómez Guedes P, Aguado Díaz M, Menéndez Álvarez S, García Garaboa A, Rodríguez González I, et al. Vali-

- dez de los nuevos criterios diagnósticos de la diabetes tipo 2. Impacto de su aplicación en un área de salud. Aten Primaria 2001;27:111-5.
12. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:671-9.
 13. Fleiss JL. The measurement of interrater agreement. In: Fleiss JL, editor. *Statistical methods for rates and proportions*. New York: John Wiley & Sons, 1981; p. 212-36.
 14. Bjorkhem I, Bergman A, Falk O, Kallner A, Lantto O, Svensson L, et al. Accuracy of some routine method used in clinical chemistry as judged by isotope dilution-mass spectrometry. *Clin Chem* 1981;27:733-5.
 15. Rej R, Jenny RW. How good are clinical laboratories? An assessment of current performance. *Clin Chem* 1992;38:1210-7.
 16. De Gortari E, Herrera M, Loría A, Terrés A, González-Salayandia MA, Hernández MA. Programa piloto en laboratorios clínicos mexicanos (III). Estrategia para evaluar la calidad de los resultados. *Salud Pública Mex* 1994;36:484-91.
 17. Watson JE, Evans RW, Germanowski J, Caggiula AW, Olson MB, Hauth BA. Quality of lipid and lipoprotein measurements in community laboratories. *Arch Pathol Lab Med* 1997; 121:105-9.
 18. Anónimo. Clinical laboratory performance on proficiency testing samples. United States, 1994. *MMWR* 1996;45:193-6.
 19. Linko S, Himberg JJ, Thienpont L, Stockl D, De Leenheer A. Assessment of the state-of-the-art trueness and precision of serum total-calcium and glucose measurements in Finnish laboratories –the QSL Finland Study. *Scand J Clin Lab Invest* 1998;58:265-8.
 20. Henderson R, Krishnan S, Webb S, Cheung CM, Nazir DJ, Richardson H. Proficiency testing of creatine kinase and creatine kinase-2: the experience of the Ontario Laboratory Proficiency Testing Program. *Clin Chem* 1998;44:124-33.
 21. Novis DA, Jones BA. Interinstitutional comparison of bedside blood glucose monitoring program characteristics, accuracy performance, and quality control documentation: a College of American Pathologists Q-Probes study of bedside blood gluco-
se monitoring performed in 226 small hospitals. *Arch Pathol Lab Med* 1998;122:495-502.
 22. Pfeiffer CM, Huff DL, Smith SJ, Miller DT, Gunter EW. Comparison of Plasma Total Homocysteine Measurements in 14 Laboratories: an international study. *Clin Chem* 1999;45: 1261-8.
 23. Pfeiffer CM, Smith SJ, Miller DT, Gunter EW. Comparison of Serum and Plasma Methylmalonic Acid Measurements in 13 Laboratories: An International Study. *Clin Chem* 1999;45: 2236-42.
 24. Ross JW, Miller WG, Myers GL, Praestgaard J. The accuracy of laboratory measurements in clinical chemistry: a study of 11 routine chemistry analytes in the College of American Pathologists Chemistry Survey with fresh frozen serum, definitive methods, and reference methods. *Arch Pathol Lab Med* 1998; 122:587-608.
 25. Westgard JO, Seehafer JJ, Barry PL. European specifications for imprecision and inaccuracy compared with operating specifications that assure the quality required by US CLIA proficiency-testing criteria. *Clin Chem* 1994; 40:1228-32.
 26. Stull TM, Hearn TL, Hancock JS, Handsfield JH, Collins CL. Variation in proficiency testing performance by testing site. *JAMA* 1998;279:463-7.
 27. Cull CA, Manley SE, Stratton IM, Neil HAW, Ross IS, Holman RR, et al. Approach to maintaining comparability of biochemical data during long-term clinical trials. *Clin Chem* 1997;43: 1913-8.
 28. Petersen PH, De Verdier CH, Groth T, Fraser CG, Blaabjerg O, Horder M. The influence of analytical bias on diagnostic misclassifications. *Clin Chim Acta* 1997;260:189-206.
 29. Fraser CG. The necessity of achieving good laboratory performance. *Diabet Med* 1990;7:490-3.
 30. Parsons PJ, Reilly AA, Esenrio-Jenssen D, Werk LN, Mofenson HC, Stanton NV, et al. Evaluation of blood lead proficiency testing: comparison of open and blind paradigms. *Clin Chem* 2001; 47:322-30.
 31. Soeken KL, Prescott PA. Issues in the use of kappa to estimate reliability. *Med Care* 1986;733-41.