Acta Bioquímica Clínica Latinoamericana
ISSN: 0325-2957
ISSN: 1851-6114
actabioq@fbpba.org.ar
Federación Bioquímica de la Provincia de Buenos Aires
Argentina

Cobbaert, Christa; Weykamp, Cas; Romijn, Fred

EQA-based evaluation of metrological traceability of clinical chemistry test results in Argentina

Acta Bioquímica Clínica Latinoamericana, vol. 53, no. 4, 2019, September-, pp. 487-497
Federación Bioquímica de la Provincia de Buenos Aires
Argentina

Homenaje a la trayectoria del Prof. Dr. Daniel Mazziotta†

EQA-based evaluation of metrological traceability of clinical chemistry test results in Argentina

Christa Cobbaert[1a*], Cas Weykamp[2b], Fred Romijn[3a]

[1] Pharmaceutical Sciences, PhD, EuSpLM (European Specialist in Lab Medicine) and Full professor in Clinical Chemistry and Laboratory Medicine.
[2] PhD.
[3] Senior research technician.

[a] Leiden University Medical Centre, Leiden, the Netherlands.
[b] MCA, Queen Beatrix Hospital, Winterswijk, the Netherlands.

* Autor para correspondencia.

Acta Bioquímica Clínica Latinoamericana
Incorporada al Chemical Abstract Service.
Código bibliográfico: ABCLDL.
ISSN 0325-2957 (impresa)
ISSN 1851-6114 (en línea)
ISSN 1852-396X (CD-ROM)

Preambule

A tribute to a highly respected Argentinian inspiring mind and ambassador of Laboratory Medicine

Prof. Mazziotta from Argentina is remembered as a warm-hearted colleague and inspiring mind who contributed to worldwide IFCC reputation in several functions. Prof. Mazziotta was very successful in progressing standardization of medical tests in Latin-America with boundless energy. In that context, Prof. Mazziotta actively participated in the Cholesterol Reference Method Laboratory Network (CRMLN) from CDC, Atlanta, Georgia, USA for improving lipid standardization in Argentina and Latin-America. During one of the yearly CRMLN meetings at CDC, about two decades ago, I first met Daniel. With Daniel's spreading enthusiasm we learned to know one another well and shared our common interest for improving metrological traceability of test results. In 2013 the idea was born to evaluate the standardization status of 16 routine clinical chemistry tests in Argentina, using value-assigned EQA-samples from the Dutch accuracy-based EQA scheme. Notwithstanding the regulatory hurdles that professor Mazziotta faced to import the EQA-specimens from the Netherlands for distribution to the Argentinian participating labs, he found a way out. The data were presented during the CALILAB congress in Mar del Plata, Argentina, hosted by Prof Dr. Mazziotta. This tribute is for keeping alive the great memories we have to this colleague-friend and great ambassador of laboratory medicine in Latin-America.

Abstract

Equivalence of results among laboratories is a major mission for medical laboratories. In the Netherlands, medical laboratories only use homogenous, commercial for general chemistry analytes, whereas in Argentina heterogenous, *home brew* test applications are common. The effect of this practice difference on test accuracy is studied using key features of the accuracy-based EQA program of the Netherlands. Six frozen, human-based, commutable poolsera, covering the (patho) physiological measuring range for 17 general chemistry analytes, were assayed by ~75 Argentinian labs and ~200 Dutch laboratories in 2014. After removal of outliers, harmonization status among laboratories was evaluated by calculating overall mean interlaboratory coefficients of variation (CVs, %) per analyte and per country for all 6 levels. Evenso, standardization status was evaluated after removal of outliers by calculating overall mean recoveries (%) as compared to the assigned target values per analyte per country for all 6 levels. Absolute median biases were compared to (minimal/desirable) biases derived from biological variation criteria. For serum enzymes *interlaboratory CVs* in the Argentinian laboratories ranged between 10 and 22%, as compared to 3-6% in the Netherlands. For serum uric acid, creatinine, glucose and total protein, interlaboratory CVs varied between 4.3 and 13.1% in Argentinian labs, as compared to <3.5% in the Netherlands. For serum electrolytes, interlaboratory CVs ranged between 1.8 and 3.8% for Na^+; 2.9-5.8% for Cl^-; 3.8-7.5% for K^+; 9.4-10.4% for Ca^{2+} and 16.2-22.3% for Mg^{2+} as compared to ≤2% (Na^+, K^+, Cl^-, Ca^{2+}) and ≤3% (Mg^{2+}) in the Netherlands. Mean recoveries in Argentinian laboratories for e.g. serum creatinine, glucose, CK, Ca^{2+} and Na^+ were 95-119%; 95-104%; 98-102%; 98-102% and 96-100% respectively, whereas min-max recovery ranges were 65-155%; 58-126%; 47-132%; 66-132% and 85-115%. In the Netherlands, absolute mean recoveries were overall 98.9% with a SD of 2.0%. Median biases in Argentinian laboratories ranged from -2.9 to 18.2%; -3.1 - 2.6%; -3.3 - 0.5%; -1.1 - 3.8% and -4.3-0% for serum creatinine, glucose, CK, Ca^{2+} and Na^+. In the Netherlands overall mean/median biases were 1.1% (SD=2.0%). Exchange of commutable, value-assigned EQA-materials was helpful for studying the harmonization and standardization status of medical tests in Argentina, and for revealing the future harmonization and standardization potential. The results clearly demonstrate that metrological traceability of test results in Argentina is on average in line with what is expected; yet, the spreading among laboratories is far too high and should be improved.

Keywords: Metrological traceability; Clinical Chemistry; Argentina; the Netherlands

Evaluación de la trazabilidad metrológica de los resultados de química clínica, basada en EQA, en la Argentina

Resumen

La equivalencia de resultados entre laboratorios es una misión importante para los laboratorios médicos. En los Países Bajos, los laboratorios médicos solo usan aplicaciones comerciales homogéneas, regulatoriamente aprobadas (CE-IVD) para analitos químicos, mientras que en la Argentina son comunes las aplicaciones heterogéneas caseras. El efecto de esta diferencia práctica en la precisión de la prueba se estudia utilizando características clave del programa EQA, basado en la precisión, de los Países Bajos. Se ensayaron seis pools de sueros, congelados, de origen humano, conmutables, que cubrían el rango de medidas (pato)fisiológicas para 17 analitos de química clínica. Estos analitos de química clínica fueron analizados por ~75 laboratorios argentinos y ~200 laboratorios holandeses en 2014. Después de eliminar los valores atípicos, el estado de armonización entre los laboratorios fue evaluado calculando los coeficientes de variación interlaboratorios medios globales (CV%) por analito y por país para los 6 niveles. No obstante, el estado de estandarización se evaluó después de la eliminación de valores atípicos mediante el cálculo de recuperaciones medias generales (%) en comparación con los valores asignados por analito por país para los 6 niveles. Los sesgos medios absolutos se compararon con los sesgos (mínimos / deseables) derivados de los criterios de variación biológica. Para enzimas séricas los CV interlaboratorio en los laboratorios argentinos oscilaron entre 10 y 22%, en comparación con 3-6% en los Países Bajos. Para el ácido úrico sérico, creatinina, glucosa y proteínas totales, los CV entre laboratorios variaron entre 4,3 y 13,1% en los laboratorios argentinos, en comparación con <3,5% en los

Países Bajos. Para los electrolitos séricos, los CV interlaboratorios oscilaron entre 1,8 y 3,8% para Na^+; 2,9-5,8% para Cl^-; 3,8-7,5% para K^+; 9,4-10,4% para Ca^{2+} y 16,2-22,3% para Mg^{2+} en comparación a ≤2% (Na^+, K^+, Cl^-, Ca^{2+}) y ≤3% (Mg^{2+}) en los Países Bajos. Las recuperaciones medias en laboratorios argentinos para, p.ej. la creatinina sérica, glucosa, CK, Ca^{2+} y Na^+ fueron 95-119%; 95-104%; 98-102%; 98-102% y 96-100% respectivamente, mientras que los rangos de recuperación min-max fueron 65-155%; 58-126%; 47-132%; 66-132% y 85-115%. En los Países Bajos, las recuperaciones medias absolutas fueron en general del 98,9% con una desviación estándar (DE) del 2,0%. La mediana de los sesgos medios de los laboratorios argentinos osciló entre -2,9 y 18,2%; -3,1 - 2,6%; -3,3 - 0,5%; -1,1 - 3,8% y -4,3-0% para creatinina sérica, glucosa, CK, Ca^{2+} y Na^+. En los Países Bajos, las medias / medianas en general fueron de 1,1% (DE=2,0%). El intercambio de los valores asignados a los materiales EQA, conmutables fue de gran ayuda para la armonización y estandarización de los ensayos médicos en la Argentina y para revelar el potencial futuro de armonización y estandarización. Estos resultados claramente demuestran que la trazabilidad metrológica de los resultados de las pruebas en la Argentina está, en promedio, de acuerdo con lo esperable; sin embargo, la dispersión entre laboratorios es muy grande y debería ser mejorada.

Palabras clave: Trazabilidad metrológica; Química Clínica; Argentina; Países Bajos

Avaliação da rastreabilidade metrológica dos resultados da química clínica, baseada em EQA, na Argentina

Resumo

A equivalência de resultados entre laboratórios é uma missão importante para os laboratórios médicos. Nos Países Baixos, os laboratórios médicos só utilizam aplicações comerciais homogêneas, aprovadas por regulações (CE-IVD) para analitos químicos, ao passo que na Argentina são comuns as aplicações heterogêneas caseiras. O efeito desta diferença prática na exatidão do teste é estudado utilizando características essenciais do programa EQA, dos Países Baixos, baseado na exatidão. Foram ensaiados seis pools de soros, congelados, de origem humana, comutáveis, que abrangiam a faixa de medidas (pato)fisiológicas para 17 analitos químicos gerais. Esses analitos químicos foram analisados por ~75 laboratórios argentinos e ~200 laboratórios holandeses em 2014. Após eliminar os valores atípicos, o estado de harmonização entre os laboratórios foi avaliado através do cálculo dos coeficientes de variação interlaboratório meios globais (CV%) por analito e por país para os 6 níveis. Não obstante, o estado de padronização foi avaliado depois da eliminação de valores atípicos pelo cálculo de recuperações médias gerais (%) se comparados com os valores atribuídos por analito por país para os 6 níveis. Os vieses médios absolutos foram comparados com os vieses (mínimos / desejáveis) decorrentes dos critérios de variação biológica. Para enzimas séricas, os CV interlaboratório nos laboratórios argentinos oscilaram entre 10 e 22%, em comparação com 3-6% nos Países Baixos. Para o ácido úrico sérico, creatinina, glicose e proteínas totais, os CV entre laboratórios variaram entre 4,3 e 13,1% nos laboratórios argentinos, em comparação com <3,5% nos Países Baixos para os eletrólitos séricos, os CV interlaboratórios oscilaram entre 1,8 e 3,8% para Na^+; 2,9-5,8% para Cl^-; 3,8-7,5% para K^+; 9,4-10,4% para Ca^{2+} e 16,2-22,3% para Mg^{2+} em comparação com ≤2% (Na^+, K^+, Cl^-, Ca^{2+}) e ≤3% (Mg^{2+}) nos Países Baixos. As recuperações médias em laboratórios argentinos para, p.ex. a creatinina sérica, glicose, CK, Ca^{2+} e Na^+ foram 95-119%; 95-104%; 98-102%; 98-102% e 96-100% respectivamente, enquanto que os intervalos de recuperação min-máx. foram 65-155%; 58-126%; 47-132%; 66-132% e 85-115%. Nos Países Baixos, as recuperações médias absolutas foram em geral de 98,9% com um desvio padrão (DE) de 2,0%. A mediana dos vieses médios dos laboratórios argentinos oscilou entre -2,9 e 18,2%; -3,1 - 2,6%; -3,3 - 0,5%; -1,1 - 3,8% e -4,3-0% para creatinina sérica, glicose, CK, Ca^{2+} e Na^+. Nos Países Baixos, as médias / medianas em geral foram de 1,1% (DE=2,0%). O intercâmbio dos valores atribuídos aos materiais EQA, comutáveis, foi de grande ajuda para a harmonização e padronização dos ensaios médicos na Argentina e para revelar o potencial futuro de harmonização e padronização. Esses resultados demonstram às claras que a rastreabilidade metrológica dos resultados dos testes na Argentina está de acordo com o esperável; a dispersão entre laboratórios ainda é muito grande e deveria ser melhorada.

Palavras-chave: Rastreabilidade metrológica; Química Clínica; Argentina; Países Baixos

Introduction

Metrology supports most of what we do, already since the start of civilization in ancient times. Nowadays, all forms of chemical and physical measurements affect the quality of all our activities in the world in which we live. Because of the need for international agreement on matters concerning metrology, an international treaty known as the Metre Convention (https://www.bipm.org/en/worldwide-metrology/metre-convention/) was established in Paris in 1875 across 17 nations. This treaty founded the International Bureau of Weights and Measures (BIPM) and remains today the basis of international agreement on units of measurement (https://www.bipm.org/en/measurement-units/). So far, seven SI-base units have been defined: the kilogram, the mole, the meter, the second, the ampere, the Kelvin and the Candela.

In the healthcare sector, good clinical practice relies on interchangeability of medical test results worldwide. Consequently, laboratory and clinical professional societies, guideline developers and regulators strive for medical test result equivalence based on traceability hierarchies to standards and methods of higher order and to SI-units when feasible, for the sake of safe and clinically effective patient care. As an answer to the publication of the European IVD Directive (IVDD) 98/79/EC on *in vitro* medical devices the Joint Committee on Traceability in Laboratory Medicine (JCTLM) was established in 2002 through a Declaration of Cooperation between the International Bureau of Weights and Measures (BIPM), the International Federation for Clinical Chemistry and Laboratory Medicine (IFCC), and the International Laboratory Accreditation Cooperation (ILAC). The Declaration of Cooperation was revised in 2012, following the revision of harmonized standards relating to higher order reference materials and methods, and re-signed in April 2016. JCTLM's goals are to provide a worldwide platform to promote and give guidance on internationally recognized and accepted equivalence of measurements in Laboratory Medicine and traceability to appropriate measurement standards. As the IVDD demands traceability of results of calibrators and controls to higher order reference materials and methods, JCTLM responded to these needs by establishing lists of suitable higher-order reference materials, higher-order reference measurement procedures and reference measurement services for laboratory medicine. With the creation of the JCTLM a framework has been established which can be used for the international recognition of such materials, procedures and measurement services from laboratories. To that end, the JCTLM framework laid down a process whereby reference materials, reference measurement procedures and reference measurement services are examined with respect to conformity with appropriate international documentary ISO-standards. The output is the JCTLM database of available higher-order reference materials and higher-order reference measurement procedures as well as reference measurement services provided by calibration laboratories (formerly named reference laboratories) that can be used by downstream users such as the IVD-industry and EQAS-organizers to meet and evaluate traceability requirements for *in vitro* diagnostic and laboratory medicine measurements. The proper use of available higher-order reference materials and reference measurement procedures demands competent calibration laboratories for specified measurands. Such calibration laboratories should be able to demonstrate their technical competence in the operation of a reference measurement procedure of higher order for a given measurand with demonstrated traceability and measurement uncertainty. The IFCC Committee on Traceability in Laboratory Medicine (C-TLM) monitors the performance of calibration labs by means of periodic ring trials (http://www.dgkl-rfb.de:81/).

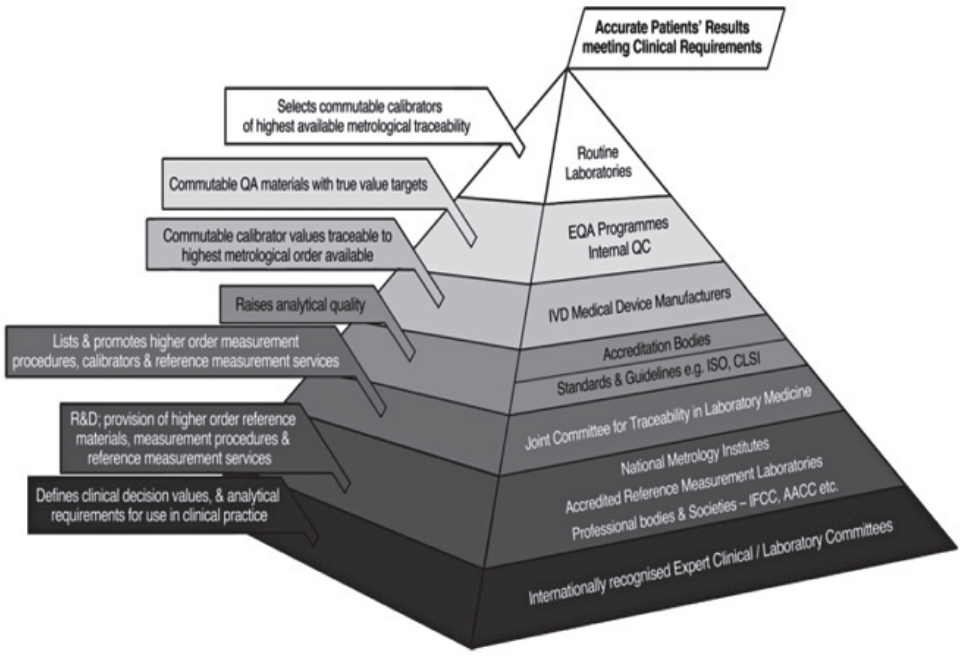
The higher-order reference materials, reference measurement procedures and reference measurement services that get JCTLM-endorsement are published in a database by the BIPM and are publicly available with links from the IFCC website and other websites as necessary. The database is for use by all stakeholders of the medical test traceability hierarchy, among them National Metrology Institutes (NMIs) and other producers of higher-order reference materials; IVD-industry; professional societies, clinical laboratories and researchers in the field of laboratory medicine; quality assurance organizations in the field of laboratory medicine; regulatory authorities; notified bodies and international organizations.

The development, deployment and quality assurance of higher-order reference materials is wholly the responsibility of the producers of the reference materials as these materials should meet their stated specifications and continue to be available. Analogously, the responsibility for reference measurement procedures meeting their stated specifications is the duty of the calibration laboratories that perform the measurements. In both cases, it is not the responsibility of the JCTLM or the member organizations of the JCTLM Executive Committee.

Although JCTLM was established in response to the European IVDD, its outreach is global. It is important to mention that the JCTLM framework has an exclusively recommendatory nature. Because the JCTLM framework does not have any binding legal effect in national and international laws, the adoption of the metrological traceability concept across the globe is variable and slow. Although successful global standardization efforts have been realized for some disease defining analytes, such as cholesterol, HbA1c, and serum

creatinine but also for serum enzymes. The global success of these standardization programmes is explained by (a) the involvement of influential clinical societies such as NCEP-ATP (cholesterol); NKDEP (creatinine) and ADA/EDA (HbA1c), (b) the availability of a reliable (inter)national infrastructure to advance complete dissemination of traceability to field labs [e.g. the CDC CRMLN for cholesterol standardization and the former IFCC C-RSE for enzyme standardization] and (c) full understanding and description of the measurand, measurement process and its uncertainty components. However, often the standardization and/or harmonization attempts of tests according to ISO 17511:2003 calibration hierarchies are arduous and unsuccessful, and not realized at a global scale. Obstacles, among others, are fragmented good will approaches and the fact that regulations for market access of IVDs differ per continent or even per region or country, which brings along enormous costs for IVD-manufacturers if tests have to go again through (S)FDA-approval and/or CE-marking because of re-standardization. Consequently, even the availability of reference materials, reference measurement procedures and reference measurement systems for global standardization is not guaranteeing successful adoption and implementation. Persistence of non-equivalence of test results due to lack of traceability and/or incomplete estimation of measurement uncertainty is harmful for patient care as it prevents the use of universal reference intervals and decision limits with unambiguous postanalytical interpretation and it impedes universal application of clinical guidelines.

Medical tests should be fit-for-clinical purpose. To be useful, medical tests should be both traceable to standards of higher order and have limited measurement uncertainty, not exceeding predefined analytical performance specifications according to the Milan consensus hierarchy (1) (2). If traceability and measurement uncertainty are adequately addressed, test results can be interpreted against their matching comparators (either reference intervals or decision limits) and be made comparable in time and space, across the globe. Developing, deploying and implementing a metrological traceability hierarchy is often experienced as a Titan's battle due to the complexity of such a process and the fact that multiple stakeholders are involved which have to be committed to the same end goal. Both the International Consortium for Harmonization of Clinical Lab Results (ICHCLR) and the IFCC Scientific Division Executive Committee give guidance to the standardization process of medical tests by prioritizing analytes that have to be standardized (www.harmonization.net) and by establishing and guiding interdisciplinary working groups and committees who standardize specific analytes (http://www.ifcc.org/ifcc-scientific-division/).



Major organizational entities underpinning the quality of patients' test results.

QA, quality assurance; R&D, research and development; EQA, external quality assessment; QC, quality control; IVD, In vitro diagnostic; ISO, International Organization for Standardization; CLSI, Clinical and Laboratory Standard Institute; IFCC, International Federation for Clinical Chemistry and Laboratory Medicine; AACC, American Association of Clinical Chemistry. Source: Metrological traceability in clinical biochemistry. Graham H. White. Ann Clin Biochem 2011; 48: 393–409 (review). Reprinted with Permission.

Materials and Methods

For monitoring the adequate implementation of the metrological traceability concept by IVD-manufacturers, trueness verification by EQA-providers who have an accuracy-based EQA-scheme that allows to evaluate analytical bias and precision separately is key (3) (4). To evaluate the equivalence of medical test results in Argentina, it was decided to set up a collaborative study between the EQA-organizations of Argentina (PEEC, Dr. Mazziotta) and the Netherlands (SKML, Dr. Cobbaert/ Dr. Weykamp). To that end a comparative EQA-survey was designed and performed using human, commutable, value-assigned trueness verifiers from the Dutch EQA-organizer, the SKML. Sixteen clinical chemistry tests were considered for which JCTLM-endorsed Reference Materials, Reference Measurement Procedures and Reference Services are in place. Moreover, the measurement uncertainty of the measured test results in EQA-samples was checked against allowable total error, and desirable bias and imprecision as derived from biological variation (Table I). The manufacturing of the EQA-samples used was done in the ISO 13485 certified lab of Dr Weykamp (Winterswijk, the Netherlands) and has been described elsewhere (4-6). Shipment of EQA-samples to Argentina was done on dry ice (-70 °C) whereas local shipment within Argentina was controlled but could not be accomplished at -70 °C.

EQA-specimens and lab methods

Six frozen, human-based, commutable poolsera, covering the (patho)physiological measuring range for 16 general chemistry analytes, were assayed by ~75 Argentinian labs (in one run during the summer of 2014) and by ~200 Dutch laboratories at biweekly intervals (during the first quarter of 2014). In the Netherlands, medical laboratories generally use homogenous, regulatory approved applications for the analysis of general chemistry analytes (the so-called CE-IVDs). The use of Lab-Developed-Tests (LDTs) with reagent components from different sources (heterogenous applications) is a general practice in Argentina.

Data-retrieval and inspection

In Argentina, visual inspection of the reported data was done by Dr. Mazziotta. Additional SPSS outlier detection was done in Leiden, using Box & Whisker plots and outliers being defined as test results exceeding 3* interquartile range. In the Netherlands, the data entry was web-based data with embedded outlier procedure, data evaluation and scoring in the SKML scoring system. For the benchmark against the Argentinian labs, the Dutch EQA-data were exported into Excel.

After removal of outliers, harmonization status among laboratories was evaluated by calculating overall

Table I. *Study parameters (N = 16) and desirable analytical performance specifications derived from biological variation.*

Medical test	*Biological variation*		*Desirable specification*		
	CV_w	CV_g	I(%)	B (%)	TE_a (%)
Urate	8.6	17.5	4.3	4.9	12.0
Creatinine	6.0	14.7	3.0	4.0	8.9
Glucose	5.6	7.5	2.8	2.3	7.0
Protein	2.8	4.7	1.4	1.4	3.6
Sodium	0.6	0.7	0.3	0.2	0.7
Chloride	1.2	1.5	0.6	0.5	1.5
Potassium	4.6	5.6	2.3	1.8	5.6
Calcium	2.1	2.5	1.1	0.8	2.6
Magnesium	3.6	6.4	1.8	1.8	4.8
Alkaline phosphatase	6.5	26.1	3.2	6.7	12.0
Creatine kinase	22.8	40.0	11.4	11.5	30.3
Glutamyltransferase	13.4	42.2	6.7	11.1	22.1
Lactate dehydrogenase	8.6	14.7	4.3	4.3	11.4
Alanine aminotransferase	19.4	41.6	9.7	11.5	27.5
Aspartate aminotransferase	12.3	23.1	6.2	6.5	16.7
Amylase	8.7	28.3	4.4	7.4	14.6

CV_w: within-person biological variability, CV_g: interperson biological variability; I: imprecision; B: bias and TE_a: total allowable error.
Source: https://www.westgard.com/biodatabase1.htm (accessed 28 July 2019)

mean interlaboratory coefficients of variation (CVs, %) per analyte and per country for all 6 levels. Even so, standardization status was evaluated after removal of outliers by calculating overall mean recoveries (%) as compared to the assigned target values per analyte per country for all 6 levels. Absolute median biases were compared to (minimal/desirable) biases derived from biological variation criteria. Harmonization and standardization status in the Netherlands anno 2014 was compared to the status in Argentina. Basic statistics and calculations were done in Excel for both Argentinian and Dutch data.

Results

Mean recoveries in Argentinian Laboratories for e.g. serum creatinine, glucose, CK, Ca^{2+} and Na^{+} were 95-119%; 95-104%; 98-102%; 98-102% and 96-100% respectively, whereas min-max recovery ranges were 65-155%; 58-126%; 47-132%; 66-132% and 85-115% (Figure 1). In the Netherlands, absolute mean recoveries were overall 98.9% with a SD of 2.0% (Table II). Median biases in Argentinian laboratories ranged from -2.9 to 18.2%; -3.1 - 2.6%; -3.3 - 0.5%; -1.1 - 3.8% and -4.3 - 0% for serum creatinine, glucose, CK, Ca^{2+} and Na^{+}. In the Netherlands overall mean/median biases were 1.1% (SD=2.0%), and generally in line with desirable biases derived from biological variation (data not shown).

For serum enzymes evaluated using six poolsera, interlaboratory CVs in the Argentinian laboratories ranged between 10 and 22%, as compared to 3-6% in the Netherlands. For serum uric acid, creatinine, glucose and total protein, interlaboratory CVs varied between 4.3 and 13.1% in Argentinian labs, as compared to <3.5% in the Netherlands. For serum electrolytes, interlaboratory CVs ranged between 1.8 and 3.8% for Na^{+}; 2.9-5.8% for Cl^{-}; 3.8-7.5% for K^{+}; 9.4-10.4% for Ca^{2+} and 16.2-22.3% for Mg^{2+} as compared to ≤ 2% (Na^{+}, K^{+}, Cl^{-}, Ca^{2+}) and ≤3% (Mg^{2+}) in the Netherlands (Figure 2) (Table II).

Table II. *Recovery, intralab and interlab CVs (in %) of ~200 Dutch medical laboratories in the first quarter of 2014 using human, commutable and value-assigned EQA-samples from the Dutch EQA-organization (SKML) (3).*

Dutch EQA data 2014.1	*Trueness*			*Precision*				*Analytical performance*		
	LUMC	*Target value*	*Consensus value*	*SD between labs*	*Your SD*	*SD-within lab*		*Recovery (%)*	*Interlab CV (%)*	*Intralab CV (%)*
Sodium, mmol/L	141.9	142.3	141.3	1.7	1.3	1.1		99.3	1.19	0.77
Potassium. mmol/L	5.23	5.23	3.19	0.08	0.06	1.06		99.2	1.53	1.15
Chloride. mmol/L	102.6	104.2	102.7	2.1	0.9	1		98.6	2.02	0.96
Creatinine. µmol/L	148.3	146.8	147.1	4	2.6	2.6		100.2	2.72	1.77
eGFR (F. 55 yr. Caucasian) mL//min/1.73m^2	27.5	28	28	0.9	0.9	1.6		100.0	3.21	2.14
Uric acid. mmol/L	0.371	0.375	0.374	0.013	0.007	0.007		99.7	3.47	1.87
Calcium. mmol/L	2.42	2.46	2.46	0.05	0.05	0.04		100.0	2.03	1.63
Magnesium. mmol/L	1.06	1.1	1.08	0.03	0.02	0.02		98.2	2.73	1.82
Total Protein. g/L	64.7	65.3	63.8	1.5	0.6	0.9		97.7	2.30	1.38
Alkaline phosphatase. U/L	157	164	155	10	5	5		94.5	6.10	3.05
Gamma-GT. U/L	78.5	75.5	79.5	2.5	2.7	1.8		101.3	3.18	2.29
ASAT. U/L	70.1	71.9	72.5	2.8	1.8	1.8		100.8	3.89	2.50
ALAT. U/L	87.4	84	85	3.9	2.2	2.2		101.2	4.64	2.62
LD. U/L	511	517	501	20	14	14		96.9	3.87	2.71
CK. U/L	223	227	225	11	3	5		99.1	4.85	2.20
Amylase. U/L	218	214	209	10	3	4		97.7	4.67	1.87
Glucose. mmol/L	12.6	12.77	12.84	0.38	0.46	0.28		100.5	2.98	2.19
							Overall mean	98.9	3.47	2.02
							Overall SD	1.95	1.56	0.70
							Overall CV	1.97	44.83	34.81

Figure 1. *Average recoveries [(%) ± 1 SD] of 16 routine clinical chemistry tests as compared to the assigned values determined by internationally recognized reference laboratories in six commutable EQA-materials. The data were produced in 2014. About ~ 200 Dutch medical laboratories and ~ 75 Argentinian laboratories participated in this comparative evaluation.*

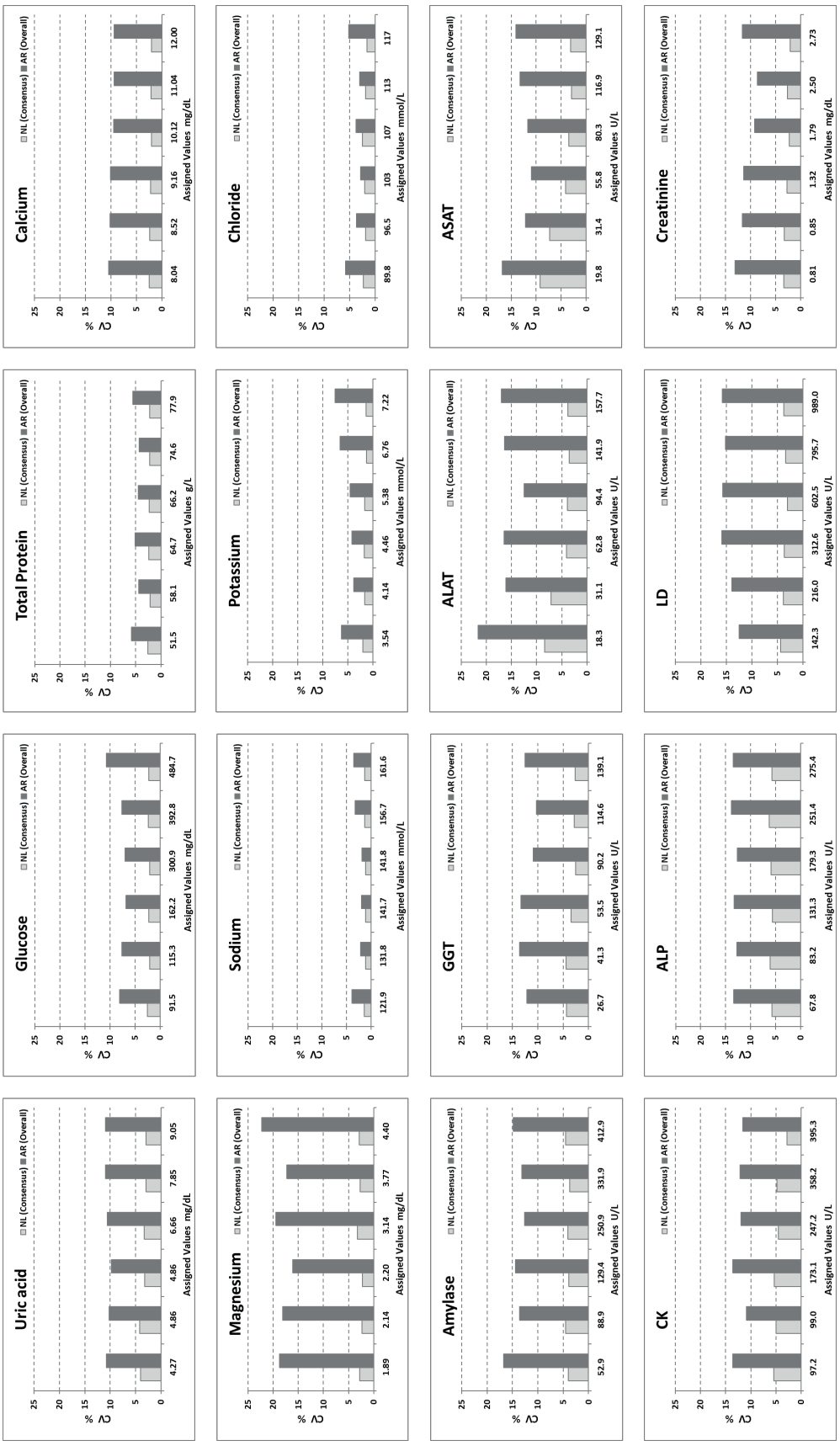


Figure 2. *Interlaboratory variation (CV, %) anno 2014 in ~ 200 Dutch medical laboratories and ~ 75 Argentinian laboratories using six commutable, value-assigned EQA-samples ranked in ascending concentration or activity.*

Discussion and conclusions

Equivalence of results among laboratories is a major mission for medical laboratories. Prof Mazziotta was very well aware of this: for cholesterol standardization his lab was the first Latin-American member of the Cholesterol Reference Method Laboratory Network of the CDC since decades. It is not surprising that this ambassador also initiated a collaboration with the Dutch SKML as the EQA-program in the Netherlands had proven to be a powerful instrument to evaluate test equivalence (4-6).

The EQA-organizations PEEC (*Programa de Evaluación Externa de la Calidad en Argentina*) and SKML (Stichting Kwaliteitsbewaking Medische Laboratoria in the Netherlands) agreed on a joint EQA-survey using commutable, value-assigned EQA-samples to document standardization status and metrological traceability in a sample of Argentinian labs for 16 general chemistry tests. Overall mean recovery was ~100% for substrates, electrolytes and CK in Argentinian labs, whereas for ALAT, ASAT, GGT and amylase there was significant underrecovery (Fig. 1). It cannot be excluded nor confirmed that preanalytical storage conditions confounded to some extent ALAT, ASAT, GGT and amylase test results as it is known by SKML that intermittent storage of EQA-samples at -70 °C is a necessity to keep enzyme activities stable. However, due to limited resources it was not feasible to arrange dry ice sending to the individual Argentinian labs. Overall Argentinian LD and ALP test results were not traceable to the IFCC Reference Measurement System, in contrast with the situation in the Netherlands.

The head-to-head comparison demonstrated that mean interlaboratory variation was 3 to 4-fold higher in Argentinian medical labs as compared to Dutch labs, average interlaboratory CVs varying between 4 and 22% in Argentinian labs with heterogenous test applications (Fig. 2). In case of serum CK ~70% of the labs met the minimum and desirable bias criteria. For substrates ~50% of the labs met the predefined bias criteria, as compared to only 10 to 15% for electrolytes (with the exception of K^+) (data not shown).

Multiple reasons contribute to insufficient implementation of medical test result traceability and large interlaboratory variation (and variable lab practices) in this sample of 75 Argentinian laboratories. Firstly, lack of formal regulations or directives for ensuring IVD test result Quality & Traceability allows deviations from good laboratory practice. Secondly, non-commutable working calibrators might have been used, especially in heterogeneous test applications. The challenge for demonstrating validity and traceability of test results in the case of Lab-Developed-Tests or off label use of commercial tests is the responsibility of the medical lab itself. Also, proof of commutability of working calibrators should have been generated by the individual lab in case of a heterogenous application. Thirdly, lower analytical quality of medical tests also contributed to insufficient traceability of test results. Examples are the limited specificity of creatinine tests based on Jaffé method principles (after all, calibration cannot compensate for non-specificity) and inadequate performance of serum magnesium tests leading to significant underrecovery. Fourthly, lack of specifications and documentation on traceability of test results in product inserts from the main providers of reagents in the Argentinian labs, also does not help the end users to get insight. Finally, individual labs ideally need guidance from EQA-organizers through the reporting of allowable performance limits to evaluate whether the tests used are fit-for-clinical-purpose.

What should be the way forward?

Legal regulation might be helpful in Argentina for ensuring IVD test traceability. Secondly, an accuracy-based EQA system is essential in Argentina to give lab professionals the tools for judging trueness and precision as compared to allowable bias and imprecision, especially when using heterogenous test applications. Key are human, commutable and value-assigned EQA-materials. Thirdly, an ISO-based Quality Management System within labs should give guidance on how to develop LDT-tests that generate traceable test results with limited measurement uncertainty. Fourthly, reference laboratories and reference measurement services are also urgently needed in Latin-American countries to improve test traceability. Finally, knowledgeable lab professionals educated on how to apply the concepts of metrological traceability and commutability are essential.

Conclusions

The collaborators conclude that exchange of human, serum-based, commutable, value-assigned EQA-materials is extremely helpful for monitoring and benchmarking the harmonization and standardization status of medical tests in countries which do not have an accuracy-based EQA-scheme in place yet, and for revealing the future harmonization and standardization potential. The head-to-head comparison on traceability and variability of test results for clinical chemistry tests between Argentina and the Netherlands clearly demonstrates that metrological traceability of clinical chemistry test results in Argentina is on average in line with the expected recovery, however the interlaboratory variation among Argentinian labs is far too high and should be improved. Therefore, the authors of this joint EQA-survey encourage the current and next generations of laboratory specialists in Argentina to continue evaluating the implementation of the metrological traceability concept for medical tests in Argentina, as a tribute to Prof. Mazziotta and his inspiring mind, for the sake of better patient care in Latin-American countries.

Acknowledgements

The authors express their sincere gratitude towards all participating medical laboratories, and especially towards the ~75 Argentinian medical laboratories who enabled this comparative evaluation study on the harmonization/standardization status of general clinical chemistry tests.

Correspondence

Prof. Dr. CHRISTA M. COBBAERT | PhD, Pharm, EuSpLM, Head of the Department of Clinical Chemistry and Laboratory Medicine | Leiden University Medical Center | Department of Clinical Chemistry and Laboratory Medicine, Postbus 9600 | Postzone E2-P | 2300 RC Leiden, the Netherlands | Room E 1-200 | Tel: 31-71-526 4483 (direct) | sein 69184 or tel.: 31-71-5262278 (secretary) | E-mail: C.M.Cobbaert@lumc.nl

References

1. Panteghini M, Ceriotti G, Jones G, Oosterhuis W, Plebani M, Sandberg S, on behalf of the Task Force on Performance Specifications in Laboratory Medicine of the European Federation of Clinical Chemistry and Laboratory Medicine (EFLM). Strategies to define performance specifications in laboratory medicine: 3 years on from the Milan Strategic Conference. Clin Chem Lab Med 2017; 55: 1.849-56.
2. Aarsand AK, Dıaz-Garzon J, Fernandez-Calle P, Guerra E, Locatelli M, Bartlett WA*, et al*. on behalf of the European Federation of Clinical Chemistry and Laboratory Medicine Working Group on Biological Variation. The EuBIVAS: within- and Between-subject biological variation sata for electrolytes, lipids, urea, uric acid, total protein, total bilirubin, direct bilirubin, and glucose. Clin Chem 2018; 64: 1.380–93.
3. Thelen MHM, Jansen RTP, Weykamp CW, Steigstra H, Meijer R, Cobbaert CM. Expressing analytical performance from multi-sample evaluation in laboratory EQA. Clin Chem Lab Med 2017; 55: 1.509-16.
4. Cobbaert C, Weykamp C, Franck P, de Jonge R, Kuypers A, Steigstra H, *et al.* Systematic monitoring of standardization and harmonization status with commutable EQA-samples - five year experience from the Netherlands. Clin Chim Acta 2012; 414: 234-40.
5. Jansen RTP, Cobbaert CM, Weykamp C, Thelen M. The quest for equivalence of test results: the pilgrimage of the Dutch Calibration 2.000 program for metrological traceability. Clin Chem Lab Med 2018; 56: 1.673-84.
6. Ceriotti F, Cobbaert C. Harmonization of external quality assessment schemes and their role - clinical chemistry and beyond. Clin Chem Lab Med 2018; 56: 1.587-90.

Recibido: 1 de agosto de 2019
Aceptado: 5 de noviembre de 2019