Original Article

Inductively Coupled Plasma Mass Spectrometry as a Reference Method to Evaluate Serum Calcium Measurement Bias and the Commutability of Processed Materials during Routine Measurements

Hai-Jian Zhao[1], Meng-Lei Ge[2], Yin Yan[1], Tian-Jiao Zhang[1], Jie Zeng[1], Wei-Yan Zhou[1], Yu-Fei Wang[1,3], Qing-Hui Meng[1,3], Chuan-Bao Zhang[1,3]

[1]National Center for Clinical Laboratories, Beijing Hospital, National Center of Gerontology, Beijing Engineering Research Center of Laboratory Medicine, Beijing 100730, China

[2]Department of Clinical Laboratory, Children's Hospital, Capital Institute of Pediatrics, Beijing 100020, China

[3]Graduate School, Peking Union Medical College, Chinese Academy of Medical Sciences, Beijing 100730, China

Abstract

Background: Measuring total serum calcium is important for the diagnosis of diseases. Currently, results from commercial kits for calcium measurement are variable. Generally, the performance of serum calcium measurements is monitored by external quality assessment (EQA) or proficiency testing schemes. However, the commutability of the EQA samples and calibrators is often unknown, which limits the effectiveness of EQA schemes. The aim of this study was to evaluate the bias of serum calcium measurements and the commutability of processed materials.

Methods: Inductively coupled plasma mass spectrometry was applied as a comparative method, and 14 routine methods were chosen as test methods. Forty-eight serum samples from individual patients and 25 processed materials were quantified. A scatter plot was generated from patient samples, and 95% prediction intervals were calculated to evaluate the commutability of the processed materials and measurement bias at three concentration levels was used to determine the accuracy of routine assays.

Results: All assays showed high precision (total coefficient of variation [CV] <2.26%) and correlation coefficients ($r > 0.99$). For all assays, the mean bias for the 48 patient samples ranged from −0.13 mmol/L to 0.00 mmol/L (−5.61–0.01%), and the ranges for the three concentrations were −0.10–0.04 mmol/L (−5.71–2.35%), −0.14–−0.01 mmol/L (−5.80–−0.30%), and −0.19–0.04 mmol/L (−6.24–1.22%). The EQA samples, calibrators, and animal sera exhibited matrix effects in some assays; human serum pools were commutable in all assays; certificate reference materials were commutable in most assays, and only GBW09152 exhibited a matrix effect in one assay; and aqueous reference materials exhibited matrix effects in most assays.

Conclusions: Biases for most assays were within the acceptable range, although the accuracy of some assays needs improvement. Human serum pools prepared from patient samples were commutable, and the other tested materials exhibited a matrix effect.

Key words: Calibration Bias; Commutability; Inductively Coupled Plasma Mass Spectrometry; Matrix Effect; Serum Calcium

Introduction

The routine methods for measuring total calcium include azo arsenic III colorimetry, O-cresolphthalein complexone (OCPC) colorimetry, and the methylthymol blue colorimetric method. The IFCC recommends isotope dilution mass spectrometry as a decisive method and atomic absorption spectrometry as a reference method.

Compared with reference methods, routine methods are easy to automate, making them suitable for clinical

Access this article online

Quick Response Code:

Website: www.cmj.org

DOI: 10.4103/0366-6999.235109

Address for correspondence: Prof. Chuan-Bao Zhang, National Center for Clinical Laboratories, Beijing Hospital, No. 1 Dahua Road, Dongcheng District, Beijing 100730, China
E-Mail: cbzhang@nccl.org.cn

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

For reprints contact: reprints@medknow.com

© 2018 Chinese Medical Journal ¦ Produced by Wolters Kluwer - Medknow

Received: 02-02-2018 **Edited by:** Li-Shao Guo

How to cite this article: Zhao HJ, Ge ML, Yan Y, Zhang TJ, Zeng J, Zhou WY, Wang YF, Meng QH, Zhang CB. Inductively Coupled Plasma Mass Spectrometry as a Reference Method to Evaluate Serum Calcium Measurement Bias and the Commutability of Processed Materials during Routine Measurements. Chin Med J 2018;131:1584-90.

laboratories. However, routine methods may exhibit weaknesses such as unsatisfactory specificity and a traceability chain[1] that is not intact, which may affect the accuracy of measurement results. Therefore, it is necessary to evaluate the performance of the routine methods commercialized by different manufacturers. Assessment of the variability of the measurement results obtained by different assays and the effectiveness of standardization is best accomplished through external quality assessment (EQA), or proficiency testing (PT), schemes.[2-4] The ideal EQA materials and calibrators need to be commutable or the commutability must be known to ensure that the objective of evaluating the comparability of different assays is fulfilled.[5,6] However, for practical reasons, most PT schemes use processed materials with unknown commutability.

We designed the present study according to the following recommendations made by the Clinical and Laboratory Standards Institute: EP14-A2[7] and EP9-A2.[8] The study involved the measurement of processed materials as well as a set of individual patient samples using comparative and routine methods. By evaluating the commutability of processed materials, we can interpret our PT results and search for possible commutable materials. In addition, the results of the 48 patient serum samples were used to assess the accuracy of routine measures and can provide insight into the potential insufficiency of these methods.

Methods

Ethical approval

The study involved using leftover patient samples, and the leftover patient samples were all de-identified during the collection. It was also ensured that the appropriate amount of serum was collected from each patient sample so that a certain volume was left for possible repetition of measurement. The use of patient samples in the present study has been reviewed and approved by the Ethics Committee of Beijing Hospital, and the need for consent with the donors was waived.

Individual patient samples

De-identified patient samples with concentrations ranging from 1.83 mmol/L to 3.10 mmol/L (measured by the Beckman Access Assay AU5400 at Beijing Hospital) were collected. Each of the samples was split into three aliquots and frozen at −80°C. Two aliquots were detected by the reference method, and one aliquot was detected by the 14 types of routine methods.

Processed materials

EQA materials

The control materials used in our 2013 EQA schemes (purchased from BIO-RAD. Co) and 2014 EQA schemes (purchased from Randox Co.) were designated EQA201331-201335 and EQA201411-201415, respectively. The matrix used was lyophilized human serum. The materials were dissolved, vialed, and frozen at −80°C.

Calibrators

Four types of calibrators, designated Imported A, Imported B, Domestic A, and Domestic B, were all lyophilized serum. The calibrators were dissolved, vialed, and frozen at −80°C.

Human serum pools

Human serum pools were prepared from 200 leftover patient serum samples collected from clinical laboratories at Beijing Hospital. The pools were then mixed and divided into two parts: low-concentration serum was prepared by adding 7% water and high-concentration serum was prepared by adding a standard reference material (NIST SRM 3109a). The target values for the low and high concentrations assigned by inductively coupled plasma-mass spectrometry (ICP-MS) were 2.06 mmol/L and 3.06 mmol/L, respectively. After dispensing, the mixed pools were blended overnight at 4°C and then vialed and frozen at −80°C.

Certified reference materials

Three reference materials were used: GBW 09152 (fresh frozen serum) purchased from the National Institute of Metrology, China, with a concentration of 2.425 ± 0.051mmol/L; BCR304 (lyophilized serum) purchased from the European Commission's Joint Research Center with a concentration of 2.201 ± 0.019 mmol/L; and SRM956C2 (fresh frozen serum) obtained from the National Institute of Standards and Technology with a concentration of 2.538 ± 0.016 mmol/L.

Aqueous reference materials

The standard material SRM3109a (Ca) was spiked into 5% nitric acid to achieve five suitable concentrations and assigned as NC 1–5.

Animal materials

The swine sera samples were obtained as gifts from the Research Center for Coronary Heart Disease, Cardiovascular Institute and Fuwai Hospital, Chinese Academy of Medical Sciences.

Comparative method (reference method)

ICP-MS was used as the comparative method with aluminum selected as an internal standard.

The method was established by the National Center of Clinical Laboratory and was described in the previous paper.[9] All of the samples (48 patient samples and 25 processed materials) were divided into 12 subsets and detected over a period of 12 days.

Test methods (routine methods)

Among the 14 routine methods, three were from foreign companies (Roche, Diasys, and Wako), and the others were from Chinese companies (Sichuan Maccura, Beijing Strong, DIRUI, Leadman, Mindray, Shanghai Fosun Long March, Kehua, Reebio, Biosino, Long March I (OCPC), and Long March II [the azo arsenic III method]). All tests were completed on a Hitachi 7180 automatic biochemical analyzer and the parameters were set in accordance with the manufacturer's instructions. Calibration was performed with

supporting calibrators before the measurements. All of the samples were detected twice, and their order of measurement was reversed in the replicate run.

Statistical analysis

The results from the 48 patient samples were used to generate a scatter plot, with the results of ICP-MS plotted on the X-axis and the results of the 14 routine methods plotted on the Y-axis. The scatter plot was used to calculate the slope, intercept, correlation coefficient, and 95% prediction intervals of each regression line. The commutability of the processed materials was evaluated by comparing the plots (for each of the processed materials) with the limits of the intervals.

The comparison results of the 48 patient serum samples were used to evaluate the measurement accuracy. The results of ICP-MS were used as the target value to calculate the bias of the 14 assays, and the regression line was used to calculate the expected bias at three concentration levels. The following two criteria were used to evaluate bias: the minimum performance standard derived from biological variation (1.3%)[10] and the expanded tolerance limits based on biological variation and the current quality condition (2%). All of the analyses were performed using Microsoft Excel 2010 (Microsoft Corporation, Redmond, WA, USA).

Results

Precision of the test methods

For the routine methods, each sample was measured twice, and the coefficient of variation (CV) was calculated from the two repeated results. With the exception of the DIRUI (1.96%) and Kehua (2.26%) assays, the mean CV of the other assays met the minimum performance criteria derived from biological variation (1.40%),[10] indicating that the precision of most of the assays was satisfactory.

Commutability of processed materials

The commutability of the processed materials was evaluated by comparing the plots with the limits of the intervals. A positive matrix effect means that the bias% was higher than the 95% prediction intervals, and a negative matrix effect means that the bias% was lower than the 95% prediction intervals. Because the concentrations of EQA201334, NC4, and NC5 were higher than those of the 48 patient samples, they were not evaluated.

The 2013 EQA samples exhibited a negative matrix effect in some assays, primarily in domestic systems. The commutability of the 2014 EQA samples was relatively high, and only 2014–2015 exhibited a positive matrix effect in some assays. The human serum pools were commutable in all assays. The analysis of the three certificated reference materials revealed the following: GBW09152 exhibited a matrix effect only in Sichuan Maccura and 956c-2 and BCR304 were commutable in all systems. The analysis of the four calibrators revealed that Imported A exhibited a negative matrix effect in some assays, Imported B exhibited a positive matrix effect in some assays, and the two domestic calibrators exhibited a matrix effect in two imported assays and were commutable in all domestic assays. The aqueous reference materials were noncommutable in most assays and could thus only be used as reference methods. The animal sera showed a matrix effect in some assays. The commutability of the processed materials is shown in Table 1.

Evaluation accuracy of routine methods

Linear regression

With the exception of the Sichuan Marker assays ($r = 0.9887$), the correlation coefficients were all >0.99. The slopes ranged from 0.8949 to 1.1078, and most were close to 1% ± 5%; the range of intercepts was −0.2964–0.2585 mmol/L. The specificity of most assays was sufficient; however, the trueness of some assays needs improvement. The linear equations, correlation coefficients, and residual standard deviation Sy.x are presented in Table 2.

Mean bias and scope

For the 14 routine methods, the mean bias and range of the 48 patient samples are listed in Table 3. For the Roche, Beijing Strong, Mindray, and Kehua assays, the mean bias met the desired bias of 2% and 1.30%. Because the samples with extremely low and high concentrations were difficult to obtain, the concentrations of the 48 fresh serum samples did not cover the three medical decision levels. Therefore, we used the lowest concentration (1.83 mmol/L), highest concentration (3.10 mmol/L), and mean concentration (2.47 mmol/L) to evaluate accuracy. For most assays, the bias at the three concentration levels met the desired bias of CLIA'88; for nearly half of the assays, the bias met the desired limit of 2%. Only a few assays met the minimum performance specification derived from biological variation (1.3%). For the Mindray assay, the bias at the three concentration levels met the desired bias of 1.3%. For the Roche and Beijing Strong assays, the bias at the three medical decision levels all met the desired bias of 2%. The prediction bias for the three medical decision levels is given in Table 4.

Trend of relative bias

For the 14 routine methods, the relative bias exhibited four primary trends. The Roche, Beijing Strong, and Mindray assays exhibited both negative and positive bias. The bias at the three concentration levels was quite small, indicating that the calibration bias can almost be ignored and that the accuracy of the methods was sufficient [Figure 1a, Roche assay]. For the Diasys, Wako, Medical System, Sichuan Maccura, Leadman, Long March II, Reebio, and Biosino assays, negative biases were dominant [Figure 1b, Sichuan Maccura]. For DIRUI and Long March I, the slope was large, indicating that negative bias existed at low concentrations, whereas positive bias existed at high concentrations [Figure 1c, Long March I]. For the Kehua assay, the slope was small, and the intercept was large; therefore, positive bias existed at low levels, and negative bias existed at high levels [Figure 1d, Kehua].

For the DIRUI, Long March I, Kehua, and Biosino assays, the bias changed with the concentration, and the mean bias did

Table 1: Commutability of processed materials

Prepared materials	Roche	Diasys	Wako	Medical system	Sichuan Maccura	Beijing strong	Didui	Leadman	Mindray	Long March I	Long March II	Kehua	Reebio	Biosino
201331	0	2	0	0	0	0	0	0	2	0	0	0	0	0
201332	0	0	0	0	0	0	0	0	0	0	0	2	0	0
201333	0	0	0	0	0	0	0	2	2	0	0	0	0	0
201335	0	0	0	0	0	0	0	0	2	0	0	2	0	0
201411	0	0	0	0	0	0	0	0	0	0	0	0	0	0
201412	0	0	0	0	0	0	0	0	0	0	0	0	0	0
201413	0	0	0	0	0	0	0	0	0	0	0	0	0	0
201414	0	0	0	0	0	0	0	0	0	0	0	0	0	0
201415	1	1	0	0	0	0	0	0	0	1	0	0	0	1
HSP1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
HSP2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Imported A	0	0	2	0	2	0	0	0	0	2	0	0	0	2
Imported B	1	1	0	0	0	1	0	0	0	0	1	0	0	0
Domestic A	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Domestic B	1	1	0	0	0	0	0	0	0	0	0	0	0	0
GBW	0	0	0	0	1	0	0	0	0	0	0	0	0	0
956C-2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BCR304	0	0	0	0	0	0	0	0	0	0	0	0	0	0
NC1	1	2	1	0	0	2	2	2	2	2	2	2	2	0
NC2	1	2	1	0	0	2	2	2	2	2	2	2	2	0
NC3	1	2	1	0	0	2	2	2	2	0	2	2	2	1
Animal	0	2	0	0	2	0	0	2	0	1	2	0	0	0

The commutability of the processed materials was evaluated by comparing the plots with the limits of the intervals. The positive matrix effect means a bias% higher than 95% prediction intervals and the negative matrix effect means a bias% lower than the 95% prediction intervals. Furthermore, "0", "1", and "2" denote commutable, positive matrix effect and negative matrix effect, respectively.

Table 2: Linear regression of the 14 evaluation assays

Assays	Regression lines	*CI* of slope	*CI* of intercept (mmol/L)	*r*	Sy.x
Roche	Y = 0.9666X + 0.0550	0.9489–0.9844	0.0141–0.0960	0.9981	0.0164
Diasys	Y = 0.9487X + 0.0387	0.9373–0.9602	−0.0108–0.0882	0.9971	0.0198
Wako	Y = 1.0002X − 0.0748	0.9848–1.0155	−0.1413–−0.0083	0.9954	0.0267
Medical system	Y = 0.9526X + 0.0613	0.9366–0.9687	−0.0082–0.1307	0.9945	0.0278
Sichuan Maccura	Y = 0.9202X + 0.0538	0.8980–0.9424	−0.0422–0.1498	0.9887	0.0389
Beijing strong	Y = 1.0187X − 0.0614	1.0057–1.0317	−0.1177–−0.0052	0.9968	0.0225
DIRUI	Y = 1.0509X − 0.1673	1.0395–1.0624	−0.2169–−0.1177	0.9977	0.0199
Leadman	Y = 0.9465X + 0.0544	0.9274–−0.9656	−0.0284–0.1371	0.9920	0.0332
Mindray	Y = 0.9630X + 0.0839	−0.9506–0.9754	0.0303–0.1376	0.9967	0.0215
Please Long March I	Y = 1.1078X − 0.2964	1.0939–1.1218	−0.3567–−0.2361	0.9969	0.0242
Long March II	Y = 0.9575X + 0.0459	0.9458–0.9693	−0.0050–0.0967	0.9970	0.0204
Kehua	Y = 0.8949X + 0.2585	0.8674–0.8901	0.2159–0.3141	0.9967	0.0207
Reebio	Y = 0.9032X + 0.1382	0.8896–0.9168	0.073–0.1971	0.9956	0.0236
Biosino	Y = 1.0392X − 0.1762	0.8896–0.9168	0.0793–0.1971	0.9980	0.0183

CI: Confidence interval.

not reflect the actual level of the assays. Therefore, careful attention should be paid to the bias at three concentration levels because the bias changed over the measurement interval in a linear manner, and the difference plot could be analyzed through ordinary linear regression.

Discussion

EQA was introduced into laboratory medicine as an educational tool to address observations that the measurement results for aliquots of the same sample are often different when measured by different laboratories. Although many different and frequently proprietary procedures are used by manufacturers to obtain PT/EQA samples with suitable analyte quantities and stability characteristics for storage and distribution, the commutability is generally unknown. The present study showed that 2013 EQA and 2014 EQA materials exhibited matrix effects in some assays. In EQA schemes, the most common procedure used to assign a target value was to categorize participant methods into "peer groups" that represented a similar technology and

Table 3: The mean bias and slope of the 48 patient samples for the 14 routine methods

Assays	Absolute bias		Relative bias	
	Absolute bias/95% *CI* (mmol/L)	Range of bias (mmol/L)	Relative bias/95% *CI* (%)	Range of bias (%)
Roche	−0.022 (−0.022–−0.021)	−0.05–0.013	−0.90 (−0.94–−0.87)	−2.48–0.74
Diasys	−0.079 (−0.080–−0.078)	−0.16–−0.04	−3.42 (−3.46–−3.39)	−5.23–−1.78
Wako	−0.074 (−0.075–−0.073)	−0.12–0.05	−3.28 (−3.34–−3.22)	−5.64–2.76
Medical system	−0.047 (−0.049–−0.046)	−0.12–0.05	−2.02 (−2.08–−1.97)	−4.94–2.34
Sichuan Maccura	−0.129 (−0.131–−0.128)	−0.16–0.04	−5.61 (−5.69–−5.54)	−8.52–1.65
Beijing strong	−0.019 (−0.020–−0.018)	−0.07–0.02	−0.85 (−0.90–−0.81)	−3.97–1.01
DIRUI	−0.050 (−0.049–−0.051)	−0.09–0.01	−2.30 (−2.35–−2.24)	−5.04–0.26
Leadman	−0.069 (−0.070–−0.067)	−0.11–0.02	−2.97 (−2.02–−2.90)	−6.07–0.84
Mindray	−0.001 (−0.002–0.000)	−0.03–0.05	−0.01 (−0.04–0.05)	−1.83–2.52
Long March I	−0.049 (−0.051–−0.047)	−0.12–0.09	−2.30 (−2.37–−2.23)	−5.81–2.92
Long March II	−0.052 (−0.053–−0.051)	−0.12–−0.01	−2.23 (−2.26–−2.19)	−3.92–−0.43
Kehua	−0.013 (−0.015–−0.012)	−0.14–0.05	−0.43 (−0.49–−0.36)	−4.41–2.65
Reebio	−0.084 (−0.086–−0.083)	−0.14–0.04	−3.58 (−3.63–−3.53)	−5.45–1.65
Biosino	−0.086 (−0.087–−0.085)	−0.12–−0.02	−3.86 (−3.91–−3.81)	−5.62–−0.99

CI: Confidence interval.

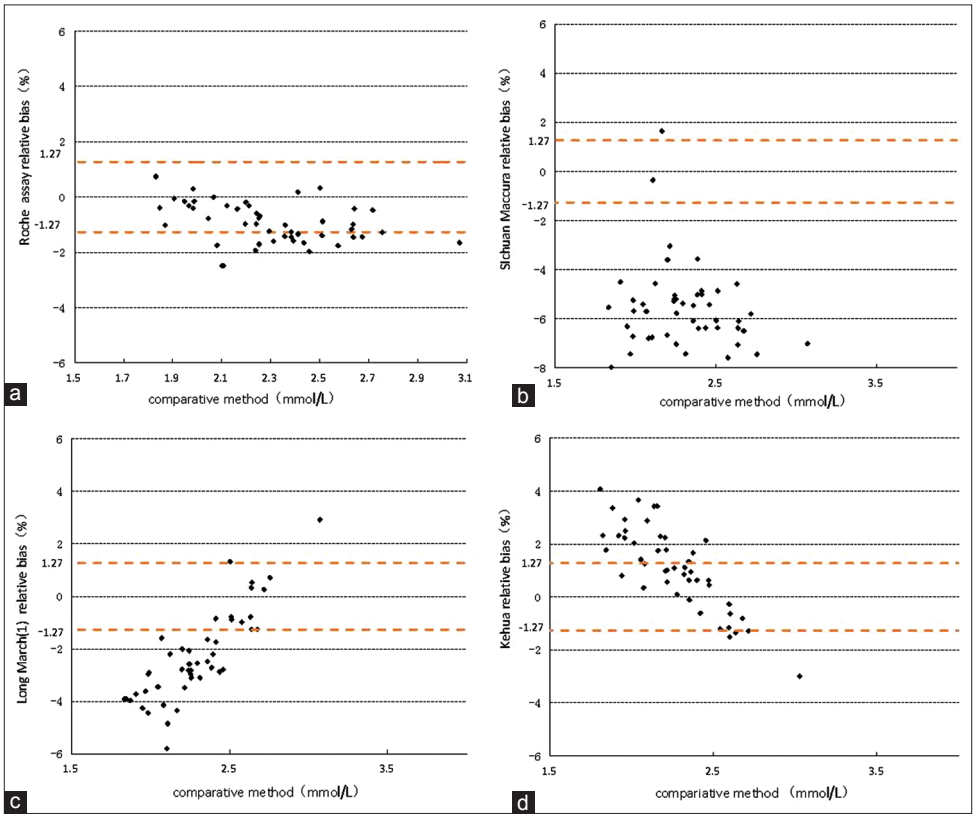

Figure 1: The relative bias of routine methods (a: Roche assay; b: Sichuan Maccura; c: Long March I; d: Kehua) compared with the reference method. The result of the comparative methods is on the X-axis, the relative bias of the routine method is on the Y-axis. The line designated 1.27% (the orange dotted line) was determined by the minimal bias specifications derived from biological variation; the line designated 2% was the expanded tolerance bias derived from biological variation and the current quality condition.

to calculate the mean or median of the peer group as the target value; however, this method may not be able to detect an error by a manufacturer when the supplied calibrators in a region are affected.[11,12] The sample volume required in EQA schemes is large, the preparation of human serum pools is difficult, and their transportation costs are high compared with those of lyophilized serum. Therefore, lyophilized serum is more practical and could not be replaced. Accordingly, these EQA schemes should include at least one human serum pool sample to enable proper interpretation of the EQA results.[13]

Secondary calibrators serve as an important medium for quantifying traceability and that should be traceable to the primary reference material or method and should be

Table 4: Expected bias for three medical decision levels

Assays	1.83 mmol/L		2.47 mmol/L	
	Absolute bias (mmol/L)	Relative bias (%)	Absolute bias (mmol/L)	Relative bias (%)
Roche	−0.01 (−0.02–0.00)	−0.33 (−0.92–0.26)	−0.03 (−0.04–−−0.02)	−1.11 (−1.48–−0.74)
Diasys	−0.06 (−0.07–−0.04)	−3.10 (−3.72–−2.30)	−0.09 (−0.10–−0.08)	−3.56 (−4.01–−3.11)
Wako	−0.07 (−0.09–−0.06)	−4.07 (−5.03–−3.11)	−0.07 (−0.09–−0.06)	−3.01 (−3.62–−2.40)
Medical system	−0.03 (−0.07–−0.04)	−1.39 (−2.39–0.39)	−0.06 (−0.07–−0.04)	−2.26 (−2.89–−1.62)
Sichuan Maccura	−0.09 (−0.12–−0.07)	−5.04 (−6.42–−3.66)	−0.14 (−0.16–−0.12)	−5.80 (−6.68–−4.93)
Beijing strong	−0.03 (−0.04–−0.01)	−1.49 (−2.30–−0.68)	−0.02 (−0.03–0.00)	−0.62 (−1.13–−0.11)
DIRUI	−0.07 (−0.09–−0.06)	−4.05 (−4.76–−3.33)	−0.04 (−0.05–−0.03)	−1.68 (−2.13–−1.23)
Leadman	−0.04 (−0.07–−0.02)	−2.38 (−3.57–−1.19)	−0.08 (−0.10–−0.06)	−3.15 (−3.91–−2.40)
Mindray	0.02 (0.00–0.03)	0.89 (0.12–1.66)	−0.01 (−0.02–0.00)	−0.30 (−0.79–0.19)
Long March I	−0.10 (−0.11–−0.08)	−5.42 (−6.28–−4.55)	−0.03 (−0.04–−0.02)	−1.22 (−1.77–−0.67)
Long March II	−0.03 (−0.05–−0.02)	−1.74 (−2.47–1.01)	−0.06 (−0.07–−0.05)	−2.39 (−2.85–−1.92)
Kehua	0.04 (0.03–0.06)	2.35 (1.65–3.06)	−0.03 (−0.05–−0.02)	−1.40 (−1.85–−0.95)
Reebio	−0.04 (−0.05–−0.02)	−2.13 (−2.98–−1.28)	−0.10 (−0.11–−0.09)	−4.09 (−4.63–−3.55)
Biosino	−0.10 (−0.12–−0.09)	−5.71 (−6.36–−5.05)	−0.08 (−0.09–−0.07)	−3.21 (−3.63–−2.80)
Meet 2%		5		6
Meet 1.3%		2		4

Assays	3.10 mmol/L	
	Absolute bias (mmol/L)	Relative bias (%)
Roche	−0.05 (−0.07–−0.03)	−1.56 (−2.19–−0.93)
Diasys	−0.12 (−0.14–−0.10)	−3.88 (−4.64–−3.11)
Wako	−0.07 (−0.11–−0.04)	−2.39 (−3.42–−1.37)
Medical system	−0.09 (−0.12–−0.05)	−2.76 (−3.84–−1.69)
Sichuan Maccura	−0.19 (−0.24–−0.15)	−6.24 (−7.73–−4.76)
Beijing strong	0.00 (−0.03–0.02)	−0.11 (−0.98–0.76)
DIRUI	−0.01 (−0.03–0.01)	−0.30 (−1.07–0.46)
Leadman	−0.11 (−0.15–−0.07)	−3.60 (−4.88–−2.32)
Mindray	−0.03 (−0.06–0.00)	−0.99 (−1.82–−0.16)
Long March I	0.04 (0.01–0.07)	1.22 (0.29–2.15)
Long March II	−0.09 (−0.11–−0.06)	−2.77 (−3.55–−1.98)
Kehua	−0.11 (−0.13–−0.09)	−3.58 (−4.34–2.82)
Ge	−0.16 (−0.19–−0.13)	−5.23 (−6.14–−4.31)
Biosino	−0.05 (−0.08–−0.03)	−1.76 (−2.47–−1.06)
Meet 2%		6
Meet 1.3%		4

The bias is the difference between the expected value calculated by the regression line and the target value (the result of ICP-MS). The range in brackets is the 95% *CI* of the bias. The two criteria used to evaluate bias were the minimum performance standard based on biological variation (1.3%) and the expanded tolerance limits based on biological variation and current measurement level (2%). ICP-MS: Inductively coupled plasma-mass spectrometry; *CI*: Confidence interval.

commutable with patient samples. The manufacturers generally validate the commutability of their calibrators to yield measurement results with traceability to the highest metrological standard.[5] In China, more than half of the laboratories use the heterogeneous system. If the calibrators lack commutability, using the same value as a constant value may affect the validity of the measurement assay. In this study, the matrices of the 14 supporting calibrators were different. For the Diasys, Medical system, DIRUI, Kehua, and Reebio assays, the matrix of the calibrators was an aqueous solution, whereas for the other assays, the matrix of the calibrators was lyophilized serum. For all 14 supporting calibrators, the matrix was different from fresh serum and the commutability was unknown. Furthermore, four calibrators (Imported A, Imported B, domestic A, and domestic B) commonly used in China showed a matrix effect in some assays. Therefore, laboratories using heterogeneous systems should verify the traceability of the measurement results.[14]

Commutability is relevant to the general materials and methods of a measurement protocol. In this study, with the exception of the aqueous reference materials, all of the processed materials were commutable in the Medical system, DIRUI, and Reebio assays, indicating that the specificity of these assays was high. For the Diasys assay, nine processed materials showed matrix effects, indicating that the specificity of the assay may need improvement. We should note that the standard to evaluate the matrix effect is based on whether a significant statistical difference was found; thus, the premise of the conclusion was that the precision of all of the assays (Sy.x) was consistent.

In the study, we evaluated the calibration bias based on the results of 48 patient serum samples. The correlation coefficient was >0.99 for most assays, and the slope was close to 1. The intercept ranged from −0.2964 mmol/L to 0.1382 mmol/L, and most of the intercept values were close to ±0.1 mmol/L. This finding indicated that the specificity of most assays was sufficient, but some assays still exhibited calibration bias. For the Roche, Beijing Strong, and Mindray assays, the bias at the three concentration levels was within 2%, indicating that the accuracy of these assays was good.

In this study, all of the assays used reagents and their supporting calibrators such that the presence of calibrator bias may originate from the inaccurately assigned value of the calibrators. For most assays, the assigned value of calibrators could be traced to the reference materials such as SRM909b, GBW09152, and CERI JCSS calcium standard solution, and the commutability of these reference materials was unknown. Only the calibrator of the Diasys assay was traced to the reference method, namely, atomic absorption spectroscopy. Using reference materials, lacking in commutability to assign the values of the routine calibrators may affect the accuracy of the method. Therefore, several researchers recommend using the reference method instead of reference materials to assign values of the routine calibrators.[15] For most manufacturers, the detailed information of the procedure to assign the values of routine calibrators is not provided; therefore, the laboratories cannot know if the traceability chain has been maintained. The traceability standards in the *in vitro* diagnostic instruments directive (Directive 98/79/EC) as well as ISO17511 and ISO18153[16] are all relevant to diagnostic equipment manufacturers.

One of the main limitations of this study is that all of the assays were performed using the Hitachi 7180 automatic biochemical analyzer. For each assay, we only used one lot reagent and calibrator, and therefore accidental errors may have occurred when performing these assays.

Because the samples with extremely low and high concentration were difficult to collect (i.e., the concentration range of the 48 fresh serum samples was 1.83–3.10 mmol/L), we can only select the processed materials within the limited concentration range for evaluation. Furthermore, the concentration range did not cover the three medical decision levels; therefore, we used the lowest concentration (1.83 mmol/L), highest concentration (3.10 mmol/L), and the concentration (2.47 mmol/L) of the patient samples to evaluate the accuracy with respect to these three levels.

In conclusion, we determined that some processed materials display the matrix effect, which should be considered during interpretation. Most routine methods have high specificity; however, the trueness of some assays needs improvement. Using the reference method to evaluate the matrix effect and calibration bias may help to promote the standardization of measurements of the serum calcium concentration.

Financial support and sponsorship

This study was supported by grants from the National High Technology Research and Development Program of China (863 Program, No. 2011AA02A102) and The Leaping Development Project of Beijing Biomedical Industry (No. Z161100001816043).

Conflicts of interest

There are no conflicts of interest.

References

1. Myers GL, Miller WG. Challenge to coordinate harmonization activities on an international level. Clin Chem 2017;63:1429-30. doi: 10.1373/clinchem.2017.275669.
2. Myers GL, Miller WG. The roadmap for harmonization: Status of the International Consortium for Harmonization of Clinical Laboratory Results. Clin Chem Lab Med 2018. pii:/j/cclm.ahead-of-print/cclm-2017-0907/cclm-2017-0907.xml. doi: 10.1515/cclm-2017-0907.
3. Badrick T, Punyalack W, Graham P. Commutability and traceability in EQA programs. Clin Biochem 2018. pii: S0009-9120(18)30127-9. doi: 10.1016/j.clinbiochem.2018.04.018.
4. Zeng J, Qi T, Wang S, Zhang T, Zhou W, Zhao H, *et al.* Commutability of control materials for external quality assessment of serum apolipoprotein A-I measurement. Clin Chem Lab Med 2018;56:789-95. doi: 10.1515/cclm-2017-0652.
5. Miller WG, Schimmel H, Rej R, Greenberg N, Ceriotti F, Burns C, *et al.* IFCC working group recommendations for assessing commutability part 1: General experimental design. Clin Chem 2018;64:447-54. doi: 10.1373/clinchem.2017.277525.
6. Young IS. The enduring importance and challenge of commutability. Clin Chem 2018;64:421-3. doi: 10.1373/clinchem.2017.284216.
7. CLSI. Evaluation of Matrix Effects: Approved Guideline. CLSI Document EP14-A2. 2nd ed. Wayne, PA: CLSI; 2005.
8. CLSI. Method Comparative and Bias Estimation Using Patient Samples: Approved Guideline. CLSI Document EP9-A2. 2nd ed. Wayne, PA: CLSI; 2006.
9. Han B, Ge M, Zhao H, Yan Y, Zeng J, Zhang T, *et al.* Determination of serum calcium levels by 42Ca isotope dilution inductively coupled plasma mass spectrometry. Clin Chem Lab Med 2017;56:51-8. doi: 10.1515/cclm-2017-0175.
10. The Minimum Performance Criteria Derived from Biological Variation. Available from: https://www.westgard.com/minimum-requirements.htm. [Last accessed on 2018 Jan 31].
11. Zhang T, Zhang C, Zhao H, Zeng J, Zhang J, Zhou W, *et al.* Determination of serum glucose by isotope dilution liquid chromatography-tandem mass spectrometry: A candidate reference measurement procedure. Anal Bioanal Chem 2016;408:7403-11. doi: 10.1007/s00216-016-9817-0.
12. Meng Q, Zhou W, Zhang C, Zeng J, Zhao H, Zhang T, *et al.* Serum triglyceride measurements: The commutability of reference materials and the accuracy of results. Clin Chem Lab Med 2017;55:1284-90. doi: 10.1515/cclm-2016-0682.
13. Zhang S, Zeng J, Zhang C, Li Y, Zhao H, Cheng F, *et al.* Commutability of possible external quality assessment materials for cardiac troponin measurement. PLoS One 2014;9:e102046. doi: 10.1371/journal.pone.0102046.
14. Ge M, Zhao H, Yan Y, Zhang T, Zeng J, Zhou W, *et al.* Performance of electrolyte measurements assessed by a trueness verification program. Clin Chem Lab Med 2016;54:1319-27. doi: 10.1515/cclm-2015-1110.
15. Perich C, Ricós C, Alvarez V, Biosca C, Boned B, Cava F, *et al.* External quality assurance programs as a tool for verifying standardization of measurement procedures: Pilot Collaboration in Europe. Clin Chim Acta 2014;432:82-9. doi: 10.1016/j.cca.2013.11.005.
16. Canalias F, Camprubí S, Sánchez M, Gella FJ. Measurement of quantities in biological samples. Metrological traceability of values for catalytic concentration of enzymes assigned to calibrators and control materials. Clin Chem Lab Med 2006; 44:333-9. DOI: 10.1515/CCLM.2006.058.

采用电感耦合等离子体色谱法作为参比方法评价血清钙测定的偏倚及参考物质的互通性

摘要

背景：血清钙的测定对相关疾病的诊断有重要作用。当前临床实验室测定血清钙的测试盒种类繁多且结果差异较大。依靠室间质量评价计划（EQA）/能力验证计划（PT）来评价血钙测量准确度的主要问题是EQA 材料的互通性是未知的，明确EQA 材料的互通性对正确解读EQA结果有重要作用。本文的主要目的是评价血清钙常规检测方法的测量偏倚及常用处理样本的互通性。

方法：采用电感耦合等离子体质谱法（ICP-MS）作为参比方法，14种常规方法作为待评价方法，同时测定48个单人份血和25个处理样本。48个单人份血的结果用来做散点图，参比方法为X轴，待评价方法为Y轴，使用最小二乘法（OLR）做回归直线，并绘制其95%置信区间。根据生物学变异导出的最低要求和行业标准评价待测方法的偏倚，根据处理样本的结果是否落在95%置信区间评价其互通性。

结果：所有待评价方法的总精密度都小于2.26%，相关系数大于0.99。对待评价常规方法，48个单人份血的偏倚范围为-0.13 mmol/L ~0.00 mmol/L （-5.61% ~-0.01%），三个医学决定水平处的偏倚是-0.10 ~0.04 mmol/L （-5.71%~2.35%）， -0.14~-0.01 mmol/L（-5.80%~-0.30%），-0.19~0.04 mmol/L （-6.24%~1.22%）。

结论：EQA 样本，校准品，和动物血清在某些方法中表现出基质效应；人血清在所有的方法中都有互通性；候选参考物质在绝大多数方法中都互通性，只有GBW09152在一个方法中显示出基质效应，水溶性参考物质在绝大多数方法中都表现出基质效应。