Variability in baseline laboratory measurements of the Brazilian Longitudinal Study of Adult Health (ELSA-Brasil)

R. Ladwig[1], A. Vigo[1], L.M.G. Fedeli[2], L.E. Chambless[3], I. Bensenor[4], M.I. Schmidt[1], P.G. Vidigal[5], C.D. Castilhos[1] and B.B. Duncan[1]

[1]Programa de Pós Graduação em Epidemiologia, Faculdade de Medicina e Hospital de Clínicas de Porto Alegre, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brasil
[2]Laboratório Clínico, Hospital Universitário, Universidade de São Paulo, São Paulo, SP, Brasil
[3]Department of Biostatistics, University of North Carolina, Chapel Hill, NC, USA
[4]Departamento de Clínica Médica, Faculdade de Medicina, Universidade de São Paulo, São Paulo, SP, Brasil
[5]Departamento de Propedêutica Complementar, Faculdade de Medicina, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brasil

Abstract

Multi-center epidemiological studies must ascertain that their measurements are accurate and reliable. For laboratory measurements, reliability can be assessed through investigation of reproducibility of measurements in the same individual. In this paper, we present results from the quality control analysis of the baseline laboratory measurements from the ELSA-Brasil study. The study enrolled 15,105 civil servants at 6 research centers in 3 regions of Brazil between 2008–2010, with multiple biochemical analytes being measured at a central laboratory. Quality control was ascertained through standard laboratory evaluation of intra- and inter-assay variability and test-retest analysis in a subset of randomly chosen participants. An additional sample of urine or blood was collected from these participants, and these samples were handled in the same manner as the original ones, locally and at the central laboratory. Reliability was assessed with the intraclass correlation coefficient (ICC), estimated through a random effects model. Coefficients of variation (CV) and Bland-Altman plots were additionally used to assess measurement variability. Laboratory intra and inter-assay CVs varied from 0.86% to 7.77%. From test-retest analyses, the ICCs were high for the majority of the analytes. Notably lower ICCs were observed for serum sodium (ICC=0.50; 95%CI=0.31–0.65) and serum potassium (ICC=0.73; 95%CI=0.60–0.83), due to the small biological range of these analytes. The CVs ranged from 1 to 14%. The Bland-Altman plots confirmed these results. The quality control analyses showed that the collection, processing and measurement protocols utilized in the ELSA-Brasil produced reliable biochemical measurements.

Key words: Multi-center studies; Laboratory measurements; Quality control

Introduction

The validity of inferences from clinical and epidemiological studies critically depends on the validity and reliability of measurements in the data collection process. Quality assurance and quality control measures are therefore needed throughout the stages of planning and data collection. For an individual patient in continued care, the use of quality specifications based on the intraperson biological variation of an analyte through a certain period of time is internationally accepted (1). For epidemiological studies, in which greater laboratory precision facilitates the detection of associations between variables, quality can be assessed through the agreement of repeated measurements from biological samples of a single participant visit (2).

Investigators in Brazil and in other low and middle income countries have increasingly been involved in multi-center studies with centralized laboratory analysis, but the necessary procedures for quality control have been infrequently reported. The ELSA-Brasil (Estudo Longitudinal de Saúde do Adulto – Brazilian Longitudinal Study of Adult Health) study utilized a central laboratory for most of its analyses, offering an opportunity to describe the methodology and results of the quality assessment of these measurements.

The resource of having a centralized laboratory in the ELSA-Brasil allowed each research center to have a team for the collection and processing of the biological material, which would be stored in cryotubes at –80°C for up to 30 days, and further transported to the central facility. The negative aspects of this strategy were the need for centralized training and certification of the teams from each center, and the increased risk of delay in the return of results to participants. The most important positive aspects were the facility for acquisition of the lab kits and the absence of inter-laboratory variability. The local processing of the samples decreased the volume of material transported, resulting in a reduction of costs.

The objective of this report is to present the results of quality control analyses of baseline laboratory measurements of the ELSA-Brasil.

Material and Methods

Data collection

The ELSA-Brasil study enrolled 15,105 participants in 6 field centers located in 3 different regions of the country from 2008 to 2010 (3). The study protocol was approved by ethics committees at each institution, and all participants gave their written informed consent. Participants underwent interviews, examinations and collection of blood and urine specimens, in approximately 6 h at the local research clinics.

As has been previously described in more detail (4), blood collection was performed in a fasting state and then, among those without a diagnosis of diabetes, 2 h after the ingestion of a 75-g oral glucose solution.

Blood was centrifuged within 30 min of collection, with aliquots then being separated in an ice bath into cryotubes previously labeled with bar codes, and stored in freezers at –80°C until transportation to the central laboratory on dry ice. Once samples were received, analysis for glucose (fasting and 2 h), glycated hemoglobin (HbA1c), creatinine, sodium, potassium, uric acid, aspartate transaminase (AST), alanine transaminase (ALT), gamma-glutamyl transferase, total cholesterol, HDL-cholesterol, LDL-cholesterol, triglycerides, thyroid stimulating hormone (TSH), insulin (fasting and 2 h) and ultrasensitive C-reactive protein, as well as serology for Chagas disease were performed.

Urine was collected over 12 h during the night, prior to the clinic visit, locally processed and stored for shipment in aliquots as described above. Determinations were performed centrally for sodium, potassium, calcium, creatinine and albumin.

Given the large number of participants, multi-center nature and diversity of measurements, the ELSA-Brasil required effective and efficient mechanisms of quality assurance and control for laboratory determinations. As previously reported (5), the main quality assurance activities were careful selection of research instruments,

centralized training and certification, pretesting and pilot studies, and preparation of procedure manuals.

For laboratory measurements, inter-assay CVs were calculated using the results of internal controls, and intra-assay CVs using data from several pilot studies on both fresh and frozen samples.

In addition, we performed test-retest analyses of specific analytes in 10% of a randomly selected sample of study participants, with determination of intra- and inter-assay variability at the central laboratory. In selected participants, a single additional blood or urine sample was collected for duplication. Extra blood samples were collected at the end of the blood drawing process. These tubes ("blind replicates"), identified only by their bar code and thus, blind for determination, were handled identically to the original tubes and sent to the central laboratory at the same time as the original tubes.

Laboratory results were transmitted to the study's data center, where the original and quality control (QC) results were matched through the bar codes to individual study participants.

Statistical analysis

Reliability was estimated through the calculation of intraclass correlation coefficient (ICC) (6), using a random effects model. A mixed model that also included a fixed order effect (QC sample minus original sample) was used for the four analytes for which the fixed effect was statistically significant at the 0.05 level. Confidence intervals were estimated by the 2.5 and 97.5 percentiles of the empirical distribution obtained with bootstrap sampling with 1000 repetitions.

We calculated the CV as the ratio of the standard deviation of measurement error to the mean of the analyzed variable.

$$CV = \frac{\sigma_{error}}{\mu}$$

Bland-Altman plots were also produced to graphically explore the agreement between measurements (7).

Analyses were performed initially on all pairs. To minimize the effect of extreme outliers, we then reanalyzed the differences after removing pairs presenting a difference between measurements >3 standard deviations from the mean or for whom the pair's mean was >5 standard deviations from the overall sample mean. Analyses were performed using SAS software (USA), version 9.3 (8).

Results

The inter- and intra-assay CVs for the analytes are reported in Table 1. This table also presents external, analytical CVs (CVa) specifications based on biological variation, for comparison. CVs obtained were generally quite small. Intra-assay CVs varied from 0.86% for HbA1C to 3.97% for urinary calcium; inter-assay CVs varied from

Table 1. Intra- and inter-assay coefficients of variation (CV). Baseline exam of the Longitudinal Study of Adult Health (ELSA-Brasil), 2008–2010.

	Intra-assay CV (%)	Inter-assay CV (%)		CV+ (%)
		Lower control	Higher control	
Glucose (mmol/L)	1.24	2.70	2.09	2.80
Glycated hemoglobin (HbA1C - proportion)	0.86	3.01	5.01	1.40
Creatinine - serum (μmol/L)	3.02	6.93	4.52	2.98
Sodium - serum (mmol/L)	#	1.44	1.28	0.50
Potassium - serum (mmol/L)	#	3.08	2.23	3.50
Aspartate transaminase (TGO/AST) (U/L)	3.17	4.68	2.25	6.15
Alanine transaminase (TGP/ALT) (U/L)	3.17	5.39	2.36	9.70
Gamma-glutamyl transferase (U/L)	2.29	3.81	2.71	6.70
Uric acid (mmol/L)	1.33	2.48	2.53	4.30
Total cholesterol (mmol/L)	1.75	1.76	1.94	2.98
Triglycerides (mmol/L)	1.60	2.02	2.56	9.95
HDL-C (mmol/L)	2.00	2.34	2.25	3.65
LDL-C (mmol/L)	#	3.28	2.57	3.90
Calcium - urine (mmol/L)	3.97	4.23	3.75	13.10
Creatinine - urine (mmol/L)	2.06	2.15	2.43	5.50
Sodium - urine (mmol/L)	2.60	1.67	1.71	14.35
Potassium - urine (mmol/L)	2.48	2.62	2.30	12.20
Microalbuminuria (g/L)	1.25	4.86	5.09	17.50
TSH (mIU/L)	3.32	4.65	5.89	9.65
Insulin (pmol/L)	3.64	7.13	7.77	10.55
C-reactive protein (nmol/L)	2.68	4.76	4.31	21.10

+Coefficient of variation based on biological variation components. Ref: SEQC. Comité de Garantia de la Calidad y Acreditación de Laboratorios. Comisión de Calidad Analítica. 2014. Available at http://www.seqc.es/es/Comisiones/18/9/102/Base_de_datos_de_Variacion_biologica_%7C_Comision_de_Calidad_Analitica_%7C_Comite_Cientifico/. #Data not available. HDL-C: high density lipoprotein cholesterol; LDL-C: low density lipoprotein cholesterol; TSH: thyroid-stimulating hormone.

1.28% for sodium to 7.77% for insulin. Most CVs were lower than the CVa, with serum creatinine, serum sodium, and glycated hemoglobin being notable exceptions.

The number of blind replicate QC samples obtained varied from 72 to 94 across analytes. Fewer QC samples were available for analytes obtained 2 h after the glucose load, as participants with diabetes were excluded from this test.

Table 2 presents QC results for all blind replicate pairs for each analyte, and Table 3 shows results for pairs after the removal of outliers. The number of pairs removed varied from 0 to 4. All analytes had an ICC above 0.93 except for serum sodium (0.50) and potassium (0.73) (Table 2). The removal of pairs with outliers (Table 3) generally produced a small improvement in ICCs values, except for serum sodium, which increased from 0.50 to 0.61.

Figure 1 presents Bland-Altman plots of the pair differences against average pair values for selected analytes. Panel A shows data for glycated hemoglobin and Panel B for fasting insulin, the latter demonstrating a small systematic difference between initial analyte and the QC sample (–3.79; 95%CI: –5.83 to –1.76). Other similarly small differences (Supplementary Figures) were seen between pairs for serum creatinine (2.34; 95%CI=1.19–3.49), serum potassium (0.16; 95%CI=0.12–0.20), and TSH (0.06; 95%CI=0.01–0.11). Although both glycated hemoglobin and insulin had outliers, their removal had only a small effect on CV and ICC values.

In order to visualize the pairs and the influence of outliers, Figure 2 shows results for alanine transaminase, prior to (Panel A) and after exclusion of outliers (Panel B). As can be seen in Panel A, a single extreme outlier, probably due to liver disease, makes QC analysis difficult. In fact, as can be seen when comparing results for this analyte in Tables 2 and 3, after outlier exclusions the CV decreased from 12.5 to 6.33%. The ICCs, prior to and after this exclusion were high with laboratory variability accounting for less than 2% of the overall variance.

Figure 3 presents the Bland-Altman plot for serum sodium, demonstrating the narrow biological range of this analyte.

Table 2. Quality control analyses of all pairs of analytes. Baseline data of the Longitudinal Study of Adult Health (ELSA-Brasil), 2008-2010.

Description	N	Mean	s^2_{total}	s^2_{error}	CV (%)	ICC	95%CI
Glucose (mmol/L)	93	6.22	2.27	0.0345	3.0	0.99	0.95–1.00
Glycated hemoglobin (HbA1C - proportion)	85	0.053	0.00004	0.000002	2.9	0.94	0.86–0.97
Creatinine - serum (μmol/L)	88	82.2	266	15.6	5.3	0.93	0.89–0.95
Sodium - serum (mmol/L)	88	143	8.68	4.37	1.5	0.50	0.31–0.65
Potassium - serum (mmol/L)	88	4.35	0.13	0.034	4.3	0.73	0.60–0.83
Aspartate transaminase (TGO/AST) (U/L)	88	29.9	2487.3	2.84	5.6	1.00	0.91–1.00
Alanine transaminase (TGP/ALT) (U/L)	88	34.9	5465	18.95	12.5	1.00	0.98–1.00
Gamma-glutamyl transferase (U/L)	88	42.2	2139	2.93	4.1	1.00	1.00–1.00
Uric acid (mmol/L)	88	0.325	0.009	0.00003	1.7	1.00	1.00–1.00
Total cholesterol (mmol/L)	88	5.36	1.21	0.0205	2.7	0.98	0.97–0.99
Triglycerides (mmol/L)	88	1.44	0.502	0.0018	2.9	1.00	0.99–1.00
HDL-C (mmol/L)	88	1.48	0.114	0.0038	4.2	0.97	0.95–0.98
LDL-C (mmol/L)	88	3.21	0.993	0.0087	2.9	0.99	0.98–1.00
Glucose 2 h post 75 g glucose load (mmol/L)	72	7.71	10.7	0.0946	4.0	0.99	0.97–1.00
Calcium - urine (mmol/L)	79	2.65	4.28	0.0081	3.4	1.00	1.00–1.00
Creatinine - urine (mmol/L)	80	9.02	28.6	0.0456	2.4	1.00	1.00–1.00
Sodium - urine (mmol/L)	80	108	2460	2.16	1.4	1.00	1.00–1.00
Potassium - urine (mmol/L)	80	34.3	406	0.62	2.3	1.00	1.00–1.00
Microalbuminuria (g/L)	94	0.01	0.0005	0.000001	10.6	1.00	0.98–1.00
TSH (mIU/L)	75	1.98	5.1	0.026	8.1	1.00	0.99–1.00
Insulin (pmol/L)	75	52.4	1220	47.3	13.1	0.96	0.94–0.98
Insulin - 2 h post 75 g glucose load (pmol/L)	74	397	96500	3190	14.2	0.97	0.91–0.99
C-reactive protein (nmol/dL)	83	26.8	1680	3.27	6.8	1.00	1.00–1.00

N: number of observed pairs; s^2_{total}: total variance; s^2_{error}: error variance; CV: coefficient of variation; ICC: intraclass correlation coefficient; 95%CI: 95% confidence interval; HDL-C: high density lipoprotein cholesterol; LDL-C: low density lipoprotein cholesterol; TSH: thyroid-stimulating hormone.

Discussion

This study demonstrates the excellent quality of the sample collection and processing, and laboratory measurement of biochemical analytes of the ELSA-Brasil study, in a sample of free-living Brazilian adults. Laboratory intra- and inter-assay variability was almost always below recommended international standards. Between-person variability accounted for more than 90% of the total variability with few exceptions, and was usually greater than 98%. For epidemiological studies involving associations between exposures and diseases, the most inclusive and relevant of the QC measures is the ICC. The ICC, also called the reliability coefficient, can be interpreted as a ratio of the between-individual variance to the total variance. The existence of noise (underlying within-person variability) as a percentage of the signal (true value of exposure, in the case the laboratory values) lowers the ICC values. Therefore, when testing epidemiological associations, the greater the noise the more difficult it is to detect real associations at a level of statistical significance. One important, and partially controllable source of this variability is the intraindividual variation, sometimes called "measurement error". The overall reliability coefficient evaluates the fraction of the population variability of a given variable that are not due to measurement errors (intra-individual variability). This measurement error could originate from biological variation or from either pre-analytical or analytical sources. Here, with samples taken from a single participant visit, the biological variation was not evaluated. Pre-analytical errors include those related to variability in collection (e.g., hemolysis), processing (e.g., delay in chilling and centrifugation leading to metabolism of glucose in the collection tube), storage (e.g., inadequate temperature maintenance) and transportation (e.g., premature thawing). Also to be considered is the possible mixing of tubes among participants during the collection processes. In this study, none of the outliers were clustered to a particular individual participant, indicating that the possibility of serious errors resulting from switching or erroneously labeling collection tubes or aliquots, or assigning laboratory results to the wrong participant, seem unlikely.

Analytical errors include laboratory determination errors (e.g., reagent variability between kits). The intra- and inter-assay CVs available from the ELSA central

Table 3. Quality control analyses of analytes after removal of outliers. Baseline data of the Longitudinal Study of Adult Health (ELSA-Brasil), 2008-2010.

Description	N	Mean	s^2_{total}	s^2_{error}	CV(%)	ICC	95%CI
Glucose (mmol/L)	91	6.22	2.29	0.0123	1.78	1.00	0.98–1.00
Glycated hemoglobin (HbA1C - proportion)	81	0.053	0.00004	0.000001	2.05	0.97	0.93–0.99
Creatinine - serum (μmol/L)	87	82.2	250	15.6	5.09	0.93	0.90–0.96
Sodium - serum (mmol/L)	86	143	8.2	3.2	1.25	0.61	0.51–0.70
Potassium - serum (mmol/L)	87	4.34	0.13	0.031	4.06	0.75	0.62–0.84
Aspartate transaminase (TGO/AST) (U/L)	87	29.9	96.3	2.86	5.66	0.97	0.90–0.99
Alanine transaminase (TGP/ALT) (U/L)	86	27.5	169	3.03	6.33	0.98	0.97–0.99
Gamma-glutamyl transferase (U/L)	87	41.1	2133	2.59	3.92	1.00	1.00–1.00
Uric acid (mmol/L)	87	0.325	0.009	0.00003	1.60	1.00	1.00–1.00
Total cholesterol (mmol/L)	88	5.36	1.21	0.0205	2.68	0.98	0.97–0.99
Triglycerides (mmol/L)	86	1.41	0.491	0.0013	2.57	1.00	1.00–1.00
HDL-C (mmol/L)	87	1.47	0.115	0.0033	3.92	0.97	0.96–0.98
LDL-C (mmol/L)	87	3.21	0.712	0.0088	2.91	0.99	0.98–0.99
Glucose 2 h post 75 g glucose load (mmol/L)	70	7.71	7.04	0.0444	2.73	0.99	0.99–1.00
Calcium - urine (mmol/L)	78	2.63	4.28	0.0075	3.27	1.00	1.00–1.00
Creatinine - urine (mmol/L)	79	8.84	28.9	0.0392	2.23	1.00	1.00–1.00
Sodium - urine (mmol/L)	78	106	2510	1.38	1.10	1.00	1.00–1.00
Potassium - urine (mmol/L)	78	32.6	396	0.38	1.89	1.00	1.00–1.00
Microalbuminuria (g/L)	90	0.0067	0.00008	0.0000002	7.29	1.00	0.99–1.00
TSH (mIU/L)	74	1.7	1.26	0.01	7.13	0.99	0.98–0.99
Insulin (pmol/L)	74	52.3	1230	40.5	12.2	0.97	0.95–0.98
Insulin - 2 h post 75 g glucose load (pmol/L)	71	365	61000	1090	9.04	0.98	0.96–0.99
C-reactive protein (nmol/dL)	81	20	489	1.81	5.88	1.00	1.00–1.00

N: number of observed pairs; s^2_{total}: total variance; s^2_{error}: error variance; CV: coefficient of variation; ICC: intraclass correlation coefficient; 95%CI: 95% confidence interval; HDL-C: high density lipoprotein cholesterol; LDL-C: low density lipoprotein cholesterol; TSH: thyroid-stimulating hormone.
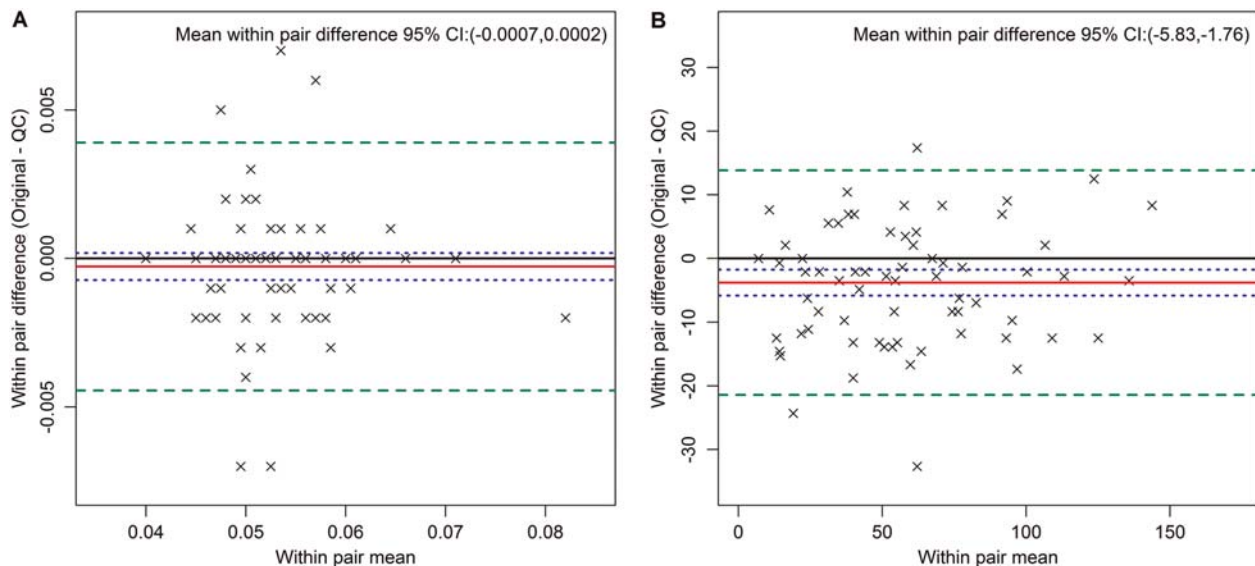

Figure 1. Bland Altman plots for glycated hemoglobin. Proportion of total hemoglobin (all pairs, *Panel A*) and insulin in pmol/L (all pairs, *Panel B*). The red line indicates the mean within-pair difference found between the original and the quality control (QC) measures. The green dashed lines indicate 2 standard deviations above and below the mean, and the blue dotted line the 95%CI for the mean of the within-pair difference.
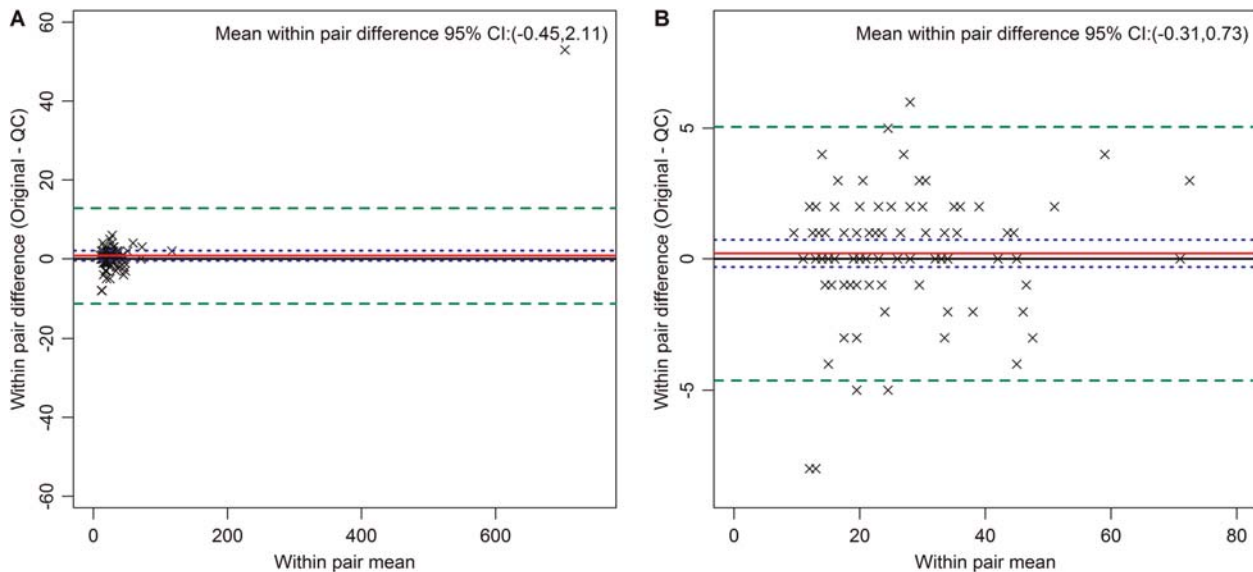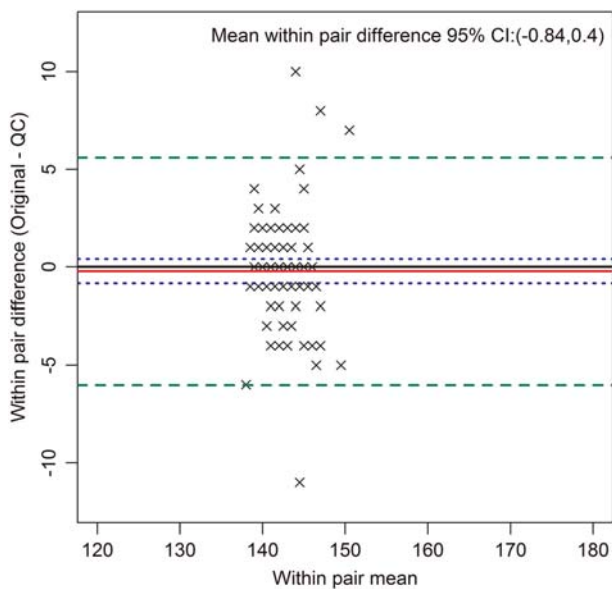
Figure 2. Bland Altman plots for alanine transaminase in U/L (all pairs, *Panel A*; after exclusion of outliers, *Panel B*). The red line indicates the mean within pair difference found between the original and the quality control (QC) measures. The green dashed lines indicate 2 standard deviations above and below this mean, and the blue dotted line the 95%CI for the mean of the within pair difference.

Figure 3. Bland Altman plot of serum sodium in mmol/L (all pairs), an example of an analyte with a narrow biological range. The red line indicates the mean within pair difference found between the original and the quality control (QC) measures, the green dashed lines indicate 2 standard deviations above and below this mean, and the blue dotted line the 95%CI for the mean of the within pair difference.

laboratory indicate very small error in this part of the measurement process, in general. With few exceptions, VCs of measured analytes were lower than the analytical

quality specifications based on components of biological variation (9). Of note, some analytes, such as insulin, are typically measured with greater variability.

Thus, the bulk of the variability in reported measurements seems likely to be originated from a combination of smaller problems in analyte collection, processing and transportation, coupled with the specific analyte susceptibility to each of these problems.

The markedly lower ICCs of the analytes serum sodium and potassium highlight an interesting phenomenon. As the CVs for these analytes were quite high, the low ICCs can be attributed to their low between-person variability, as illustrated by the narrow distribution of the values in the Bland-Altman plots for serum sodium (Figure 3) and potassium (Supplementary Figures). The ranges of these analytes, vital to body function, are tightly controlled within narrow limits. Despite this, their ICCs indicated that their relatively high measurement error would make it more difficult to find statistically significant associations involving either of them in the ELSA sample.

The CV, popular in reports of measurement error, is of utility in determining the extent to which the analyte value of a given participant will vary, in relation to the analyte's mean value, from one measurement to the next – a relevant issue in the determination of the analyte's clinical utility. However, as CV evaluation doesn't take other variabilities into consideration, it has a lesser role in evaluating the quality of measurements in epidemiologic studies of associations. In this study, the blind replicate CVs were higher than the intra- and inter-assay CVs, which was caused by the inclusion of not only laboratory

measurement variability, but also pre-analytical sources of error usually present in multi-center studies with centralized measurements.

The Bland-Altman plots provide a visual analysis of these QC results. One can see potentially systematic differences among replicates, such as that found for fasting insulin (Figure 2). This difference, when present, was always small, and most likely due either to the fact that the QC sample was always drawn last, or to chance, given the multiple comparisons performed. The plots show no evidence of increased or decreased variability as a function of the analyte average value, which might be present, for example, if inadvertent thawing had led to loss of analytes. The $+/-$ 2 standard deviations boundaries were relatively close to the means, once again indicating a small measurement error.

The decision to report results with and without outliers is not always easy. For example, it seems logical to use the ICC and CV obtained for alanine transaminase after exclusion of the outlier seen in the Bland-Altman plot, given that the mean of the outlier pair was more than 30 times the mean of the remaining pairs, leading to a potentially large weighting of this single value. The outlier value was apparently obtained from a participant with liver disease and, given the large difference within the pair ($>$50 U/L), likely outside the range of precision of the measurement technique. For precision evaluation in an epidemiological study in which the overwhelming majority of participants have values within or near the normal range, inclusion of such outliers makes interpretation of the findings difficult. As can be seen when comparing the Bland-Altman plots before and after the outlier exclusion, it obscures the analysis of measurement error of values within the relevant range. In contrast, outliers in other analytes, both for average values and those for differences between the QC sample and its pair, were close to the exclusion boundaries. As such, these values should probably be considered measurements within the relevant range of study, and their differences should be included in our final estimates of measurement variability.

Two limitations of our analysis merit mention. First, as QC samples were collected, processed, and shipped simultaneously with their corresponding pair samples, some of the measurement variability occurring during these steps could have been underestimated. Additionally, an important aspect of variability – temporal biological variability – was not assessed, as participants were not recalled to repeat measurements on another day. Thus, for a thorough evaluation of the measurement reliability for epidemiological studies, consideration of this within-participant day-to-day variability would require incorporation of data from other studies.

In conclusion, these analyses of measurement variability are important for QC documentation in multi-center studies during which strict QC measures are necessary to guarantee accurate results. The biochemical analytes of the ELSA-Brasil here reported were measured with high reliability. Based on this, they should serve well as exposure variables and co-variables for most studies of associations, especially given the large sample size of the cohort.

Supplementary material

Click here to view [pdf].

Acknowledgments

References

1. Kenny D, Fraser CG, Hyltoft Petersen P, Kallner A. Consensus agreement. *Scan J Clin Lab Invest* 1999; 59: 585–586, doi: 10.1080/00365519950185409.

2. Szklo M, Nieto F. *Epidemiology, beyond the basics.* 3rd edn. Boston: Jones and Barllett Publishers; 2012.

3. Aquino EM, Barreto SM, Bensenor IM, Carvalho MS, Chor D, Duncan BB, et al. Brazilian Longitudinal Study of Adult Health (ELSA-Brasil): objectives and design. *Am J Epidemiol* 2012; 175: 315–324, doi: 10.1093/aje/kwr294.

4. Fedeli LG, Vidigal PG, Leite CM, Castilhos CD, Pimentel RA, Maniero VC, et al. [Logistics of collection and transportation of biological samples and the organization of the central laboratory in the ELSA-Brasil]. *Rev Saúde Pública* 2013; 47 (Suppl 2): 63–71, doi: 10.1590/S0034-8910.2013047 003807.

5. Schmidt MI, Griep RH, Passos VM, Luft VC, Goulart AC, Menezes GM, et al. [Strategies and development of quality assurance and control in the ELSA-Brasil]. *Rev Saúde Pública* 2013; 47 (Suppl 2): 105–112, doi: 10.1590/S0034-8910.2013047003889.

6. Lachin JM. The role of measurement reliability in clinical trials. *Clin Trials* 2004; 1: 553–566, doi: 10.1191/17407 74504cn057oa.

7. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; 1: 307–310, doi: 10.1016/S0140-6736(86)90837-8.

8. Statistical Analysis System (SAS). [Computer program]. Version 9.3. Cary: SAS Institute Inc.; 2011.

9. Fraser CG, Stevenson HP, Kennedy IMG. Biological variation data are necessary prerequisites for objective autoverification of clinical laboratory data. *Accred Qual Assur* 2002; 7: 455–460, doi: 10.1007/s00769-002-0526-3.