J Bras Patol Med Lab. 2019; 55(4): 348-359.

10.5935/1676-2444.20190032

Longitudinal analysis of laboratory tests in a population-based study

Análise longitudinal da qualidade dos exames laboratoriais em um estudo de base populacional

Anna Cecília Roncalio; Caroline Galgowski; Mayra Alice Corrêa; Leonardo André Lange; Celina Noriko Yamanaka; Tatiani Karini Rensi; Caio Mauricio M. de Cordova

Universidade de Blumenau, Blumenau, Santa Catarina, Brazil.

ABSTRACT

Clinical laboratories must have quality management systems that guarantee the reliability of their results. Furthermore, in longitudinal studies, it is important that the generated data be analyzed over time, as an additional measure of quality control (QC), seeking to identify fluctuations not explainable by biological variables. Applying this strategy to the SHIP-Brazil study was the aim of this work. Thus, we analyzed the results of fasting glucose, post-load glucose and glycated hemoglobin in the participants of the SHIP-Brazil study, from July 2014 to November 2016, in relation to the following aspects: difference of each individual's results from the mean of the month over time, mean and median of results over time, and percentage of results above the reference values over time, with trend line. According to the observed data, in order to guarantee the correct association of laboratory parameters with different health and disease conditions, the fasting blood glucose and post-load glucose measurements in the samples collected during the first months of the study are consistent, despite being performed in two different facilities. This QC strategy has proved very useful, and may even be used by clinical laboratories in their routine, observing result fluctuations of their population in the course of time, being capable of eventually detect bias even before the routine practices of internal and external quality control.

Key words: quality control; cohort study; population-based study; laboratory tests; diabetes *mellitus*; HbA1C.

RESUMO

Os laboratórios clínicos devem ter sistemas de gestão da qualidade que garantam a fidedignidade dos resultados. Ainda, em estudos longitudinais, é importante que os dados gerados sejam analisados ao longo do tempo, como medida adicional de controle de qualidade (CQ), buscando identificar flutuações não explicáveis por variáveis biológicas. Aplicar essa estratégia ao estudo SHIP-Brasil foi o objetivo deste trabalho. Analisamos os resultados de glicemia em jejum, glicemia pós-sobrecarga e hemoglobina glicada nos participantes do estudo SHIP-Brasil, no período de julho de 2014 a novembro de 2016, em relação aos seguintes aspectos: diferença dos resultados de cada indivíduo em relação à média do mês ao longo do tempo; média e mediana dos valores ao longo do tempo; e porcentagem de resultados acima dos valores de referência ao longo do tempo, com linha de tendência. De acordo com os dados observados, as dosagens de glicemia em jejum e glicemia pós-sobrecarga, nas amostras coletadas nos primeiros meses de estudo, apresentam-se consistentes, mesmo tendo sido realizadas em dois laboratórios distintos. Essa estratégia de CQ mostrou-se bastante útil e pode, inclusive, ser utilizada pelos laboratórios clínicos em sua rotina, observando as flutuações de resultados de sua população ao longo do tempo, detectando, eventualmente, desvios antes mesmo das práticas rotineiras de controle interno e externo da qualidade.

Unitermos: *controle de qualidade; estudo de coorte; estudo populacional; exames laboratoriais; diabetes* mellitus*; HbA1C.*

First submission on 08/12/18; last submission on 08/20/18; accepted for publication on 03/19/19; published on 08/20/19

RESUMEN

Los laboratorios clínicos deben tener sistemas de gestión de calidad que garanticen la fiabilidad de los resultados. Además, en estudios longitudinales, es importante que los datos generados sean analizados a lo largo del tempo, como medida adicional de control de calidad (QC), buscando identificar fluctuaciones no explicables por variables biológicas. Aplicar esa estrategia al estudio SHIP-Brasil es el objetivo de este trabajo. Analizamos los resultados de glucemia en ayunas, prueba de sobrecarga de glucosa y hemoglobina glucada en los participantes del estudio SHIP-Brasil, en el período de julio de 2014 a noviembre de 2016, acerca de los siguientes aspectos: diferencia de los resultados de cada individuo con respecto a la media del mes con el tiempo; media y mediana de los valores con el tiempo; y porcentaje de resultados por encima de los valores de referencia con el tempo, con línea de tendencia. Según los datos observados, las mediciones de glucemia en ayunas y tras una sobrecarga de glucosa, en las muestras recolectadas en los primeros meses de estudio, resultan consistentes, a pesar de haber sido realizadas en dos laboratorios distintos. Esa estrategia de QC se presentó muy útil y puede, incluso, ser utilizada por laboratorios clínicos en su rutina, observando fluctuaciones de resultados de su población a lo largo del tiempo, detectando, eventualmente, desviaciones aún antes de las prácticas rutinarias de control interno y externo de calidad.

Palabras clave: *control de calidad; estudio de cohorte; estudio poblacional; pruebas de laboratorio; diabetes* mellitus*; HbA1C.*

INTRODUCTION

The Study of Health in Pomerode (SHIP-Brazil), conducted by the University of Blumenau, Santa Catarina, Brazil, aims at stratifying the genetic and risk factors existing (or not) in the genesis of the most varied diseases. To this end, laboratory, anthropometric, genetic and image exams are necessary to ascertain these factors and monitor their evolution.

The study in Pomerode is multicentric, having as a model the SHIP Project conducted in the region of German Pomerania[1]. That study has been going on for 15 years at the University of Greifswald, and the cooperation of researchers from there will also allow the use of data obtained in Germany in the current project.

Pomerode is a city founded by Pomeranian immigrants that keep German tradition and culture alive. From the health point of view, it presents some particularities, such as high suicidal and skin cancer rates, besides high mortality indices by cardiovascular diseases, among others, combined with social indicators considered of developed countries. The SHIP-Brazil study seeks to investigate whether different environmental, social conditions, as well as those of access to sanitation, health and education, bring differences in the conditions of life and health of genetically similar populations, but subjected to a different life environment. The studied population amounts to 3,091 individuals, that is, 11.13% of the 27,759 inhabitants of Pomerode[2]. However, to answer the epidemiological and analytical questions, it is necessary that data generated by the studies be absolutely reliable, and tools of quality control (QC) can detect analytical or pre-analytical variations as time passes, from the most varied origins, regardless of the quality management systems of the laboratories themselves. Applying this strategy to the SHIP-Brazil study was the objective of this work.

This article shows the analysis of fasting glucose, post-load glucose, and glycated hemoglobin (HbAlC) results, as examples, in the participants of the SHIP-Brazil study, in the period from July 2014 to November 2016, concerning the following aspects: difference of each subject's means over time, mean and median of values over time, and percentage of results above reference values over time, with trend line.

MATERIAL AND METHODS

Sample population

The population of SHIP-Brazil is a simple random sample composed of seven strata of age and sex beginning from 20 years of age, with intervals of 10 years; the population was considered by sex for each age group, with prevalence of 50%, precision of 5% [confidence interval (CI) 95%], 20% for losses and 20% for analysis of confounding, totaling 3,091 subjects estimated of the population of 27,759 inhabitants of Pomerode[2]. Individuals older than 20 years that have lived in the municipality for at least six months were included in the study; those with any physical or mental limitation that prevents them to answer the questionnaires or perform the tests, or refuse to sign the free informed consent, were excluded. A lengthy general questionnaire was applied, with socioeconomic, demographic, lifestyle, medical assistance and

mental variables. Ultrasonography examinations of heart and carotids, complete abdomen, and calcaneus were carried out. Individuals' participation in the SHIP-Brazil study was registered and approved by the Ethics Research Committee of the University under no. 33/2012. In the current study of quality control, results from 1,081 individuals were evaluated, although a portion of the participants did not undergo all the tests, once they are optional at the interview moment, especially the post-load glucose. The sample collection for the SHIP-Brazil study and the participants' exams began in July 2014.

Biological samples

The complete SHIP-Brazil study encompasses samples of whole blood, serum, mononuclear cells of peripheral blood, plasma, urine, saliva, nasal swab, oropharyngeal swab, tongue scrap, gingival pocket samples (when detected) and stool[1]. At the current project, we analyzed the results obtained from samples of serum and whole blood. We adopted a procedure to orientate participants as to preparation for exams, as well as for collection, transportation, processing, and storage of samples at the study biorepository. All those procedures, as well as the training of exam and collection staff, were validated by the responsible staff of the SHIP study in Germany.

Laboratory measurements

HbAlC measurements were taken by affinity chromatography at an automated system D10 (Biorad, Hercules, California, USA); and the glucose measurements, by the Hexokinase II method (Siemens AG, Erlangen, Germany), at an automated instrument Mindray BS-120 (Laboratory A) or Siemens ADVIA 1650 (Laboratory B). The exams were carried out at a small university laboratory (Laboratory A), from July 2014 to September 2015; and at a private hospital medium-sized laboratory (laboratory B), from December 2015 to November 2016. A comparability study between both laboratories was not performed due to financing issues and management difficulties of the study. However, both laboratories have established quality management systems with internal controls and participation in external programs of proficiency – Laboratory B is about to be certificated and acts within a hospital certified by Organização Nacional de Acreditação (ONA).

Statistical analysis

The obtained data were treated by descriptive statistics and analysis of variance (ANOVA), aided by the Microsoft Excel 2016 (Microsoft) and ezANOVA (Copyright© 2007, by Chris Rorden). The following aspects were evaluated: monthly mean and median in relation to moving mean and median of all results over time; percentage of results above reference values (RV) each month, with trend curves.

RESULTS AND DISCUSSION

Observing the results of fasting glucose in relation to the mean (**Figure**), significant differences are not perceptible over time (**Table**). However, from May 2015 and during the three following months, a tendency of decrease can be noted of mean and median of results, as well as the percentage of values above RV, what is also observed in the results of the oral glucose tolerance test (OGTT) and HbAlC, revealing a profile of individuals more tolerant to glucose in this period of time.

In relation to OGTT results, we observed higher values of mean and median in the first month of the study followed by higher HbA1C values. An average increase of HbA1C values is also perceptible in the month of March 2015, similarly accompanied by increased mean and median of fasting glucose.

Generally speaking, we observed a slight decrease of all parameters over time, in a gradual way. This occurred due to a strategy of the recruitment staff, which initially invited the participants of more advanced age, so that, once they got to know the study, and, as expected, approved the exam routine, could encourage other inhabitants of Pomerode, facilitating adherence to recruitment.

The Figure demonstrates the analysis of results of fasting glucose exams, post-load glucose, HbA1C and OGTT, presenting the result of each individual in relation to general mean, mean, and median in each month, and percentage of results above the RV.

The HbA1C and glucose are measured based on laboratory screening tests and the diagnosis of diabetes *mellitus*. Because of this, results must be precise and reliable, supported by population studies aimed at determining conditions of life, health, and sickness associated with the different pathological manifestations. Those laboratory tests, however, can suffer pre-analytical interference, which comprise especially chronobiologic variation, gender, age, position at the collection moment, physical activity, fasting, diet, use of medications and drugs of abuse, tourniquet application and collection procedure, prior diagnostic and/or therapeutic procedures, drug infusion, use of separating gel tube, hemolysis, and lipemia[3-6].

Clinical laboratories develop and implement quality management systems, with routines of internal control and

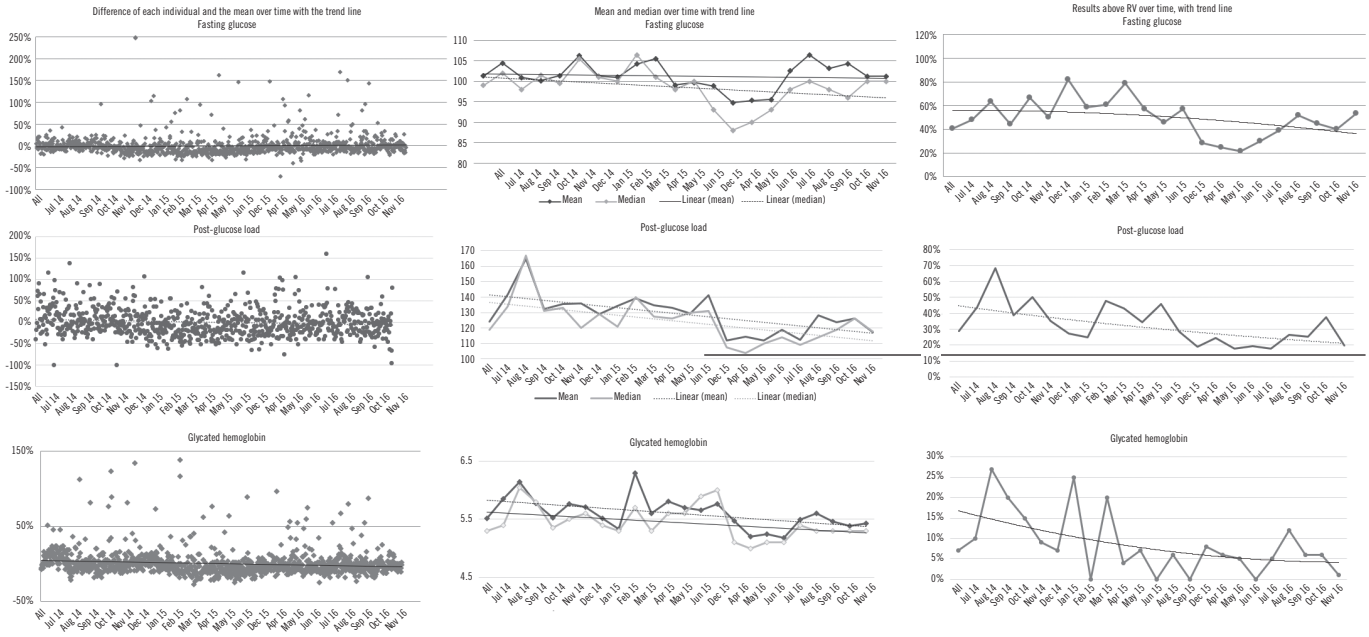

FIGURE – *Graphical representation of results in analysis of fasting blood glucose (*n *= 1,081), post-glucose load (*n *= 806), and HbA1C (*n *= 1,141), with mean and median in each month, besides the difference of value of each individual compared with the general mean and percentage of results above the RV*

RV: reference value.

TABLE – Descriptive statistics of mean, median and values above RV over time for fasting glucose, post-glucose load, and HbA1C in the first months of the study

FG	All months	Jul 14	Aug 14	Sep 14	Oct 14	Nov 14	Dec 14	Jan 15	Feb 15	Mar 15	Apr 15	May 15	Jun 15	Aug 15	Sep 15	Dec 15
n	1.081	25	19	18	12	20	22	12	23	14	35	37	7	-	-	156
Mean	100.6	101.3	104.5	100.9	100.2	101.4	106.3	101.3	101	104.3	105.5	99.1	99.7	-	-	98.8
Median	97	99	102	98	101.5	99.5	105.5	101	100	106.5	101	98	100	-	-	93
p	-	0.592	0.1794	0.805	0.3962	0.5027	0.0613	0.323	0.4589	0.0888	0.2625	0.9065	0.6198	-	-	0.485
SD	23.5	10.5	12.8	9.3	5.3	12	7.8	9.4	6.3	9.1	18.4	9.6	5.2	-	-	34.1
CV%	23%	10%	12%	9%	5%	12%	7%	9%	6%	9%	17%	10%	5%	-	-	35%
% > RV	41%	48%	63%	44%	67%	50%	82%	58%	61%	79%	57%	46%	57%	-	-	28%
p	-	> 1	> 1	> 1	> 1	> 1	> 1	> 1	> 1	> 1	> 1	> 1	> 1	-	-	> 1
OGTT	**All months**	**Jul 14**	**Aug 14**	**Sep 14**	**Oct 14**	**Nov 14**	**Dec 14**	**Jan 15**	**Feb 15**	**Mar 15**	**Apr 15**	**May 15**	**Jun 15**	**Aug 15**	**Sep 15**	**Dec 15**
n	806	25	19	18	12	20	22	12	23	14	35	37	7	-	-	116
Mean	124	141.6	164.9	132.3	135.6	136.1	129.1	134.5	139.2	134.6	133.3	129.5	141.1	-	-	112
Median	119	134	167	131	133	120	129	121	140	127.5	126	130	131	-	-	107.5
p	-	0.0698	0.0037**	0.0591	0.0452*	0.5184	0.0944	0.4768	0.019*	0.1589	0.2068	0.0816	0.1229	-	-	0.0731
SD	39.5	46.2	43.7	38.1	32.2	50	29.2	40	36.3	23.2	31.2	39.1	50.2	-	-	34.3
CV%	32%	33%	27%	29%	24%	37%	23%	30%	26%	17%	23%	30%	36%	-	-	31%
% > RV	29%	44%	68%	39%	50%	35%	27%	25%	48%	43%	34%	46%	29%	-	-	19%
p	-	0.1498	0.1481	0.1477	0.146	0.1484	0.149	0.146	0.1492	0.1466	0.153	0.1535	0.1445	-	-	0.1871
HbA1C	**All months**	**Jul 14**	**Aug 14**	**Sep 14**	**Oct 14**	**Nov 14**	**Dec 14**	**Jan 15**	**Feb 15**	**Mar 15**	**Apr 15**	**May 15**	**Jun 15**	**Aug 15**	**Sep 15**	**Dec 15**
n	1.141	39	22	35	20	23	29	8	30	25	47	56	11	69	3	51
Mean	5.52	5.85	6.15	5.78	5.53	5.77	5.71	5.52	5.33	6.3	5.6	5.81	5.7	5.66	5.77	5.47
Median	5.3	5.4	6.05	5.8	5.35	5.5	5.6	5.4	5.3	5.7	5.3	5.6	5.6	5.9	6	5.1
p	-	0.4811	0.0293*	0.0751	0.8524	0.3255	0.1846	0.7284	0.4823	0.2061	0.8497	0.1919	0.1854	0.1509	0.0963	0.6201
SD	0.92	0.6	0.6	0.77	0.7	1.37	0.93	0.76	0.53	1.82	0.83	1.08	0.27	0.69	0.21	1.52
CV%	17%	10%	10%	13%	13%	24%	16%	14%	10%	29%	15%	19%	5%	12%	3%	28%
% > RV	7%	10%	27%	20%	15%	9%	7%	25%	0%	20%	4%	7%	0%	6%	0%	8%
p	-	0.3759	0.3699	0.3744	0.3692	0.3703	0.3724	0.3651	0.3728	0.3709	0.3788	0.3821	0.3662	0.387	0.3635	0.3802

FG	Apr 16	May 16	Jun 16	Jul 16	Aug 16	Sep 16	Oct 16	Nov 16
n	65	141	57	59	91	100	95	73
Mean	94.8	95.3	95.5	102.5	106.5	103.2	104.3	101.2
Median	88	90	93	98	100	98	96	100
p	0.1943	0.1949	0.1735	0.6648	0.3537	0.6266	0.7935	0.4429
SD	24.3	24.7	11.9	27.8	24.9	25.9	24.3	11.6
CV%	26%	26%	12%	27%	23%	25%	23%	11%
% > RV	25%	21%	30%	39%	52%	45%	40%	53%
p	> 1	> 1	> 1	> 1	> 1	> 1	> 1	> 1
OGTT	**Apr 16**	**May 16**	**Jun 16**	**Jul 16**	**Aug 16**	**Sep 16**	**Oct 16**	**Nov 16**
n	41	51	41	45	64	79	69	56
Mean	1,145	111.8	118.8	112.2	128.3	123.9	126.3	117.2
Median	104	110	114	109	114	119	126	117.5
p	0.1697	0.0575	0.279	0.0668	0.9673	0.9899	0.2045	0.2394
SD	39.1	35	42.3	30.5	49.1	38.4	32.4	36.3
CV%	34%	31%	36%	27%	38%	31%	26%	31%
% > RV	24%	18%	20%	18%	27%	25%	38%	20%
p	0.1555	0.1591	0.1554	0.1569	0.1632	0.1691	0.1643	0.1608
HbA1C	**Apr 16**	**May 16**	**Jun 16**	**Jul 16**	**Aug 16**	**Sep 16**	**Oct 16**	**Nov 16**
n	64	135	57	59	91	101	95	71
Mean	5.2	5.25	5.18	5.49	5.6	5.46	5.39	5.43
Median	5	5.1	5.1	5.4	5.3	5.3	5.3	5.3
p	0.1724	0.2196	0.1475	0.7961	0.8497	0.8459	6,393	0.7583
SD	0.85	0.86	0.358	0.883	0.932	0.83	0.841	0.395
CV%	16%	16%	7%	16%	17%	15%	16%	7%
% > RV	6%	5%	0%	5%	12%	6%	6%	1%
p	0.3851	0.4135	0.3826	0.3653	0.3680	0.3663	0.3663	0.3646

The values of p *(ANOVA) are indicated for eventual differences of values of the month in analysis compared with cumulative data of all results.*
RV: reference values; FG: fasting glucose; OGTT: oral glucose tolerance test; SD: standard deviation; CV%: coefficient of variation; ANOVA: analysis of variance; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

participation in external programs of proficiency, in order to guarantee quality of their results. However, these routines are sometimes unable to detect variations over time, whether they are due to variability, bias of reagent lots and/or analytical systems, or even, as can occur in longitudinal studies, pre-analytical variations, such as staff turnover, changes in procedures and collection materials, etc. Thus, other strategies of data analysis are necessary for QC in cohort studies, such as evaluation over time of the difference of result of each individual in relation to the mean results of a specific period of time (month to month, for example); mean and median of results over time; percentage of results above or below, when this is the case, of limits of RV established for each test. Each of these parameters must be analyzed individually and, if only one of them is out of the global trend, results must be considered suspect of bias and preferably repeated to ensure their reliability[7].

In our study, we could observe that the biochemical tests evaluated over time did not present a marked variability.

In the last months of the study, because of operational difficulties with the initially hired laboratory, laboratory tests were conducted at an outsourced laboratory. Even so, results presented consistent, and variation of creatinine values is explained by the more advanced age of patients in the first months. Somehow these findings are explained by the recent standardization of calibrators and biochemical assays, traceable to reference materials quantified by methods considered gold standard for each analyte, fostered by the International Federation of Chemical Chemistry and Laboratory Medicine (IFCC), what did not occur in the other specialties of the clinical laboratory. We also see that quality management systems established by both laboratories have apparently contributed to the consistency of their results, what is essential for longitudinal studies. In their final phase, the laboratory results of the SHIP-Brazil study will only be released for association analysis after control of all the exams carried out and the eventual repetition of inconsistent results.

Other similar studies also expressed concern with the quality of their laboratory data, without leaving them just in charge of internal and external routine strategies of laboratory control. The Brazilian Longitudinal Study of Adult Health (ELSA-Brazil), for example, searched for a strategy to analyze intra- and inter-assay variability, as well as the retest of a random sampling, which proved effective in result assurance[8]. The Indonesia Family Life Survey and the Longitudinal Aging Study in India used an external control strategy similar to those offered by proficiency

providers; this proved adequate to their necessities[9]. The strategy we used in this work was similar to that used in the original SHIP study in Germany[10]. Regardless of the used method, for which there is still no consensus, it is essential that authors of studies continuously perform these analyses to detect eventual deviations. A classical study detected, for example, that total cholesterol presented significant changes during three years, even with the laboratory keeping its certification[11].

REFERENCES

1. Voelzke H, Alte D, Schmidt CO, et al. Cohort profile: The study of health in Pomerania. Int J Epidemiol. 2011; 40: 294-307.

2. IBGE – Instituto Brasileiro de Geografia e Estatística. Censo 2010. Available at: https://censo2010.ibge.gov.br/. [accessed on: Nov 12, 2018].

3. Clinical and Laboratory Standards Institute (CLSI). Procedures for the collection of diagnostic blood specimens by venipuncture; approved standard. 6th ed. CLSI document H3-A6. Wayne, PA: CLSI; 2007.

4. Lima-Oliveira G, Lippi G, Salvagno GL, Montagnana M, Picheth G, Guidi GC. Impact of the phlebotomy training based on CLSI/NCCLS H03-A6 – procedures for the collection of diagnostic blood specimens by venipuncture. Biochem Med (Zagreb). 2012; 22(3): 342-51.

5. Lima-Oliveira G, Lippi G, Salvagno GL, Montagnana M, Picheth G, Guidi GC. The effective reduction of tourniquet application time after minor modification of the CLSI H03-A6 blood collection procedure. Biochem Med (Zagreb). 2013; 23(3): 308-15.

6. Lippi G, Plebani M, Di Somma S, Cervellin G. Hemolyzed specimens: a major challenge for emergency departments and clinical laboratories. Crit Rev Clin Lab Sci. 2011; 48(3): 143-53.

7. Algeciras-Schimnich A, Bruns DE, Boyd JC, Bryant SC, La Fortune KA, Grebe SKG. Failure of current laboratory protocols to detect lot-to-lot reagent differences: findings and possible solutions. Clin Chem. 2013; 59: 1187-94.

8. Ladwig R, Vigo A, Fedeli LM, et al. Variability in baseline laboratory measurements of the Brazilian Longitudinal Study of Adult Health (ELSA-Brasil). Braz J Med Biol Res. 2016; 49(9): e5381. doi: 10.1590/1414-431X20165381.

9. Hu P, Herningtyas EH, Kale V, et al. External quality control for dried blood spot-based C-reactive protein assay: experience from the Indonesia family life survey and the longitudinal aging study in India. Biodemography Soc Biol. 2015; 61(1): 111-20. doi: 10.1080/19485565.2014.1001886.

10. Lüdemann J, Piek M, Wood WG, Meyer S, Greiner B, John U, Hense HW. Methoden zur Qualitätssicherung im medizinischen Untersuchungsbereich epidemiologischer Feldstudien: Die Study of Health in Pomerania (SHIP). Gesundheitswesen. 2000; 62(4): 234-43.

11. Whitney CW, Lind BK, Wahl PW. Quality assurance and quality control in longitudinal studies. Epidemiological Rev. 1998; 20(1): 71-80.

CORRESPONDING AUTHOR

Caio Mauricio Mendes de Cordova ID 0000-0001-6090-0367
e-mail:cmcordova@furb.br