Validity of informed birth weight. Study of Cardiovascular Risk in Adolescents (ERICA) - Rio de Janeiro

Bruno Mendes Tavares [1]
Carlos Henrique Klein [2]
Katia Vergetti Bloch [3]

[1] Instituto de Ciências Biológicas. Universidade Federal do Amazonas. Avenida General Rodrigo Octávio, 6200. Coroado I. Manaus, AM, Brasil. CEP: 69.080-900. E-mail: brutav@gmail.com
[2] Escola Nacional de Saúde Pública Sérgio Arouca. Fundação Oswaldo Cruz. Rio de Janeiro, RJ, Brasil.
[3] Instituto de Estudos em Saúde Coletiva. Universidade Federal Rio de Janeiro. Rio de Janeiro, RJ, Brasil.

Abstract

Objectives: to verify the agreement between birth weight information referred by the guardians of the adolescents who participated in the Study of Cardiovascular Risks in Adolescents (Portuguese acronym ERICA) and the birth weight data of the adolescents identified in the National Information System on Live Births (Portuguese acronym Sinasc).

Methods: probabilistic record linkage of 1,668 records was conducted between the ERICA databases and the Sinasc databases from 1996 to 2002, both from the state of Rio de Janeiro. The agreement between the informed birth weight and the one registered in Sinasc was estimated by the Intra-class Correlation Coefficient (ICC), Bland-Altman plot, Cohen's Kappa index, and Gwet's agreement coefficient.

Results: the ICC was = 0.89; CI95% =0.88-0.90 and the higher the mother's educational level was, the higher it became. There was also an elevated agreement between the birth weight classification in the low (< 2,500 g), adequate (2,500 to 3,999 g) and elevated (≥ 4,000 g) birth weight categories, Gwet's Agreement Coefficient = 0.91; CI95%= 0.89-0.92.

Conclusions: the results showed satisfying agreement between the birth weight referred by the parent/guardian of the adolescents, and the ones registered in SINASC, this agreement being directly proportional to the mother's educational level.

Key words *Birth weight, Adolescent, Epidemiological studies*

http://dx.doi.org/10.1590/1806-93042017000400006

Rev. Bras. Saúde Matern. Infant., Recife, 17 (4): 705-715 out-dez., 2017   **705**

Introduction

The rise in the prevalence of non-communicable chronic diseases has stimulated the interest in studies with auto-referred information on risk factors, whether by adults, adolescents or parents/guardians of children and adolescents, due to cost reduction and logistic complexity.

The birth weight is an important predictor on morbidity and child mortality, due to its association with growth impairment, vulnerability to infectious diseases, developmental delay and mortality.[1] Furthermore, birth weight has been associated with risk factors for cardiovascular diseases, such as over-weight and obesity, in several studies.[2-4] In epidemiological studies, the information on birth weight depends on the memory of the interviewee, while in clinical trials there may be problems due to the lack of communication between maternities and pediatric services.[5]

In recent years, the number of researches in the health department that use the method of record linkage has been rising. The record linkage is a process of identifying and joining records from different databases that belong to the same entity.[6] A new database is created for that, with the variables from the linked databases, improving the capacity to answer the research's questions by aggregating several pieces of information through common data.[7,8]

Considering the importance of assessing the precision of the birth weight information obtained years after birth, the present study had the objective of verifying the agreement between the birth weight information referred by the parent/guardian of the adolescents participating in the Study of Cardiovascular Risks in Teenagers (Portuguese acronym ERICA), in the state of Rio de Janeiro, and the birth weight information of the adolescents registered in the National Information System on Live Births (Portuguese acronym Sinasc). Furthermore, it was investigated if socio-economic and demographic characteristics had influenced the degree of agreement between birth weight referred by a parent/guardian and the one registered in Sinasc.

Methods

The linkage was conducted through the probabilistic record linkage method between the ERICA database, the Parent's/Guardian's Questionnaire, filled by parent/guardian of the adolescent participants born in the state of Rio de Janeiro, and the annual databases from the National Information System on Live Births (Portuguese acronym Sinasc) from 1996 to 2002, also from the state of Rio de Janeiro, obtained from the State Department of Vital Health Data of Rio de Janeiro.

ERICA was a school-based, national, multi-center, sectional study that assessed adolescents from 12 to 17 years old enrolled in public and private schools in Brazilian cities with over 100,000 inhabitants. Further details on the protocol of the study can be found in Bloch *et al.*[9] The data collection from ERICA took place in the period of March 2013 to December 2014 and this study was approved by the ethics committee of research in the Federal University of Rio de Janeiro.

The probabilistic linkage was conducted using the *software OpenRecLink*,[6] a system of database linkage that uses a probabilistic technique following the Fellegi and Sunter model[10] to combine observed pairs in two databases where there are no unequivocal key points, these being of common identification. The ERICA and Sinasc databases went through standardization steps, blocking and record pairing, following the steps and parameters recommended by Camargo Jr. and Coeli.[11] Initially, characters that did not compose the mother's name, such as accents, numbers and unusual characters that could interfere in the linkage process were removed from both databases.

Seventeen blocking steps were conducted, applying different keys through the combination of phonetic codes: Soundex from the first name, Soundex from the last name, gender, day, month and year of birth, and area code from the city they were born, ranging from more strict and less strict key combinations. For the record pairing, variations of the mother's name and birth date were used.

The scores ranged from -6,877 to 10,693 (inferior and superior limits), and, applying predefined rules, all the records with score values above zero were manually reviewed at the end of each blocking step, for the attribution of the pair's status (true or false). At the end of the process, out of 2,671 adolescents from ERICA who went through the linkage, 1,822 (over 68.2%) were identified in Sinasc.

The nutritional state of the adolescents was classified by anthropometric indices BMI ($BMI=weight/height^2$), expressed in Z-score values, using a reference from the WHO.[12] The cut-off points used were: very low weight Z-score < -3; low weight Z-score ≥ -3 and < -2; eutrophication Z-score ≥ -2 and ≤ 1; overweight Z-score > 1 and ≤ 2; obesity Z-score > 2. This information was used to compare subgroups of adolescents with and without the Guardian's Questionnaire and linked and non-

linked.

Three different classification methods were used for the socioeconomic classification. An economic classification criterion (socioeconomic criteria according the Brazilian Association of Research Companies – Portuguese acronym ABEP) that combines possession of assets with the education level of the householder,[13] grouping the five original classes in three, A=A1 and A2, B=B1 and B2, and C= C, D and E; the mother's educational level, classified in 4 categories: Incomplete primary school (including illiterates also), complete primary school, complete secondary school and complete tertiary school and the administrative nature of the adolescent's school, public or private.

For birth weight classification, were considered the weights registered in Sinasc to be in the category of low birth weight (Portuguese acronym BPN) those lower than 2,500 g, adequate birth weight (Portuguese acronym APN) those of higher or equal vale to 2,500 g and lower than 4,000 g, and elevated birth weight (Portuguese acronym EPN) those of higher or equal value to 4,000 g.[14,15]

The differences between birth weights informed by the guardians and birth weights registered in Sinasc were compared in the categories of socioeconomic indices and birth weight categories through the qui-square test. The means of the absolute differences in values referred by the guardians registered in Sinasc (birth weights informed by the guardians - the birth weights registered in Sinasc) were described with means and standard deviations.

The agreement between the birth weights informed by the guardians of the adolescents in ERICA with the ones registered in Sinasc was estimated by intra-class correlation coefficient (ICC),[16] to obtain a summary measure of agreement between referred information and registered information, and through the Bland-Altman graphic analysis.[17] The agreement in birth weight classifications in the low, adequate and elevated categories was also verified by the Cohen's *Kappa* index[16,18] of the Brennan-Prediger Coefficient or G-Index[19,20] and Gwet's Agreement Coefficient.[21,20] *Kappa* is the most frequently used index to assess agreement, however, its method of calculating the expected agreement by chance, according to the algorithm proposed by Cohen, results in overestimation.[21] Furthermore, *Kappa* is highly influenced by the table's marginal distribution, tending to be zero when there is a large heterogeneity in prevalence, such as the case of the prevalence of the different weight categories described in the literature.[22] STATA software version 14.0 and Win Pepi version 11.53[23] were used in the graphic and statistic analysis.

This study was approved by the ethics committee of research in the Institute of Collective Health Studies (Portuguese acronym IESC) of Federal University of Rio de Janeiro (Portuguese acronym UFRJ), under the number CAAE 26857814.7.0000.5286. Report Nº 652.120 in 2014. The nominal databases from Sinasc were obtained through a solicitation to the State Health Department of Rio de Janeiro, data's confidentiality being guaranteed.

Results

Of 5,042 adolescents assessed in ERICA in the state of Rio de Janeiro, 355 individuals born in the year 1995 were excluded because this year was not included in the analysis due to structural differences in Sinasc. Adolescents that did not bring back the Guardian's Questionnaire (n=1,702), or without the mother's name on it (n=39), necessary for the linkage process, were also excluded, as were 275 adolescents who were not born in the State of Rio de Janeiro, leaving 2,671 records of adolescents to be linked with the Sinasc databases. In the Sinasc databases from Rio de Janeiro, 1996 to 2002, there were 102,550 live births excluded who did not have the mother's name, resulting in 1,699,591 records. Out of the 2,671 Guardian's Questionnaire from ERICA adolescents submitted to linkage, 1,822 were identified in Sinasc, meaning that, the probabilistic relation presented a sensitivity of 68.2%. 154 records out of the 1,822 pairs identified as real were excluded for not containing birth weight information (one from Sinasc and 153 from ERICA), leaving 1,668 adolescents in the final sample.

The 849 non-linked teenagers had a mean 15.2 (standard deviation=1.4) years of age, most of them were female, had adequate weight and studied in a public school (Table 1). Between the pairs identified as real, most were also female, had adequate weight and studied in a public school. The 1,668 individuals in the final sample had a mean of 14.3 (SD=1.4) years of age, most were female, 19.5% was overweight and 10.5% obese, almost half of them had a mother with a level of education equal to or inferior to primary school, and most studied in public schools. In that same sample, 5.2% of the adolescents were from class A and one fourth were from classes C and D from the ABEP economical classification (2011), and, also, according to Sinasc records, 7.4% were born underweight (BPN) and 4.8% overweight (APN).

In the Guardian's Questionnaire, the information was given by mothers (79%), fathers (11%) grandparents (2.1%) siblings (0.7%), stepparents (0.2%), other relatives (1.0%) and other non-relatives (0.2%), however, 5.7% did not inform their relation with the adolescents.

In the category without the information about the mother's level of education in the Guardian's Questionnaire, 66.3%, of the respondents was the mother herself. In the other categories about the mother's level of education, this percentage varied from 76.6% (incomplete primary school) to 81.8% (complete secondary school). It was yet observed that the percentages of adolescents who studied in a public school in this category without information about the mother's level of education was 94.4% and in the other categories it varied from 31.0% (tertiary school) to 94.4% (incomplete primary school).

In total, approximately 38% of the weights informed by the guardians were a perfect match to the weight in Sinasc. The percentages of these answers without discrepancies were higher in

Table 1

Distribution of sociodemographic variables, gender, birth weight and nutritional state, in the linked clusters, non-linked, final sample, and in adolescents without Guardian's Questionnaire. Rio de Janeiro, 2013 and 2014.

Characteristics	Adolescents with Guardian's Questionnaire						Adolescents without Guardian's Questionnaire		*p*
	Linked		Non-Linked		Final Sample				
	n	%	n	%	n	%	n	%	
Total	1822	100.0	849	100.0	1668	100.0	1702	100.0	
Gender									< 0.001
Male	723	39.7	518	39.0	671	40.2	884	51.9	
Female	1099	60.3	331	61.0	997	59.8	818	48.1	
Nutritional state									0.04
Low and very low weight	46	2.5	25	3.0	41	2.5	33	1.94	
Adequate	1241	68.1	596	70.2	1134	68.0	1232	72.4	
Overweight	341	18.7	148	17.4	317	19.0	273	16.0	
Obesity	194	10.7	80	9.4	176	10.5	164	9.6	
Mother's educational level									< 0.001
Incomplete Primary School	439	25.9	206	24.3	392	23.5	208	12.2	
Complete Primary School	416	22.8	165	19.4	384	23.0	242	14.2	
Complete Secondary School	661	36.3	295	34.8	629	37.7	424	24.9	
Complete Tertiary School	181	9.9	98	11.5	171	10.3	294	17.3	
No information	125	6.9	85	10.0	92	5.5	534	31.4	
Type of school									< 0.001
Public	1511	82.9	723	85.1	1375	82.4	1279	75.2	
Private	311	17.1	126	14.8	293	17.6	423	24.9	
ABEP									< 0.001
A	94	5.2	58	6.8	86	5.2	149	8.8	
B	583	32.0	287	33.8	549	32.9	580	34.1	
C and D	169	25.7	220	25.9	420	25.2	329	19.3	
No information	676	37.1	284	33.5	613	36.7	644	37.8	
Birth weight									
Low birth weight	143	7.9	0	0.0	123	7.4	-	-	
Adequate birth weight	1595	87.5	0	0.0	1465	87.8	-	-	
Elevated birth weight	84	4.6	849	100.0	80	4.8	-	-	

Source: ERICA and Sinasc
ABEP: Socioeconomic criteria according to the Brazilian Association of Research Companies.
*Pearson's Chi-square (Between Final Sample and Adolescents without Guardian's Questionnaire).

mothers with complete secondary or tertiary school, and the higher discrepancies (differences higher than 100g) in mothers who did not complete primary school or in the category of no information on mother's level of education (Table 2).

The ICC, in the sample's total, demonstrates that high agreement was achieved, that is, low variability between the weights referred by the guardians and the ones registered in Sinasc of each adolescent when compared to the existing variability between them. The agreement is a little higher if analyzed only in the subset in which the mothers answered the questionnaire (ICC=0.90; CI95%= 0.88 to 0.91) and higher if we exclude 45 pairs with greater differences than 1.0 kg (most likely they are not pairs) , ICC = 0.93 (0.92 to 0.94). The elevated degree of agreement is also found in all the categories of analyzed variables, with the exception of the elevated birth weight category. There is an increasing gradient on the agreement according to how high the mother's educational level is (Table 3).

The agreement measures in the birth weight classifications in the low, adequate and elevated categories (Table 4) represented by Gwet first-order agreement coefficient (AC1) allow the observation of an elevated degree of total agreement in all categories of the sociodemographic variables. Cohen's Kappa agreement measure presented lower agreement values and Brennan-Predigger coefficient presented values closer to Gwet first-order agreement coefficient.

The elevated agreement between referred birth weights and birth weights registered in Sinasc can also be seen in the Bland-Altman graphic (Figure 1), in which the largest part of the point cloud inside the agreement limit of 95% can be found, and which has the largest concentration along the reference line of absence of difference between the referred birth weights and the Sinasc birth weights. The highest frequency of values in overestimation occurred between 3,000 g and 4,200 g of mean birth weights referred and registered in Sinasc, on the other hand, in underestimation, the highest frequency occurred between 2,200 and 3,000g of mean birth weights referred and registered.

Table 2

Differences between birth weight referred by the guardians of the adolescents and the real weight registered in Sinasc, total, and according to sociodemographic and birth weight classification variables. Rio de Janeiro, 2013 and 2014.

Variables	Diference*										Total		p**
	≤ -100 g		- 1 a - 99 g		0 g (None)		1 a 99 g		≥ 100g				
	n	%	n	%	n	%	n	%	n	%	n	%	
Total	338	20.3	175	10.5	632	37.9	153	9.2	370	22.2	1.668	100.0	
ABEP													0.483
A	20	23.3	12	14.0	28	32.6	8	9.3	18	20.9	86	100.0	
B	103	18.8	62	11.3	212	38.6	63	11.5	109	19.9	549	100.0	
C and D	88	21.0	40	9.5	165	39.3	32	7.6	95	22.6	420	100.0	
No information	127	20.7	61	10.0	227	37.0	50	8.2	148	24.1	613	100.0	
Mother's level of education													0.023
Incomplete Primary School	98	25.0	29	7.4	137,0	35.0	31	7.9	97	24.7	392	100.0	
Complete Primary School	72	18.8	51	13.3	137,0	35.7	38	9.9	86	22.4	384	100.0	
Complete Secondary School	109	17.3	65	10.3	257,0	40.9	69	11.0	129	20.5	629	100.0	
Complete Tertiary School	36	21.1	22	12.9	68,0	39.8	11	6.4	34	19.9	171	100.0	
No information	23	25.0	8	8.7	33,0	35.9	4	4.4	24	26.1	92	100.0	
Type of school													0.734
Public	276	20.1	139	10.1	529	38.5	125	9.1	306	22.3	1375	100.0	
Private	62	21.2	36	12.3	103	35.2	28	9.6	64	21.9	293	100.0	
Birth weight (g)													0.352
≤ 2.500 (BPN)	20	16.3	15	12.2	51	41.4	10	8.1	27	22.0	123	100.0	
2.500 - 3.999 (APN)	308	21.0	154	10.5	551	37.6	134	9.2	318	21.7	1465	100.0	
≥ 4.000 (EPN)	10	12.5	6	7.5	30	37.5	9	11.3	25	31.3	80	100.0	

Source: ERICA and Sinasc; * Referred birth weight minus the real birth weight registered in Sinasc; ** Qui-square test; ABEP= Socioeconomic criteria according to the Brazilian Association of Research Companies; BPN= Low birth weight; APN= Adequate birth weight; EPN= Elevated Birth Weight.

710

Rev. Bras. Saúde Matern. Infant., Recife, 17 (4): 705-715 out-dez., 2017

Table 3

Means, standard deviations (SD), mean differences, mean absolute differences and intraclass correlation coefficients (ICC) of the birth weights (g) referred by the guardians of the adolescents and registered in Sinasc, total and according to sociodemographic and birth weight variables. Rio de Janeiro, 2013 and 2014.

Characteristics	Mean referred birth weight (SD)	Mean Sinasc birth weight (SD)	Mean differences (SD)†	CI95%	Mean absolute differences (DP)††	CI95%	ICC	CI95%
Total	3202.8 (600.7)	3184.6 (517.3)	18.2 (354.5)*	1.2 – 35.2	185.6 (302.5)	171.1 – 200.2	0.89	0.88 – 0.90
Questionnaire informant								
Mother	3203.2 (601.7)	3168.5 (517.6)	34.7 (342.2)*	16.3 – 53.2	177.8 (294.4)	161.9 – 193.6	0.90	0.88 – 0.91
Father	3235.3 (614.2)	3268.9 (520.2)	-33.7 (345.4)	-84.1 – 16.7	198.2 (284.5)	156.9 – 239.4	0.90	0.86 – 0.92
Other	3122.7 (631.1)	3177.8 (573.5)	-55.1 (540.5)	-187.0 – 76.7	287.5 (459.7)	177.4 – 397.7	0.74	0.59 – 0.85
No information	3191.8 (538.1)	3251.6 (448.6)	-59.8 (355.7)	-132.3 – 12.6	199.5 (299.9)	138.4 – 260.5	0.85	0.77 – 0.90
ABEP								
A	3272.3 (563.2)	3264.0 (507.5)	8.26 (353.1)	- 67.5 – 84.0	202.0 (288.9)	140.9 – 263.1	0.88	0.82 – 0.92
B	3177.9 (627.5)	3163.4 (551.1)	14.5 (328.7)	- 13.1 – 42.1	169.4 (282.0)	145.8 – 193.0	0.92	0.90 – 0.93
C and D	3208.8 (629.6)	3184.8 (508.5)	23.9 (428.9)	-11.6 – 72.6	213.1 (372.8)	177.5 – 248.8	0.84	0.80 – 0.87
No information	3211.4 (559.8)	3192.4 (492.7)	19.0 (319.7)	-6.3 – 44.4	179.1 (265.4)	158.0 – 200.1	0.90	0.88 – 0.91
Mother's level of education								
Incomplete Primary School	3206.8 (620.4)	3197.3 (519.8)	9.46 (392.2)	-29.5 – 48.4	221.4 (323.7)	189.3 – 253.5	0.87	0.84 – 0.89
Complete Primary School	3222.0 (578.6)	3206.1 (490.9)	15.9 (367.3)	-20.9 – 52.8	185.4 (317.4)	153.6 – 217.2	0.87	0.84 – 0.89
Complete Secondary School	3204.0 (596.6)	3180.5 (514.9)	23.5 (317.3)	-1.3 – 48.4	154.7 (278.0)	133.0 – 176.5	0.91	0.90 – 0.92
Complete Tertiary School	3160.0 (606.2)	3144.7 (572.2)	15.3 (319.4)	-32.9 – 63.5	177.3 (265.8)	137.4 – 217.2	0.92	0.89 – 0.94
No information	3177.6 (630.4)	3143.3 (527.6)	34.3 (431.7)	-55.2 – 123.7	261.3 (344.3)	190.0 – 332.6	0.84	0.76 – 0.90
Type of school								
Public	3200.0 (608.1)	3178.8 (522.3)	21.2 (366.2)*	1.9 – 40.6	189.7 (313.9)	173.1 – 206.3	0.88	0.87 – 0.90
Private	3216.0 (565.1)	3212.0 (493.3)	4.0 (293.4)	-29.7 – 37.8	166.6 (241.3)	139.0 – 194.3	0.92	0.91 – 0.93
Birth weight (g)								
≤ 2.500 (BPN)	2149.1 (595.8)	2048.6 (361.2)	100.5 (456.8)*	19.0 – 182.1	205.9 (419.7)	131.6 – 280.1	0.72	0.59- 0.80
2.500 - 3.999 (APN)	3231.1 (462.8)	3223.0 (348.7)	8.1 (345.0)	-9.5 – 25.8	183.8 (292.1)	168.9 – 198.8	0.79	0.76- 0.80
≥ 4.000 (EPN)	4305.6 (347.8)	4229.4 (224.2)	76.1 (327.0)*	3.4 – 148.9	187.9 (277.5)	127.0 – 248.7	0.54	0.28 – 0.70

Source: ERICA and Sinasc; † Mean difference: difference between the referred mean values and the mean values registered in Sinasc; †† Mean absolute differences: mean absolute differences of the referred values and the values registered in Sinasc.
ABEP= Socioeconomic criteria according to the Brazilian Association of Research Companies;
SD= standard deviation; CI= Confidence Interval; BPN= Low birth weight; APN=Adequate birth weight; EPN= Elevated Birth Weight; * *p*<0,05 second t paired test.

Table 4

Kappa, G-index and Confidence intervals 95% between birth weight classifications in the Low, Adequate and Elevated categories, referred by the guardians of the adolescents registered in Sinasc, total and according to sociodemographic variables. Rio de Janeiro, 2013 and 2014.

Characteristics	*Kappa** (CI95%)	G-index** (CI95%)	AC1*** (CI95%)
Total	0.69 (0.65 – 0.74)	0.88 (0.86 – 0.90)	0.91 (0.89 – 0.92)
ABEP			
A	0.71 (0.50 – 0.91)	0.88 (0.79 – 0.96)	0.91 (0.84 – 0.97)
B	0.75 (0.68 – 0.83)	0.90 (0.86 – 0.93)	0.92 (0.89 – 0.94)
C and D	0.64 (0.54 – 0.73)	0.84 (0.80 – 0.89)	0.88 (0.84 – 0.91)
No information	0.67 (0.59 – 0.76)	0.89 (0.86 – 0.92)	0.92 (0.89 – 0.94)
Mother's level of education			
Incomplete Primary School	0.62 (0.52 – 0.73)	0.85 (0.80 – 0.89)	0.88 (0.85 – 0.92)
Complete Primary School	0.65 (0.55 – 0.76)	0.87 (0.83 – 0.91)	0.90 (0.87 – 0.93)
Complete Secondary School	0.80 (0.73 – 0.86)	0.92 (0.90 – 0.95)	0.94 (0.92 – 0.96)
Complete Tertiary School	0.72 (0.58 – 0.85)	0.87 (0.80 – 0.93)	0.90 (0.84 – 0.95)
No information	0.50 (0.29 – 0.72)	0.76 (0.64 – 0.87)	0.81 (0.71 – 0.90)
Type of school			
Public	0.69 (0.64 – 0.74)	0.87 (0.85 – 0.90)	0.90 (0.89 – 0.92)
Private	0.73 (0.61 – 0.84)	0.90 (0.86 – 0.95)	0.93 (0.89 – 0.96)

Source: ERICA and Sinasc;
ABEP= Socioeconomic criteria according to the Brazilian Association of Research Companies
CI= Confidence interval; * Cohen's *Kappa*; ** Brennan-Prediger Coefficient; *** Gwet's agreement coefficient.

Discussion

There was no difference between the birth weight referred by the guardian and the weight in Sinasc in less than half of the sample of adolescent participants of ERICA in the State of Rio de Janeiro (38%). However, when including cases in which the difference was lower than 100g, the similarity between the referred weights and the registered weights elevated to 58%. This means that in only a little more than half of the sample, weight information referred by the parents could be used as a substitute to the register at the time of birth. It should be considered that this should be expected, since the temporal distance between the confronted pieces of information was at least 12 years, the age of the youngest adolescents who participated in ERICA.

As for the classification of birth weight adequacy, the similarity or agreement between the information referred during adolescence and the one registered at birth was elevated. Therefore, the utilization of weight referred by guardians can result in a relative imprecision in attributing birth weight in each individual, but it is satisfactory as for the individual classification in the traditional classification categories (low, adequate or elevated birth weight). Silva *et al*.[24] evaluated the quality of

data on live births using the linkage method to verify the agreement between the Sinasc information and the data from a perinatal survey, and they verified a good reliability for birth weight data in the Sinasc, with high agreement (*kappa*=0.94) between low birth weight and perinatal research.

It was verified in this study the positive association of the mother's level of education with the precision of the referred birth weight information. The highest precision was given in the answers of mothers with highest degree of education, just like the highest imprecision was found in the information given by mothers with a lower degree of education. The agreement measured by the ICC of the guardians who did not inform, for any reason, is similar to the one found in the category of mothers with a lower educational level. Still, through the analysis of the confidence intervals of the mean absolute differences, it is not possible to confirm that there is relevance in the differences of these means in the categories of economic variables, mother's educational level, type of school or birth weight.

Reports of birth weight validation studies conducted in Brazil are scarce. In a cohort study conducted in Pelotas, Victora *et al*.[5] verified that in 61.2% of the cases, the informed weight was exactly the same as the real weight. However, in this study the children's age varied between 9 to 15 months, in

Figure 1

Graphic analysis (Bland-Altman) of the mean difference and Confidence interval (95%) of the referred values minus the values registered in Sinasc, in relation to the mean values referred and registered in Sinasc in adolescents born and assessed in the Study of Cardiovascular Risk in Adolescents – (ERICA) in the State of Rio de Janeiro, 2013 and 2014.
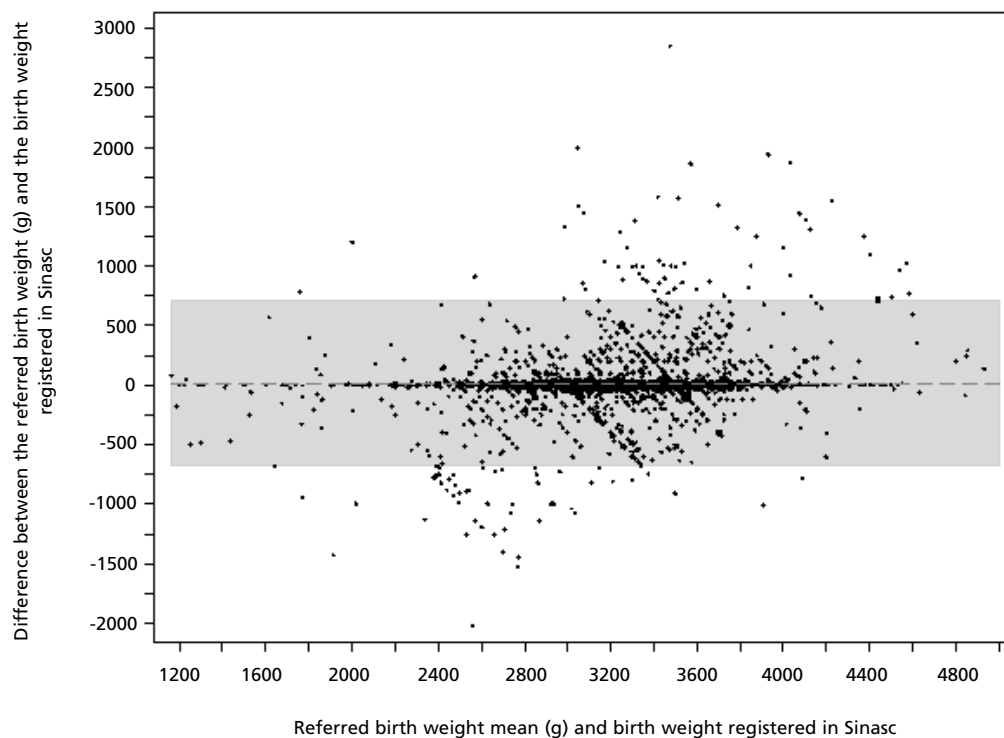
Sinasc= System of Information on Live Births.

a way that the chronological distance between the pieces of information was much shorter than in the present study with the adolescents from ERICA. Still in the Victora *et al.*[5] study, 79% of the answers presented a difference smaller than 100g, and 89.5% inferior to 250g, without any relevant tendency in any way, of weight increase or reduction. The same study also observed that the mothers who had never been to the school presented incorrect answers with a higher frequency than the others did. Victora *et al.*,[5] considered hospital records as the real weight, in this study we considered the weights registered in Sinasc as the true weight. Filha *et al.*[25] verified high concordance level (ICC over 0.90) in the birth weight registered in Sinasc, comparing them with data obtained from an interview with the puerperae and a survey in hospital records. Almeida *et al.*[26] also compared information from live birth certificates, which are used to fill Sinasc, with those obtained in interviews with the mothers, and the registries in the hospital records, and concluded that Sinasc is an excellent information source to identify low birth weight births.

The highest frequencies of answers with differences higher than 100g being present in the category of guardians who did not answer the mother's level of education indicates that these mothers were probably in a low level of education, since the percentage of children who studied in public school, whose mothers did not answer, was practically identical to the mothers with a referred lower level of education. Furthermore, between those who did not answer the mother's level of education, it was observed a smaller proportion of mothers as informants, which may have contributed to the discrepancies between the weights registered in Sinasc and referred in ERICA. Victora *et al.*[5] also observed a trend towards a decrease in the accuracy of information among mothers with lower level of education.

On the sample's total, the difference between mean referred and registered weights was inferior to 20 grams. However, some variation of these differences between the categories of the variables analyzed in this study occurred. Still, the highest difference reached approximately 100 grams only in case of the adolescents who were born with low weight, and the highest mean weight was referred by the guardians. This could be expected as a tendency of the guardians to attribute a better weight value to those who were born with this deficiency. In a smaller degree, the same happened to those who were born overweight, as well as with those who were born with adequate weight, in general there is an overall modest overestimation of birth weight by the guardians. This can be the result of a historical popular culture that babies who are big or have elevated weight are "healthy", becoming something for the parents to be proud of, despite the absence of evidences of that in literature. Some studies have reported that childhood obesity and overweight is not perceived correctly by those guardians, and that too many of these guardians do not care about children's overweight.[27,28] A German study evaluated the maternal perception of the child's silhouette and showed that some of the mothers of younger children preferred the child to be with overweight silhouettes.[29]

In the cases of school types, it was verified that in public schools the mean overestimation was approximately five times bigger than in private ones, even though the highest difference has been only a little over 20 grams. This finding might be related to the more elevated frequencies of adolescents of lower socioeconomic classes in public schools and of mothers with a low level of education. Victora *et al*.[5] verified greater errors and overestimation of birth weight responses in families with lower family income.

The present study found a high overall agreement between the referred weights and the weights registered in Sinasc according to ICC, which means that the variability between the values referred in ERICA and registered in Sinasc in each individual were considerably lower than the variance between birth weights of all adolescents from the sample. This is important to qualify studies of association between variables from ERICA in which one of them is the birth weight, increased to what has already been observed relating to small differences between the means. Elevated degree of agreement was also observed by Gwet's agreement coefficient, when classified birth weight in the low, adequate and elevated weight categories. This means that

there is high agreement between the categories of BPN, APN and EPN referred and registered in Sinasc. In fact, Gayle *et al*.[30] had already suggested that maternally reported birth weights are sufficiently accurate for research and programmatic purposes when birth certificate information is not readily available.

Despite the agreements measured by Cohen's *Kappa* being only reasonable, the agreements according to Gwet's coefficient were highly elevated, this indicator being more stable and reliable than the usual *Kappa*, because it is less dependent of prevalence distribution and homogeneous marginal, such as the case of the frequencies of the three weight categories in the ERICA sample.[21]

The results of this study show that the groups of adolescents who brought the Guardian's Questionnaire, linked and non-linked, and those who did not bring, were homogeneous in relation to several characteristics and therefore, the losses of questionnaires of the guardians and non-linked records did not bias the results. It is necessary to highlight that, the questionnaires with birth weight information were delivered by the teenagers, that took them to their respective guardians to answer at home, and they would give them back later. That means it is possible to assume that some of the guardians have consulted vaccination cards, characterizing a possible information bias. The impact of this possible information bias is the increase in the accuracy of birth weight data, increasing the concordance between the birth weight data referred to Sinasc.

This study possesses some limitations that must be mentioned. The final sample is not representative in relation to adolescents of the State of Rio de Janeiro, due to ERICA's complex sampling design, that does not contemplate representativeness in the Federation Units. It should be added to this the fact that only 36.1% of the 5,042 adolescent participants of ERICA in the State of Rio de Janeiro were assessed in this study, mainly due to losses when returning the Guardian's Questionnaire, and the merely reasonable efficiency of the linkage process. The Guardian's Questionnaire was self-filled, which might have led to losses and misinterpretation due to illegibility and equivocated transcription, besides the dependency on the memory of the interviewee (memory bias), after a minimum period of 12 years. It is yet possible to assume that there are typing or transcription mistakes in the Sinasc databases. Another source of mistakes is the process of database linkage, since it is a probabilistic

relationship, with imperfect sensibility and specificity, despite the manual revision of the doubtful pairs. To assess the accuracy and the quality of the linkage method, it is necessary to compare results with another source of information considered gold standard.[6-8] In a study to assess the sensibility of the probabilistic relationship between the data of the Pro-Saúde Study and the Sinasc databases, Coutinho *et al.*[7] verified a sensibility for the identification of births of 60.9% in reduction strategy, and of 72.8% in amplified strategy. The present study presented intermediary sensibility (68.2%) between these extremes. Therefore, this study is susceptible to information biases that can influence and distort the association measures.

The birth weight values referred in ERICA, in the State of Rio de Janeiro, by the guardians of adolescents from 12 to 17 years old, presented reasonable degrees of agreement with the values registered soon after birth in Sinasc, in individual levels. However, the agreements between mean weights were very high, which means that these can be used safely in population studies. The results also suggest that the mother's level of education is associated with the agreement between referred birth weight and the birth weight registered in Sinasc; the higher the level of education, the higher the precision and agreement of this measurement will be.

For better safety on the application of the present findings and recommendations, it is important, however, the conduction of similar studies in other regions of the country.

References

1. Kramer MS. The epidemiology of adverse pregnancy outcomes: An overview. J Nutr. 2003; 133: 1592S-96S.

2. Monteiro POA, Victora CG, Barros FC, Monteiro LMA. Birth size, early childhood growth, and adolescente obesity in a Brazilian birth cohort. Int J Obes. 2003; 27: 1274–82.

3. Rossi CE, Vasconcelos FAG. Birth weight and obesity in children and adolescents: a systematic review. Rev Bras Epidemiol. 2010; 13: 246-58.

4. Yu ZB, Han, SP, Zhu GZ, Zhu C, Wang XJ, Cao XG, Guo XR. Birth weight and subsequent risk of obesity: a systematic review and meta-analysis. Obes Rev. 2011; 12: 525-42.

5. Victora CG, Barros FC, Martines J, Béria JU, Vaughan JP. As mães lembram o peso ao nascer de seus filhos?. Rev Saúde Pública. 1985; 19: 195-200.

6. Camargo Jr. KR, Coeli CM. Going open source: some lessons learned from the development of OpenRecLink. Cad Saúde Pública. 2015; 31: 257-63.

7. Coutinho RGM, Coeli CM, Faerstein E, Chor D. Sensibilidade do linkage probabilístico na identificação de nascimentos informados: Estudo Pró-Saúde. Rev Saúde Pública. 2008; 42: 1097- 1100.

8. Silva JPL, Travassos CMR, Vasconcellos MM, Campos LM. Revisão sistemática sobre encadeamento ou linkage de bases de dados secundários para uso em pesquisa em saúde no Brasil. Cad Saúde Coletiva. 2006; 14: 197-224.

9. Bloch KV, Szklo M, Kuschnir MCC, Abreu GA, Barufaldi L, Klein CH, et al. The study of cardiovascular risk in adolescents - ERICA: rationale, design and sample characteristics of a national survey examining cardiovascular risk factor profile in Brazilian adolescents. BMC Public Health. 2015; 15: 94-103.

10. Fellegi I, Sunter A. A Theory for Record Linkage. Journal of the American Statistical Association 1969; 64: 1183-210.

11. Camargo Jr. KR, Coeli CM. OpenRecLink Guia do Usuário. Rio de Janeiro, 2012 [acesso em: 01 Out. 2015]. Disponível em: http://sourceforge.net/projects/reclink/files/?source= navbar

12. Onis MD, Onyango AW, Borghi E, Siyam A, Nishida C, Siekmann J. Development of a WHO growth reference for school-aged children and adolescents. Bull World Health Organ. 2007; 85: 649-732.

13. ABEP (Associação Nacional de Empresas de Pesquisa). Critério de Classificação Econômica Brasil; 2011.

14. WHO (World Health Organization). Physical status: the use and interpretation of anthropometry. Technical Report Series. Geneva; 1995.

15. WHO (World Health Organization). Guidelines on optimal feeding of low birth-weight infants in low and middle-income countries. Geneva: World Health Organization; 2011 [acesso em: 21 out. 2015]. Disponível em: http://www.who.int/child_adolescent_health/documents/9789241548366.pdf.

16. Fleiss JL, Cohen J. The equivalence of weighted kappa and the intraclass correlation coeficiente as measures of agreement. Educ Psychol Measures. 1973; 33: 613-9.

17. Bland J, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986; 1: 307-10.

18. Wongpakaran N, Wongpakaran T, Wedding D, Gwet KLA. Comparison of Cohen's Kappa and Gwet's AC1 when calculating inter-rater reliability coefficients: a study conducted with personality disorder samples. BMC Med Res. 2013; 13: 1-7.

19. Brennan RL, Prediger DJ. Coefficient Kappa: some uses, misuses, and alternatives. Educational and Psychological Measurement 1981; 41:687-99.

20. Gwet KL. Handbook of inter-rater reliability: the definitive guide to measuring the extent of agrément among raters. 2ª Edição. Gathersburg: Advances Analytics; 2010.

21. Gwet KL. Kappa statistic is not satisfactory for assessing the extent of agreement between raters. Statistical Methods for Inter-Rater Reliability Assessment, n. 1; 2002. [acesso em 25 fev. 2016]. Disponível em:

http:\\www.agreestat.com/research_papers/kappa_statistic_
is_not_satisfactory.pdf

22. Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. J Clin Epidemiol. 1993; 46: 423-9.

23. Abramson JH. WINPEPI updated: computer programs for epidemiologists, and their teaching potential. Epidemiol Perspect Innov. 2011; 8: 1-9.

24. Silva AAM, Ribeiro VS, Júnior AFB, Coimbra LC, Silva RA. Evaluation of data quality from the Information System on Live Births in 1997-1998. Rev Saúde Pública 2001; 35(6):508-14.

25. Filha MMT, Gama SGN, Cunha CB, Leal MC. Reliability of birth certificate data in Rio de Janeiro, Brazil, 1999-2001. Cad Saúde Pública. 2004; 20 (Supl. 1): S83-S91.

26. Almeida MF, Alencar GP, França Jr I, Novaes HMD, Siqueira AAF, Schoeps D, Campbell O, Rodrigues LC. Validation of birth certificates based on data from a case-control study. Cad Saúde Pública. 2006; 22 (3): 643-52.

27. Wright DG. Parental perception of preschool child body weight. J Pediatr Nurs. 2011; 26(5):435-445.

28. Santos DFB, Strapasson GC, Golin SDP, Gomes EC, Wille GMFC, Barreira SMW. The implications of family lack of concern and perception of obesity in childhood in the city of Curitiba, State of Paraná, Brazil. Ciênc Saúde Coletiva. 2017; 22 (5): 1717-24.

29. Warschburger P, Kroller K. Childhood overweight and obesity: maternal perceptions of the time for engaging in child weight management. BMS Public Health. 2012; 12: 295.

30. Gayle HD, Yip R, Frank MJ, Nieburg P, Binkin NJ. Validation of Maternally Reported Birth Weights among 46,637 Tennessee WIC Program Participants. Public Health Reports 1988; 103 (2): 143-7.