ORIGINAL ARTICLE

RELIABILITY OF ANTHROPOMETRIC MEASUREMENTS CONDUCTED IN NATIONAL PHYSICAL FITNESS STANDARD (SEGAK) ASSESSMENTS AMONG SCHOOL-AGED ADOLESCENTS IN TERENGGANU, MALAYSIA

Nurzaime Zulaily[1], Aryati Ahmad[1], Mohd Razif Shahril[1], Fadzli Syed Abdullah[2] and Amran Ahmed[3]

[1]Faculty of Health Sciences, Gong Badak Campus, Universiti Sultan Zainal Abidin, Kuala Nerus, Terengganu, Malaysia,
[2]Faculty of Informatics & Computing, Universiti Sultan Zainal Abidin, Tembila Campus, Besut, Terengganu, Malaysia,
[3]Institute of Engineering Mathematics, Pauh Putra Campus, Universiti Malaysia Perlis, Arau, Perlis, Malaysia.

Corresponding author: Aryati Ahmad
Email: aryatiahmad@unisza.edu.my

ABSTRACT

School-based health programs implemented by the Malaysian Ministry of Education (MOE) through the National Physical Fitness Standard (SEGAK) assessments provided an important platform in health status monitoring among schoolchildren. However, to date, there is still no reliability study conducted on this method. Therefore, this study aimed to determine the reliability of the anthropometric data collected by physical education (PE) teachers in the SEGAK assessments. Anthropometry measurements of standard six school adolescents involved in the Health of Adolescents in Terengganu study were taken by trained researchers using a standardised protocol. The anthropometrics data were then compared with PE teachers' measurements from the SEGAK assessments obtained from the specifically developed Health Monitoring System database. Reliability of the anthropometric measurements were analysed using Pearson's correlation test, Intraclass Correlation Coefficients (ICC), Bland-Altman plot and Cohen's Kappa statistics. Intraclass correlation coefficient between teacher-measured and researcher-measured values shows good correlation in weight (ICC = 0.93), height (ICC = 0.98) and BMI (ICC = 0.91). The Bland-Altman plot showed a relatively small difference in mean of weight, height, and BMI between teacher-measured and researcher-measured value. The mean difference between teacher-measured and researcher-measured value of weight, height, and BMI were 1.8kg, 0.1cm, and 0.8kg/m^2 respectively. Overall, Cohen's Kappa statistics showed substantial agreement (κ = 0.642) in BMI categorisation between the two measurements. Findings from reliability analysis conducted affirmed that anthropometrics assessments conducted by PE teachers in SEGAK assessments are reliable to be used for identification of body weight status among school children and adolescents particularly in Terengganu, Malaysia.

Keywords: Obesity, adolescent health, measurement issues, reliability

INTRODUCTION

Obesity among children and adolescents has become epidemic worldwide and more evident in developing countries including in Malaysia. Recently, Malaysian National Health and Morbidity Survey (NHMS) reported that the prevalence of obesity among children aged below 18 years has increased from 5.4% in 2006 to 11.9% in 2015[1,2]. A similar trend was reported in Terengganu in which the prevalence increased by almost two-fold (*i.e.* from 5.0 % in 2006 to 10.6 % in 2015) within a decade[1,2]. Therefore, in combating the obesity epidemic, one of the measures taken by the Malaysian government was by implementing school-based health-related programs including the National Physical Fitness Standard (SEGAK) assessments. The SEGAK program which had been implemented nationwide since 2008, involved Malaysian primary and secondary schoolchildren aged 10 to 17 years.

Previous studies asserted that school-based BMI measurement is important as a screening and surveillance tool for risk factors of certain diseases among children and adolescents[3,4]. SEGAK assessment is a mandatory test which is carried out bi-annually in March and August by trained Physical/Health Education (PE) teachers at school. PE teachers from schools nationwide were provided with hands-on training and modules by Ministry of Education (MOE) to conduct the SEGAK assessments to their students. The PE teachers were taught on the appropriate measurement technique to conduct the weight and height measurements together with the fitness assessments. For this study, anthropometric measurements of weight and height were conducted by PE teachers in each school. Although the PE teachers in each school have been provided with training and standard protocol module to follow, the measurements were still subjected to error. Despite BMI measurement in SEGAK assessments had been conducted nationwide for almost a decade, to date, there is still no reliability study conducted on this method.

Therefore, this study aimed to determine the reliability of the anthropometrics data collected by PE teachers in the SEGAK assessments.

METHODS

Sampling

Data on anthropometrics measurement of weight, height and BMI were obtained from 200 students who were enrolled in the Health of Adolescents in Terengganu study (HATs). This study employed random sampling method in which the 200 students were randomly selected from four selected primary schools in Besut and Kuala Terengganu districts. HATs' project is a school-based health status survey of 10 to 17 years old school adolescents conducted in primary and secondary schools throughout seven districts of Terengganu, Malaysia from November 2014 until May 2015[5].

Data collection

For the purpose of reliability study, a total of 200 school adolescents enrolled in HATs were randomly selected from four primary schools (n = 50 for each school). Weight and height were measured by trained researchers using calibrated scales (Tanita BC-587 Innerscan Body Composition Monitor and Seca 217 Portable Stadiometer) to the nearest 0.1 kg and 0.1 cm, respectively. In the school settings, the weight were measured by PE teachers using calibrated analogue health scales provided by the MOE, whilst for the height, PE teachers used wall-mounted height measuring chart. The anthropometric data from the researchers were then compared with the PE teacher's measurements which were extracted from the specifically developed database named Health Monitoring System (HEMS) [6.] The PE teacher's anthropometric measurements were from the second term of 2015 SEGAK assessments. Both the teacher-measured and researcher's measurements were taken less than two weeks apart.

BMI classification

BMI-for-age z-score of each student was then calculated using WHO AnthroPlus software[7]. Age of each participant was calculated to the precise days by subtracting the date of birth from the date of measurement while the BMI were calculated by dividing body weight in kilograms (kg) by height in metre squared (m[2]). Age- and sex- specific cut-off points according to 2007 WHO classifications[8] were used for BMI categories. The interpretation of the cut-offs classifies overweight as having z-score > +1SD (equivalent to BMI 25 kg/m[2] at 19 years), obesity as having z-score > +2SD (equivalent to BMI 30 kg/m[2] at 19 years) and thinness as having z-score <-2SD.

Ethical Approval and Permission

This study had obtained ethical approval from the Universiti Sultan Zainal Abidin Human Research Ethics Committee (UHREC) (Reference: UniSZA.N/1/628-1Jld.2 (11)). Permission to conduct the study and publish the data was obtained from the Malaysian Ministry of Education and Terengganu State Education Department.

Statistical analysis

The relationship between teacher-measured and researcher's measurements was assessed using Pearson's correlation coefficient where values of more than 0.7 were considered to have a strong correlation. Paired sample t-test was conducted to determine any significant difference between the two measurements. Intraclass correlation coefficient (ICC) was conducted to assess the overall reliability of the anthropometric measurements. The interpretation of ICC values were 0.41 to 0.60, fair; 0.61 to 0.80, moderate; and 0.81 to 1.0, substantial reliability[9]. Bland-Altman plot was used to explore the agreement between teacher-measured and researcher's measurements of weight and height. The agreement between the teacher-measured and researcher's measurements was quantified using the difference between the two measurements of weight and height for the same participant [$i.e.$ Limits of Agreement (LOA)]. The degree of agreement between teacher-measured and researcher-measured BMI categories was assessed using Cohen's kappa statistics. Guidelines for kappa statistics classification were: 0.81-1.00, almost perfect; 0.61-0.80, substantial; 0.41-0.60, moderate; 0.21-0.40, fair; and <0.21, poor agreement[10]. Statistical significance was accepted at P<0.05. Statistical analysis were carried out using IBM SPSS Statistics for Windows, Version 22.0 software (IBM Corporation, Armonk, N.Y. USA) and GraphPad Prism v5 for Windows (GraphPad Software Inc., San Diego, CA, USA).

RESULTS

Distributions of participants are presented in Table 1. A total of 200 standard six school adolescents with almost equal gender proportion ($i.e.$ male: 45.5%, female: 54.5%) were involved in this study. Larger percentages of participants were gathered from urban as compared to rural schools location (62.5% vs. 37.5%).

Correlation test

Table 2 presents the differences in anthropometric characteristics between teacher-measured and researcher-measured values. The difference in weight was significant between teacher-measured and researcher-measured values (1.8 kg, P=0.001) but not significant in height (0.1cm, P=0.679). The resulting mean difference in BMI between teacher-

measured and researcher-measured values were significantly different (0.8 kg/m2, P<0.001). The significant difference found in mean weight but not in mean height between teacher-measured and researcher-measured values indicating there was a bias in weight measurement. The intraclass coefficient for weight, height and BMI were 0.93, 0.98 and 0.91, respectively, which indicates substantial reliability. The correlations between teacher-measured and researcher-measured values tested using Pearson's correlation test showed strong association for weight and height (Pearson's, r=0.966, P <0.001) and (Pearson's, r = 0.864, P <0.001), respectively.

Table 1: Participant's distribution by school location and gender

	Frequency	Percent
Overall	200	100
Gender		
Male	91	45.5
Female	109	54.5
School location		
Rural	75	37.5
Urban	125	62.5

The differences in the BMI categories between teacher-measured and researcher-measured values were compared in Figure 1. Differences between prevalence of teacher-measured and researcher-measured values of overweight and obesity were 1.0% and 5.5%, respectively. Additionally, teacher-measured values reported higher prevalence of thinness and normal BMI than researcher-measured values by 4.0% and 1.5%, respectively.

Bland and Altman plot
The Bland-Altman plots for differences between teacher-measured and researcher-measured values on height, weight and BMI are shown in Figure 2. The mean difference indicates that teacher-measured weight and height were about 1.81 kg and 0.12 cm lower than researcher's measurement. Mean ± SD differences and the 95% limits of agreement between the teacher-measured and researcher-measured values were as follow: height (-0.12 cm ± 3.98), 95% limit of agreement, -0.77 to +7.91 cm; weight (-1.81 ± 3.84 kg) 95% limit of agreement, -5.71 to +9.34 kg. On the average, wide limits of agreement were observed for both height and weight measurement. Nonetheless, there was also evidence of random error for both height, weight and BMI measurements. Overall, 96.5% of difference in weight measurement and 93% of difference in height measurement of the participants fell within the LOA.
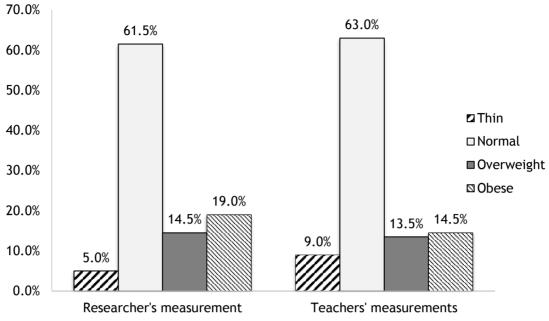
Cohen's Kappa statistic
Cohen's kappa (κ) statistics found substantial agreement between in BMI categories between teacher-measured and researchers-measured (κ=0.642, P <0.001) for overall participants (Table 3). By comparing between genders, male showed a higher level of agreements (κ=0.721, substantial agreements) as compared to female (κ=0.584, moderate agreement). In addition, there were significant correlations between teacher-measured and researchers-measured of BMI categories among overall participants (r=0.845, P < 0.001) and both male (r=0.882, P < 0.001) and female participants (r= 0.811, P < 0.001).

Table 2: Anthropometric characteristics of teacher-measured and researcher-measured values of the study population (n = 200)

	Measurement			ICC (95%CI)	Pearson's correlation coefficients (r)	P-value
	Researcher-measured	Teacher-measured	Differences			
Weight (kg)	44.4 ±14.4	42.6 ± 13.0	1.8 ± 3.8*	0.93 (0.90-0.94)	0.966	< 0.001
Height (cm)	148.7± 7.3	148.5 ± 7.8	0.1 ±3.9	0.98 (0.97-0.98)	0.864	< 0.001
BMI (kg/m²)	19.9 ± 5.7	19.1 ± 4.7	0.8 ± 2.2**	0.91 (0.89-0.93)	0.929	< 0.001

Data are mean ± SD, P-value for Pearson's correlation. Mean of weight/BMI between researcher-measured and teacher-measured values (**P<0.001,*P<0.05; Paired sample t-test). ICC= Intraclass correlation coefficient
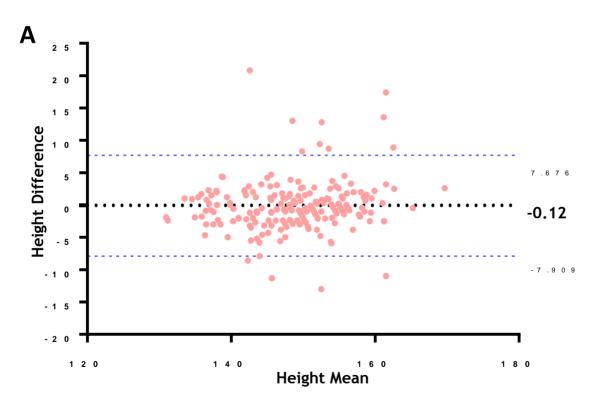
Figure 1: Percentage of BMI categories between teachers and actual measurements.

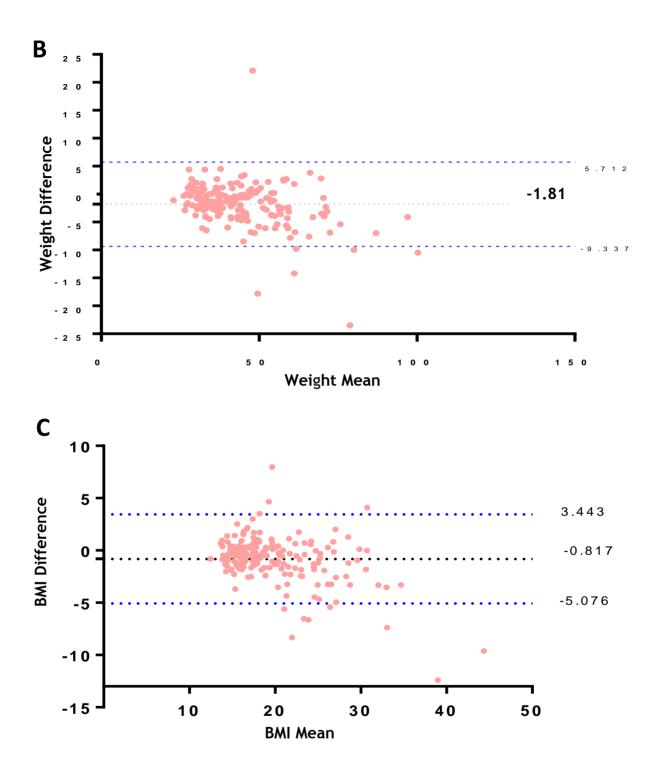**Figure2: Bland-Altman plots for differences between teacher-measured and researcher-measured values: (A) Height**

Figure 2: (*continued*) Bland-Altman plots for differences between teacher-measured and researcher-measured values: (B) Weight and (C) BMI

Table 3: Agreement levels between teacher-measured and researcher-measured BMI categories between genders.

	Kappa value (κ)[a]	P-value	Correlation (r)[b]	Kappa interpretation
Overall	0.642	< 0.001	0.845	Substantial
Male	0.721	< 0.001	0.882	Substantial
Female	0.584	< 0.001	0.811	Moderate

[a]Cohen's Kappa statistics for comparing the BMI categories between teacher-measured and researchers-measured among participants. [b]Teacher-measured BMI categories vs. researcher-measured BMI categories in overall, male and female participants (Spearman's correlation test). Kappa interpretation: 0.81-1.00, almost perfect; 0.61-0.80, substantial; 0.41-0.60, moderate; 0.21-0.40, fair; and <0.21, poor agreement[10].

DISCUSSION

The purpose of this study was to determine the reliability of teacher-researchers measurement and the extent of agreement between teachers-measured and researcher-measured anthropometric values. To our knowledge, this is the first attempt to evaluate the reliability of school-based anthropometric data focusing on PE teachers in Terengganu state. Findings from the present study suggest that PE teachers were able to provide reliable anthropometric measurements.

Similarly, in large scale multi-centre epidemiological studies, reliability study is important to be conducted as a part of data quality assurance purpose. The reliability of anthropometric measurement conducted by multiple observers had been assessed in previous studies worldwide including in the ALPHA project[11], ToyBox study[12], Healthy Lifestyle in Europe by Nutrition in Adolescence (HELENA) project[13], Identification and prevention of Dietary and lifestyle-induced health Effects In Children and Infants (IDEFICS) study[14] and FUPRECOL Study[15]. While reliability is the measure of ability to produce consistent with minimal variability within subjects, ensuring the data obtained are reliable to be used is crucial. In this current study, the inter-observer reliability between two independent observers (i.e. PE teachers and researchers) in conducting weight, height and BMI were assessed. In the community setting, Johnson et al. (2009)[16] assessed the reliability of anthropometric measurement from existing United Kingdom growth monitoring program named Personal Child Health Record (PCHR). Similar to SEGAK assessment, despite the PCHR program that has been conducted since 1990, lack of information had previously been established. Therefore, reliability assessment is conducted to ensure the routine data collected are reliable.

Overall, in agreement with previous study[17], teachers-measured weight and height met excellent reliability criterion (i.e. based on ICC values) suggesting that PE teachers' measurement was reliable. Nonetheless, despite high reliability was found in the PE teacher's measurements, the absolute mean differences criteria were not met (i.e. height: ≤ 0.25 cm and weight ≤ 0.25 kg). In the present study, the significant difference reported in the mean differences of weight, but not in height between the teacher-measured and researcher-measured values suggesting bias in measurement. Teachers' measurement underestimated the weight by 1.8 kg. Although a standardised protocol has been issued for anthropometric measurement, errors may still occur due to inter-observer or intra-observer differences. In addition, the discrepancies in measurements could be due to systematic mis-measurement resulted from equipment error or by human errors (i.e. recording/ transcription errors).

Differences in measurements may also be attributed by the daily diurnal variation in mass and stature. Claessens et al. suggested that daily weight fluctuation was nearly 1 kg resulting from changes in body fluid and GI contents[18]. On the other hand, stature is affected by time of the day. The value was found higher in early morning and decreased gradually during the day[18]. The impact of timing of anthropometric measurement on weight and height differences among school children had been extensively studied by Routen et al. (2010)[19]. By using data from the National Child Measurement Programme conducted in England, the author found that the timing of height and weight measurements (i.e. in the morning vs. in the afternoon) produced different readings. Similar to the finding reported by a previous study[18], Routen et al. also found that the stature to be decreased in the afternoon, whilst in girls, afternoon measurement produced greater weight thus increased the BMI and BMI percentile. This finding suggests that, there is a greater weight variation propensity in girls as compared to boys although the reasons for the gender dependency remain unclear. In addition, the lower agreement of BMI categories between teacher-measured and

researcher-measured among girls as compared to boys might possibly resulted from this gender-dependent weight variation.

Limit of agreement (LOA) in Bland-Altman plot was conducted to provide visual judgement on the degree of agreement between the two measurements. The present study also found that the LOA between the two measurements was wide suggesting the agreement between teachers- and researcher's measurements were low. Although the mean difference in weight and height between teacher-measured and researcher-measured values of overall participants appeared small, there were substantial differences found for some individuals. The differences between PE teachers'-measured and researcher-measured values ranging between -0.77 to +7.91 cm for height and -5.71 to +9.34 kg for weight. Since the measurements were taken within two weeks' time, it may have attributed to the differences observed. When comparing the BMI categories between teachers-measured and actual measurements, there was a substantial agreement between the two measurements. This finding showed that the BMI categories reported by PE teachers were reliable to be used as a screening tool to detect the risk of obesity among school adolescents.

To our knowledge, this was the first study conducted to determine the reliability of SEGAK data particularly among standard six school adolescents in Terengganu. Hence, findings from the reliability study proved that teachers-measured anthropometric data from SEGAK assessment can be used for identification of body weight status among school children and adolescents. This study which was conducted among school adolescents aged 12 years old however might not represent the older age group (i.e. among secondary school adolescents). Thus, further study is required to evaluate the reliability among the older age group. Besides, the equipment used by teachers and researchers were different, thus potentially contributed to the equipment bias. However, this study reflected the real scenario and setting during the SEGAK assessments at schools. Despite the limitations that were identified in this study, there were several recommendations that can be conducted to improve the implementation of SEGAK assessments. First, comprehensive training on measurement procedure, data managements, sources of bias and equipment handling should be provided to the PE teachers prior to the assessments. Apart from the training module provided by the MOE, PE teachers should also be provided with training video on the appropriate method of weight and height measurements. In addition, to minimise the biological variation, a standardised timing of anthropometric assessment should be implemented as a simple procedural changes in the SEGAK modules. Secondly, periodic reliability assessments and equipment maintenance should be conducted to reduce the measurement errors. Lastly, the effectiveness of the SEGAK assessments as the school-based BMI monitoring system can be improved by involvement of parents/caregivers. Parents who were notified on their child BMI status and provided with education materials were likely to find the school-based BMI monitoring system beneficial in improving obeso-protective behaviour and home environmental strategies[20].

CONCLUSION

In conclusion, SEGAK data submitted showed substantial agreement and correlation when compared with the researcher-measured anthropometric values. Reliability analyses conducted in this study asserted that anthropometric data obtained from SEGAK assessments was proven to be a reliable source for identification of body weight status among school children and adolescents. In turn, it is suggested that SEGAK assessments to be utilised as the school-based BMI monitoring system.

ACKNOWLEDGEMENT

REFERENCES

1. Institute of Public Health (IPH) 2015. National Health and Morbidity Survey 2015 (NHMS 2015). Vol. II: Non-Communicable Diseases, Risk Factors & Other Health Problems. Kuala Lumpur; 2015.

2. Institute of Public Health (IPH) 2011. National Health and Morbidity Survey 2011 (NHMS 2011). Vol.2: Non-Communicable Diseases. Kuala Lumpur; 2011.

3. Nihiser AJ, Lee SM, Wechsler H, McKenna M, Odom E, Reinold C, Thompson D, Grummer-Strawn L. Body mass index measurement in schools. J Sch Health. 2007;77(10):651–671.

4. Ruggieri DG, Bass SB. A Comprehensive Review of School-Based Body Mass Index

Screening Programs and Their Implications for School Health : Do the. J Sch Heal. 2015;85(1):61–72.

5. Aryati A, Nurzaime Z, Nor Saidah AM, Shahril MR, Wafa SW, Amin RM, Fadzli SA, Ahmed A. Body weight status of school adolescents in Terengganu, Malaysia: a population baseline study. BMC Public Health

6. Fadzli SA, Nor Saidah AM, Aryati A, Wafa SW, Shahril MR, Nurzaime Z, Amin RM, Ahmed A. HEMS : Automated Online System for SEGAK Analysis and Reporting. Int J Softw Eng Its Appl. 2016;10(10):89–104.

7. World Health Organization (WHO). WHO AnthroPlus for Personal Computers Manual Software for assessing growth of the world's children. 2009. p. 1–45.

8. World Health Organization (WHO). WHO Child Growth Standards [Internet]. 2007 [cited 2016 Oct 28]. Available from: http://www.who.int/growthref/en/

9. Shrout PE. Measurement reliability and agreement in psychiatry. Stat Methods Med Res. 1998;7:301–317.

10. Viera AJ, Garrett JM. Understanding Interobserver Agreement: The Kappa Statistic. Fam Med. 2005;37(5):360–363.

11. España-Romero V, Artero EG, Jimenez-Pavón D, Cuenca-Garcia M, Ortega FB, Castro-Piñero J, Sjöstrom M, Castillo-Garzon MJ, Ruiz JR. Assessing health-related fitness tests in the school setting: reliability, feasibility and safety; the ALPHA Study. Int J Sports Med. 2010;31(7):490-7.

12. De Miguel-Etayo P, Mesana MI, Cardon G, De Bourdeaudhuij I, Góźdź M, Socha P, Lateva M, Iotova V, Koletzko BV, Duvinage K, Androutsos O, Manios Y, Moreno LA; ToyBox-study group. Reliability of anthropometric measurements in European preschool children: the ToyBox-study. Obes Rev. 2014;15 Suppl 3:67-73.

13. Ortega FB, Artero EG, Ruiz JR, Vicente-Rodriguez G, Bergman P, Hagströmer M, Ottevaere C, Nagy E, Konsta O, Rey-López JP, Polito A, Dietrich S, Plada M, Béghin L, Manios Y, Sjöström M, Castillo MJ; HELENA Study Group. Reliability of health-related physical fitness tests in European adolescents. The HELENA Study. Int J Obes (Lond). 2008;32 Suppl 5:S49-57.

14. Stomfai S, Ahrens W, Bammann K, Kovács E, Mårild S, Michels N, Moreno LA, Pohlabeln H, Siani A, Tornaritis M, Veidebaum T, Molnár D; IDEFICS Consortium. Intra- and inter-observer reliability in anthropometric measurements in children. Int J Obes (Lond). 2011;35 Suppl 1:S45-51.

15. Ramírez-Vélez R, Rodrigues-Bezerra D, Correa-Bautista JE, Izquierdo M, Lobelo F. Reliability of Health-Related Physical Fitness Tests among Colombian Children and Adolescents: The FUPRECOL Study. PLoS One. 2015;16;10(10):e0140875.

16. Johnson W, Cameron N, Dickson P, Emsley S, Raynor P, Seymour C, Wright J. The reliability of routine anthropometric data collected by health workers: a cross-sectional study. Int J Nurs Stud. 2009 1;46(3):310-6.

17. Berkson SS, Espinola J, Corso KA, Cabral H, Mcgowan R, Chomitz VR. Reliability of Height and Weight Measurements Collected by Physical Education Teachers for a School-Based Body Mass Index Surveillance and Screening System. J Sch Health. 2013;83(1):21–27.

18. Claessens A, Beunen G, Malina R. Anthropometry, physique, body composition and maturity. Paediatr Exerc Sci Med. 2008. p. 23–36.

19. Routen AC, Edwards MG, Upton D, Peters DM. The impact of school-day variation in weight and height on National Child Measurement Programme body mass index-determined weight category in Year 6 children. Child Care Health Dev. 2011;37(3):360-7.

20. Bailey-Davis L, Peyer KL, Fang Y, Kim J-K, Welk GJ. Effects of Enhancing School-Based Body Mass Index Screening Reports with Parent Education on Report Utility and Parental Intent To Modify Obesity Risk Factors. Child Obes. 2017;13(2):164-171.