

# Ordinary Least Products Regression: A Robust Statistical Tool for Assessing Agreement Between Measures Attended by High Variability

Marcus Nascimento-Ferreira (✉ [marcus1986@usp.br](mailto:marcus1986@usp.br))

Universidade de Sao Paulo Faculdade de Medicina Hospital das Clinicas <https://orcid.org/0000-0002-6344-1044>

Augusto De Moraes

Universidade de Sao Paulo

Heraclito Carvalho

Universidade de Sao Paulo

---

## Technical advance

**Keywords:** agreement, continuous variables, fixed bias, methods comparison, proportional bias.

**Posted Date:** September 10th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-71462/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

## **Ordinary least products regression: a robust statistical tool for assessing agreement between measures attended by high variability**

**Running title: Agreement on behavioral sciences.**

Marcus Vinicius Nascimento-Ferreira, PhD<sup>1\*</sup>; Augusto César Ferreira De Moraes, PhD<sup>1,2\*</sup>; Heráclito Barbosa Carvalho, PhD<sup>1\*</sup>.

<sup>1</sup>YCARE (Youth/Child cArdiovascular Risk and Environmental) Research Group, Faculdade de Medicina, Universidade de Sao Paulo, Sao Paulo, Brazil; <sup>2</sup>Department of Epidemiology School of Public Health, University of Sao Paulo, Sao Paulo, SP, Brazil.

\*Address: Faculdade de Medicina da Universidade de São Paulo, Avenida Doutor Arnaldo, 455 - Cerqueira César, São Paulo - SP, 01246-903. Corresponding author tel.: +55 11 3061 7074; fax: +55 11 3061 7074. E-mail address: marcus1986@usp.br

### **Authorship**

Marcus V. Nascimento-Ferreira, Augusto César Ferreira De Moraes and Heráclito B Carvalho were responsible by article design, statistical analysis and scientific review.

### **Acknowledgments**

None.

### **Competing interests**

The authors declare that they have no competing interests.

### **Funding details**

Marcus V. Nascimento-Ferreira received a Ph.D. Student Internships abroad scholarship from the National Council of Technological and Scientific Development (CNPq, proc. 200340/2015-8) and a

Brazilian Ph.D. Student scholarship from the São Paulo Research Foundation (FAPESP, proc.2016/18436-8 and 2017/11732-3). Augusto César Ferreira De Moraes awarded by Young Investigator grant from FAPESP (proc. 2017/20317-0 and 2019/02617-1). Heráclito B Carvalho received a research grant from São Paulo Research Foundation (FAPESP, proc. 2014/11468-6) and an advanced scientist scholarship from National Counsel of Technological and Scientific Development (CNPq; proc. 300951/2015-9).

### **Availability of data and materials**

The simulated database is available in a supplementary file.

### **Abstract**

**Background:** Agreement (and disagreement) assessments are essential steps in the evaluation of new and existing methods. We aimed to provide a statistical approach to assess systematic disagreement between two measures/methods when both are attended by random error and high variability.

**Methods:** We applied ordinary least products (OLP) regression and the Bland-Altman method in six simulated pairs of samples. In OLP regression, fixed bias defined if 95% confidence intervals (CIs) of the intercept did not include 0. Proportional bias was defined if 95% CIs of the slope did not include 1. As a comparator, we assessed fixed and proportional bias by the Bland-Altman method. **Results:** We found divergence between studied statistical method outcomes only for measures with low variability (coefficient of variation,  $CV < 25,0\%$ ).

**Conclusion:** OLP regression is a simple and powerful tool for detecting systematic disagreement when the measures are attended by high variability, as well as behavioral variables.

**Keywords:** agreement; continuous variables; fixed bias; methods comparison; proportional bias.

## 1. Introduction

Agreement (and disagreement) assessments are essential steps in the evaluation of new and existing methods<sup>1-3</sup>. In agreement studies, measures from one or more methods compared with the measures obtained with one reference method on the same subjects<sup>3</sup>. However, biostatisticians insist that what should be sought is not an agreement between methods or measurers but disagreement or bias<sup>4</sup>. Disagreement estimates are essential for two main reasons: (i) to calibrate one method or measure against another and (ii) to detect bias, fixed or proportional<sup>1,4</sup>. Fixed bias means that one method gives values that are higher (or lower) than those from the other by a constant amount<sup>5</sup>. Proportional bias means that one method gives values that are higher (or lower) than those from others by an amount that is proportional to the level of the measured variable<sup>5</sup>.

In this sense, although there are still inappropriate applications of statistical methods in some studies, the Bland-Altman approach is the most popular method used to assess fixed and proportional bias in agreement research<sup>6</sup>. Conversely, Ludbrook argues that in favor of using ordinary least products (OLP) regression for comparing methods of measurement<sup>1,4,5,7</sup>. The author argument least products regression analysis is suitable for calibrating one approach against another; as well as, and it is also a sensitive technique for detecting and distinguishing fixed and proportional bias between methods<sup>4,5</sup>. However, this theory was built or at least supported based on a dataset (or simulated dataset) with low variability<sup>4,5,7</sup>. In these studies, the author suggests that OLP regression is a simple statistical tool to assess fixed and proportional using simulated measures of blood pressure (in mmHg). However, all tests were conducted in datasets with coefficient of variation (a variability estimator) lower than 25.0% and sometimes these tests showed divergent outcome (diagnosis) of systematic disagreement, fixed and proportional bias, in comparison with Bland-Altman method.

On the other hand, the agreement between measures from emotion (e.g., Well-Being) or behaviors (e.g., physical activity, sedentary behavior, sleep time), commonly assessed by subjective methods, present more considerable variability<sup>8-11</sup> than blood pressure measures<sup>12</sup>. A key point in behavioral sciences is an understanding of the sources of variability in behavior: i) systematic variability explains differences in behavior that result from manipulation of the independent variable and ii) unrelated variability represents differences in behavior that result from a systematic error (confounds) and random error<sup>13</sup>. Thus, the researcher attempts to maximize systematic variability and to eliminate systematic error and to minimize random error<sup>13</sup>. In this sense, it will be essential that the instrument has reduced fixed and proportional bias even in situations of the high variability of measurements.

In this article, we present the ordinary least products (OLP) regression intercept and slope confidence intervals (CIs) rationale and their implementation, and interpretation for assessing potential sources of systematic disagreement between two measures from different behaviors. We hypothesized that the OLP regression could also be a simple and powerful tool for detecting systematic disagreement, even in cases in which the measures present high variability.

## **2. Materials and methods**

To assess properties of OLP regression for assessing systematic disagreement, fixed and proportional bias, we have used a simulated scenario with hypothetical measures based on the literature and created with Stata 14 software (Stata Corp., College Station, TX, USA). The Bland-Altman method was used as a comparator to assess systematic disagreement. The hand-held calculator conducted both statistical analyses.

### **Simulated dataset scenario**

To perform examples of agreement assessments, we simulated six pairs of samples (N=100) using random normal distribution (SEED 010101). To generate the measures, the mean and standard deviation were retrieved from the literature<sup>11,14,15</sup>. The measures were (moderate-to-vigorous) physical activity in minutes/day, sedentary behavior in hours/day and sleep time (duration) in hours/day. To recreate typical agreement research<sup>15-17</sup> in each pair, one measure was a simulation of the variable assessed by questionnaire in youth, and the second one was the same simulation, however, assessed by accelerometer.

In this simulated scenario (described in Table 1), we simulated three pairs of measures with high correlation ( $r > 0.9$ )<sup>18</sup>, similar to Ludbrook studies<sup>4,5</sup>; as well as, we simulated three pairs of measures with low correlation ( $r < 0.3$ )<sup>18</sup>, similar to behavior studies<sup>10,11,14</sup>. In this case, the correlation was an outcome of agreement<sup>4,5,7</sup>. Also, among these variables, we also simulated two pairs of measures with low variability [coefficient of variation (CV) < 25.0%] as control agreement, similar to blood pressure measures where OLP regression was validated<sup>4,5</sup>. The simulated dataset presented in Supplementary Table 1.

### **Ordinary least products regression**

The OLP regression, also known as standard major axis regression<sup>5</sup>, is a technique designed to assess systematic disagreement obtained from two subjects or methods attended by random error<sup>4</sup>, especially when it is impossible to decide which should be regarded as a dependent<sup>5</sup>. Using the OLP regression intercept (constant) and slope confidence intervals (CIs). We can examine potential sources of systematic disagreement between two measures: fixed and proportional bias.

The OLP regression assumptions are<sup>1</sup>:

- The relationship between  $x$  (tested method) and  $y$  (reference method) is linear;
- Both  $x$  and  $y$  may be measured with an error or attended by random error;
- Errors are independent and normally distributed;

- The variance of the errors is constant (homoscedasticity);
- The standard deviations of  $x$  and  $y$  values are identical ( $S_x = S_y$ );
- The scales of measurement of  $x$  and  $y$  are the same.

### ***Fixed and proportional bias assessment***

In OLP regression, the sum of the products of the vertical and horizontal distances of the  $x$  (tested method) and  $y$  (reference method) values from the line minimized<sup>1,5</sup>. A detailed discussion about OLP regression models is available elsewhere<sup>1,4,5</sup>. The general formula for the OLP regression line is

$$E(Y) = a' + b'x$$

Where  $E(Y)$  is the predicted value of  $y$  as a function of  $x$ , in this sense, the OLP regression coefficients and their CIs can calculate several ways<sup>5</sup>. Here, we used one way to assess OLP coefficients and two ways to evaluate their 95% CIs that seem to be less sophisticated approaches<sup>1,4,19</sup>. Initially, it is necessary to find values of  $b$  (slope) in the ordinary least square (OLS) models where they are estimated by OLS regression of  $y$  on  $x$  and  $x$  on  $y$ , respectively.

$$E(y) = a + bx \quad (1)$$

$$E(x) = a + by \quad (2)$$

so

$$bx.y = r \frac{Sy}{Sx} \text{ or } by.x = r \frac{Sx}{Sy} \quad (3)$$

in addition

$$Sx = \sqrt{\frac{\sum(x - \bar{x})^2}{n - 1}} \quad (4)$$

$$S_y = \sqrt{\frac{\sum (y - \bar{y})^2}{n - 1}} \quad (5)$$

where  $r$  is the Pearson (product-moment) correlation coefficient<sup>20</sup>, obtained via

$$r = \frac{\sum ((x - \bar{x})(y - \bar{y}))}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}} \quad (6)$$

Secondly, the coefficient of the slope,  $b'$ , of the OLP regression is given by

$$b' = \frac{b_{y \cdot x}}{r} \quad (7)$$

Finally, because the OLP regression line passes through the point mean  $x$  and mean  $y$ , the intercept,  $a'$ , can be obtained from the formula<sup>5</sup>:

$$a' = \bar{y} - b'\bar{x} \quad (8)$$

To calculate the OLP regression intercept and slope (95%) CIs, it is necessary to assess the standard error based on standard major axis (SMA) regression (also called reduced major axis or RMA regression) technique<sup>21</sup>. CIs can be calculated by the Student's t-distribution and z-distribution method with  $n-2$  degrees of freedom<sup>22,23</sup>. However, there are other complex methods to calculate OLP regression intercept and slope CIs. For a detailed discussion about this topic, please see Ludbrook<sup>5</sup>. In the SMA technique, the OLP intercept and slope standard errors,  $SE$ , can be assessed by<sup>21</sup>

$$SE(a') = S_y \sqrt{\frac{1 - r^2}{n} \left(1 + \frac{\bar{x}^2}{S_x^2}\right)} \quad (9)$$

and

$$SE(b') = b' \sqrt{\frac{1 - r^2}{n}} \quad (10)$$

Where  $r$  is the Pearson (product-moment) correlation coefficient between  $x$  and  $y$ . Finally, the intercept and slopes CIs can be assessed using critical values of Student's-t (statistic) distribution [degrees of freedom (d.f.) =  $n-2$ ,  $\alpha = 0.05$ ] by the equations<sup>21,22</sup>

$$95\%CI(a') = a' \pm t \frac{\alpha}{2} \times SE(a') \quad (11)$$

and

$$95\%CI(b') = b' \pm t \frac{\alpha}{2} \times SE(b') \quad (12)$$

Another way used to assess OLP coefficients and their CIs can be using values of the z-distribution method ( $\alpha = 0.05$ ) based on normal sample distribution and the central limit theorem<sup>24</sup>

$$95\%CI(a') = a' \pm z \frac{\alpha}{2} \times SE(a') \quad (13)$$

and

$$95\%CI(b') = b' \pm z \frac{\alpha}{2} \times SE(b') \quad (14)$$

Finally, to interpret the sources of disagreement, fixed bias was defined if 95% CIs of the intercept did not include 0. Proportional bias was defined if 95% CIs of the slope did not include 1<sup>4</sup>.

### **Bland-Altman method**

Agreement between two methods or measures can be quantified using the differences between observations made using the two methods on the same subjects. *Bland* and *Altman* established a method to quantify the agreement between two quantitative measurements by constructing limits of agreement<sup>25</sup>. In the Bland-Altman method, the calculation of the 95% CIs for the mean difference provide a fixed bias estimate; whereas, the presence or absence of the correlation between averages and differences indicate proportional bias<sup>26</sup>. Another important inference assessed by Bland-Altman is the 95% limits of agreement (LOA, not addressed in this study) estimated by mean difference  $\pm 1.96$  standard deviation of the differences<sup>27</sup>. These limits indicate the range within which a single new observation taken from the same population would be expected to lie<sup>26</sup>. In this sense, LOA is not exactly a parameter for fixed and proportional bias. A comprehensive description of LOA characteristics has published elsewhere<sup>26,27</sup>. Moreover, the Bland-Altman method based on the

assumption that the differences are normally distributed<sup>27</sup>. Thus, investigators also should check the normality of disagreement between the tested and reference measure or method<sup>27</sup>.

### ***Fixed and proportional bias assessment***

In the Bland-Altman method, the Bland-Altman (average difference or LOA) plot and analysis are used to compare two measurements of the same variable<sup>27</sup>. That is, it is a method comparison technique.

$$average = \frac{x + y}{2}$$

$$difference = x - y$$

Firstly, it is necessary to estimate the average ( $a$ ) and difference ( $d$ ) between the tested ( $x$ ) and reference ( $y$ ) method or measure using the equations<sup>25,27</sup>

$$mean\ average\ (\bar{a}) = \frac{\sum(\frac{x + y}{2})}{n} \quad (15)$$

and

$$mean\ difference\ (\bar{d}) = \frac{\sum(x - y)}{n} \quad (16)$$

Thus, to assess fixed bias, mean differences, and their 95% CIs should be calculated<sup>27</sup>. The formula for 95% CIs for the mean difference is as follows<sup>27</sup>

$$95\%CI(\bar{d}) = \bar{d} \pm t_{0.5} \times SE(\bar{d}) \quad (17)$$

where  $t$  is the value of Student's t-distribution corresponding to two-sided  $\alpha = 0.05$  at  $n-1$  and  $n$  is the sample size, and  $SE$  is the standard error of the mean for difference<sup>27</sup> assessed by

$$SE(\bar{d}) = \frac{SD}{\sqrt{n}} \quad (18)$$

and the  $SD$  is the sample standard deviation for differences<sup>26</sup>

$$SD_{diff.} = \sqrt{\frac{\sum(d - \bar{d})^2}{n - 1}} \quad (19)$$

Besides, to examine proportional bias, the correlation between average and difference should be checked<sup>27</sup>. The (Pearson product-moment) correlation coefficient can be obtained via<sup>20</sup>

$$r = \frac{\sum((d - \bar{d})(a - \bar{a}))}{\sqrt{\sum(d - \bar{d})^2 \sum(a - \bar{a})^2}} \quad (20)$$

The level of significance (p-value, p) of the correlation coefficient assessed by the formula

$$t = \frac{r}{\frac{\sqrt{1 - r^2}}{\sqrt{n - 2}}} \quad (21)$$

where  $t$  is the value of Student's t-distribution corresponding to two-sided with  $df = n - 2$ <sup>20</sup>.

In disagreement assessment, an obvious starting point is a difference between measurements (fixed bias) of the two methods on the same subject<sup>27</sup>. Additionally, if the difference increases or decreases (proportional bias) progressively as the average value increases is another essential aspect of the disagreement diagnosis<sup>26</sup>. If the difference limits include 0, then it is assumed that no fixed bias is present<sup>25,27</sup>. On the other hand, if there is no correlation ( $p > 0.05$ ) between average and difference, the differences are uniform (homoscedasticity), and there is no proportional bias; however, if the differences increase or decrease progressively (heteroscedasticity,  $p \leq 0.05$ ), there is proportional bias<sup>25,27</sup>.

### 3. Results

Table 2 shows the summary of OLP regression disagreement analysis. The six hypothetical pairs of measures were then re-analyzed by the Bland-Altman method (Table 3), and the disagreement outcomes from the OLP and Bland-Altman methods shown in Table 4. In this sense, we found

partial divergence between studied statistical method outcomes only for measures with low variability ( $CV < 25,0\%$ ), also called control agreement. In addition, we found complete divergence when the measures presented low variability ( $CV < 25,0\%$ ) and low correlation ( $r \leq 0.1$ ). An example of the statistical implementation is presented in the Appendix.

#### **4. Discussion**

In this study, we provide a statistical procedure to assess systematic disagreement between two measures or methods when both attended by random error and high variability. Thus, the step-by-step of concepts and implementation of both methods were detailed. Practical examples of the OLP regression and Bland-Altman method implementation applied in a simulated scenario. The main novelty of our study is to demonstrate that when the measures attended by high variability, the OLP regression, and Bland-Altman method produce the same outcomes for fixed and proportional bias (sources of systematic disagreement). In this sense, our findings suggest that the OLP regression can detect fixed and proportional bias between two measures or methods attended by random error and high variability, as well as behavior variables (e.g., physical activity, sedentary behavior and sleep time).

##### **Fixed and proportional bias assessments**

In our simulated scenario, we assessed the disagreement between six pairs of measures: physical activity, sedentary behavior, and sleep time. As we hypothesized, we found the same outcomes (diagnoses) for fixed and proportional bias for all agreements with high variability. On the other hand, for measures with low variability, we found divergent outcomes of fixed and/or proportional bias between OLP regression and Bland-Altman method (Table 4). Our results are in line with Ludbrook studies<sup>4,5</sup>, where outcomes for fixed and proportional bias cannot be equal between OLP

regression and Bland-Altman method for blood pressure measures with high correlation and low variability (in a simulated scenario).

The divergence between OLP regression and Bland-Altman method can also be seen in a real sample (with high correlation and low variability) <sup>19,28</sup>. A potential explanation for this finding is attributable to the presence of a large amount proportional bias <sup>26</sup>. Briefly, if there is proportional bias, then the mean difference will almost inevitably deviate from 0 <sup>4,26</sup>. Thus, there is a risk that fixed bias will be overdiagnosed <sup>4</sup>. The exception to this is if proportional bias is in one direction (e.g., proportional bias > 0) and fixed bias in the opposite direction. Then the mean difference may be close to 0, and fixed bias will be underdiagnosed <sup>4</sup>.

Additionally, we hypothesized another potential explanation based on the evaluation of gene expression measurements <sup>29,30</sup>, where the coefficient of variation (variability) used as a degree of reproducibility. In this case, the comparison of measure agreement on the same set of subjects can lead to a comparison of dependent measures. And, the low variability can reinforce this dependence and confound systematic disagreement diagnoses. As suggested by Tan et al. <sup>29</sup>, a dependency between the datasets would confound any inferences we could make about the differences.

OLP regression is a simple way to evaluate the disagreement between two methods or measures, where the sum of the products of the vertical and horizontal deviations of  $x$ ,  $y$  values from the line minimized<sup>5</sup>. Likewise, the Bland-Altman method is a simple way to evaluate bias between the mean differences and to estimate an agreement interval<sup>27</sup>. Thus, note that OLP regression (as well as the Bland-Altman method) defines if there is fixed and/or proportional bias significantly different from 0. These statistical procedures do not say whether those differences are acceptable or not. Acceptable disagreements must be defined *a priori* and based on clinical necessity, biological considerations, or other goals.

### **Limitations and strengths**

Some restrictions should be noted in this study. Firstly, the OLP regression confidence intervals are assessed by the hand-held calculator. Although least products analysis is more difficult to execute by hand-held calculator than the Bland-Altman method, the bias diagnosis is easier to interpret. Performing OLP regression in statistical software can eliminate this difficulty. Besides, we achieved different values of confidence intervals coming from the t-distribution (used in this study) and z-distribution methods; however, the diagnosis/outcomes were the same. Some methodologists favor the use of more precise CIs for effect sizes; for example, CIs estimation based on *bootstrap* or *bias-corrected bootstrap* methods<sup>5,22</sup>. However, in previous experimental agreement assessments, we achieved more approximated (but not the same) by t-distribution confidence intervals compared with estimations made by statistical software<sup>4</sup>.

Moreover, in both cases, t-distribution and z-distribution, the diagnosis agreed with those found by Ludbrook<sup>4</sup>. Besides, our results and discussion were grounded in a simulated scenario, which can be considered a potential limitation. Studies in a real sample should be conducted to confirm our findings.

Despite these limitations, our study has several strengths. The present study provides a simple statistical method made by hand-held calculator to assess sources of disagreement, fixed and proportional bias, as well as an analytical method to evaluate disagreement when the reference method is sometimes known as the 'gold standard', but this does not, or should not, imply that it measured without error. Our study also provides an alternative to quantify proportional bias and to deal with cases in which there is a full presence of proportional bias affecting fixed bias estimation for the Bland-Altman method, which has been extensively used worldwide<sup>6</sup>. And, finally, this is the first time that OLP regression compared with the Bland-Altman method for assessing systematic disagreement when the measures attended by high variability.

In conclusion, ordinary least products regression is a simple and powerful tool for detecting systematic disagreement when the measures attended by high variability, as well as behavioral

variables. The convergence between least products regression and Bland-Altman method seems to improve when the variables accompanied by high variability. Our findings, in a simulated scenario, demonstrate that ordinary least products regression using intercept-slope confidence intervals interpretations could achieve an acceptable convergence for systematic disagreement outcome in comparison with the most popular statistical method used in agreement studies, providing scientific evidence that ordinary least products regression could be implemented as an analytical technique as Bland-Altman method at measures/methods attended by random error and high variability.

## References

1. Ludbrook J. Linear regression analysis for comparing two measurers or methods of measurement: but which regression? *Clin Exp Pharmacol Physiol*. 2010;37(7):692-699.
2. Watson PF, Petrie A. Method agreement analysis: a review of correct methodology. *Theriogenology*. 2010;73(9):1167-1179.
3. Bossuyt PM, Reitsma JB, Bruns DE, et al. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Ann Intern Med*. 2003;138(1):W1-12.
4. Ludbrook J. Statistical techniques for comparing measurers and methods of measurement: a critical review. *Clin Exp Pharmacol Physiol*. 2002;29(7):527-536.
5. Ludbrook J. Comparing methods of measurements. *Clin Exp Pharmacol Physiol*. 1997;24(2):193-203.
6. Zaki R, Bulgiba A, Ismail R, Ismail NA. Statistical methods used to test for agreement of medical instruments measuring continuous variables in method comparison studies: a systematic review. *PLoS One*. 2012;7(5):e37908.
7. Ludbrook J. A primer for biomedical scientists on how to execute model II linear regression analysis. *Clin Exp Pharmacol Physiol*. 2012;39(4):329-335.
8. Lee Y, Hofferth SL, Flood SM, Fisher K. Reliability, Validity, and Variability of the Subjective Well-Being Questions in the 2010 American Time Use Survey. *Soc Indic Res*. 2016;126(3):1355-1373.
9. Nascimento-Ferreira MV, De Moraes ACF, Toazza-Oliveira PV, et al. Reliability and Validity of a Questionnaire for Physical Activity Assessment in South American Children and Adolescents: The SAYCARE Study. *Obesity (Silver Spring)*. 2018;26 Suppl 1:S23-S30.
10. Nascimento-Ferreira MV, Collese TS, de Moraes ACF, Rendo-Urteaga T, Moreno LA, Carvalho HB. Validity and reliability of sleep time questionnaires in children and adolescents: A systematic review and meta-analysis. *Sleep Medicine Reviews*. 2016;30:85-96.
11. Nascimento-Ferreira MV, Moraes ACF, Rendo Urteaga T, Oliveira PVT, Moreno LA, Barbosa Carvalho H. Impact of methodological approaches in the agreement between subjective and objective methods for assessing screen time and sedentary behavior in pediatric population: a systematic review. *Nutr Hosp*. 2019;36(2):449-462.
12. Chen W, Chen F, Feng Y, Chen A, Zheng D. Quantitative Assessment of Blood Pressure Measurement Accuracy and Variability from Visual Auscultation Method by Observers without Receiving Medical Training. *Biomed Res Int*. 2017;2017:3537079.
13. Popović ZB, Thomas JD. Assessing observer variability: a user's guide. *Cardiovasc Diagn Ther*. 2017;7(3):317-324.
14. Nascimento-Ferreira M, De Moraes A, Toazza Oliveira P, et al. Assessment of physical activity intensity and duration in the paediatric population: evidence to support an a priori hypothesis and sample size in the agreement between subjective and objective methods *Obesity Reviews*. 2018.
15. Nascimento-Ferreira MV, Collese TS, de Moraes AC, Rendo-Urteaga T, Moreno LA, Carvalho HB. Validity and reliability of sleep time questionnaires in children and adolescents: A systematic review and meta-analysis. *Sleep Med Rev*. 2015;30:85-96.

16. Nascimento-Ferreira MV, De Moraes ACF, Toazza Oliveira PV, et al. Assessment of physical activity intensity and duration in the paediatric population: evidence to support an a priori hypothesis and sample size in the agreement between subjective and objective methods. *Obes Rev.* 2018.
17. Hidding LM, Altenburg TM, Mokkink LB, Terwee CB, Chinapaw MJ. Systematic Review of Childhood Sedentary Behavior Questionnaires: What do We Know and What is Next? *Sports Med.* 2016.
18. McHorney CA, Ware JE, Raczek AE. The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Med Care.* 1993;31(3):247-263.
19. Berkman C, Pereira K, Nardi G, et al. Agreement between i-STAT and YSI 2300 devices to determine lactate concentrations in dogs undergoing exercise. *Comparative Exercise Physiology.* 2016;12(2):75 - 82.
20. Dunn KW, Kamocka MM, McDonald JH. A practical guide to evaluating colocalization in biological microscopy. *Am J Physiol Cell Physiol.* 2011;300(4):C723-742.
21. KERMACK KA, HALDANE JB. Organic correlation and allometry. *Biometrika.* 1950;37(1-2):30-41.
22. Feingold A. Confidence interval estimation for standardized effect sizes in multilevel and latent growth modeling. *J Consult Clin Psychol.* 2015;83(1):157-168.
23. Greenland S, Senn SJ, Rothman KJ, et al. Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *Eur J Epidemiol.* 2016;31(4):337-350.
24. Barde MP, Barde PJ. What to use to express the variability of data: Standard deviation or standard error of mean? *Perspect Clin Res.* 2012;3(3):113-116.
25. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986;1(8476):307-310.
26. Ludbrook J. Confidence in Altman-Bland plots: a critical review of the method of differences. *Clin Exp Pharmacol Physiol.* 2010;37(2):143-149.
27. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res.* 1999;8(2):135-160.
28. Ibrahim J, Berk BC, Hughes AD. Comparison of simultaneous measurements of blood pressure by tail-cuff and carotid arterial methods in conscious spontaneously hypertensive and Wistar-Kyoto rats. *Clin Exp Hypertens.* 2006;28(1):57-72.
29. Tan PK, Downey TJ, Spitznagel EL, et al. Evaluation of gene expression measurements from commercial microarray platforms. *Nucleic Acids Res.* 2003;31(19):5676-5684.
30. Shoukri MM, Colak D, Kaya N, Donner A. Comparison of two dependent within subject coefficients of variation to evaluate the reproducibility of measurement devices. *BMC Med Res Methodol.* 2008;8:24.

## Appendix

### Simulated scenario: Fixed and proportional bias for a hypothetical variable

Considering the previous dataset (Supplementary Table 1), we present how to assess fixed and proportional bias using the variable physical activity (questionnaire vs. accelerometer comparison). Two techniques to achieve these parameters described. The first use OLP regression, and the second one uses the Bland-Altman method. All the calculations below were made with the help of a hand-held calculator.

### Ordinary least products regression

As we described above, the first step is to find the slope ( $b'$ ) of OLP regression. Thus, we performed a correlation between the two methods or measures

$$r = \frac{1823113.00}{\sqrt{(1897513.00)(2227996.00)}} \cong 0.8864 \quad (6)$$

Secondly, we found the  $b$  coefficient of OLS regression

$$b_{y.x} = 0.8864 \frac{138.4442}{150.0167} \cong 0.8180 \quad (3)$$

after that, we found the  $b'$  coefficient of OLP regression

$$b' = \frac{0.8180}{0.8864} \cong 0.9229 \quad (7)$$

and, the intercept is

$$a' = 259.1707 - 0.9229 * 255.7836 \cong 23.1060 \quad (8)$$

Where the accelerometer is the independent variable, and the questionnaire is the dependent variable. To calculate SMA regression standard errors, we used

$$SE(a') = 138.4442 \sqrt{\frac{1 - 0.8864^2}{100} \left(1 + \frac{255.7836^2}{150.0167^2}\right)} \cong 12.6681 \quad (9)$$

and

$$SE(b') = 0.9229 \sqrt{\frac{1 - 0.8864^2}{100}} \cong 0.0427 \quad (10)$$

In this sense, we calculate the intercept and slope CIs by the equations, and we consulted the table of Student's t-distribution (using d.f. =  $n-2 = 98$  and  $\alpha = 0.05$ , thus  $\alpha / 2 = .025$ , the two-sided t-value was 1.9845):

$$CI_{95\%}(a') = 23.1060 - 1.9845 \times 12.6681 \cong -2.0338 \quad (11)$$

$$CI_{95\%}(a') = 23.1060 + 1.9845 \times 12.6681 \cong 48.2458 \quad (11)$$

$$CI_{95\%}(b') = 0.9229 - 1.9845 \times 0.0427 \cong 0.8381 \quad (12)$$

$$CI_{95\%}(b') = 0.9229 + 1.9845 \times 0.0427 \cong 1.008 \quad (12)$$

### **Bland-Altman method**

In the Bland-Altman method, we firstly calculated

$$\text{mean average } (\bar{a}) = \frac{25747.71}{100} \cong 257.4771 \quad (15)$$

and

$$\text{mean difference } (\bar{d}) = \frac{338.7092}{100} \cong 3.387 \quad (16)$$

Secondly, we found the *SE* of the mean difference 95% CIs, where

$$SD_{diff.} = \sqrt{\frac{\sum(d - \bar{d})^2}{n - 1}} \cong 69.6467 \quad (19)$$

and

$$SE(\bar{d}) = \frac{69.6467}{\sqrt{n}} \cong 6.9647 \quad (18)$$

Thus, we assessed the 95% CIs, using *t* as the value of Student's t-distribution (using d.f. = n-1 = 99 and  $\alpha = 0.05$ , the two-sided t-value was 1.9842)

$$95\%CI(\bar{d}) = 3.387 - 1.9842 \times 6.96467 \cong -10.4322 \quad (17)$$

$$95\%CI(\bar{d}) = 3.387 + 1.9842 \times 6.96467 \cong 17.2064 \quad (17)$$

And, finally, we calculated the correlation coefficient between average and difference

$$r = \frac{-165241.5}{\sqrt{(480215.7)(1942701)}} = -0.1711 \quad (20)$$

We assessed the significance level by

$$t = \frac{-0.1711}{\sqrt{\frac{1 - (-0.1711)^2}{100 - 2}}} = -1.719 \quad (21)$$

Based on Student's t-distribution, the significance level (using d.f. = n-2 = 98 and t = -1.719, two-sided t-value) was 0.0887.



## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [1.SupplemantaryTable1Simulateddatabase.xlsx](#)