Author’s response to reviews

Title: Reliability and validity of a sedentary behavior questionnaire for South American pediatric population: SAYCARE Study.

Authors:

Augusto César De Moraes (augustocesar.demoraes@usp.br)
Marcus Nascimento-Ferreira (marcus1986@usp.br)
Claudia Lucia M Forjaz (cforjaz@usp.br)
Juan Carlos Aristizabal (juan.aristizabal@udea.edu.co)
Leticia Azzaretti (leazzaretti@hotmail.com)
Walter Nascimento Junior (walterjuniorufpi@hotmail.com)
Maria Miguel-Bergues (marymiguelberges@hotmail.es)
Estela Skapino (estalb@hotmail.com)
Carlos Delgado (cdelgado.isn@gmail.com)
Luis Moreno (lmoreno@unizar.es)
Heraclito Carvalho (heracc@gmail.com)

Version: 2 Date: 30 Nov 2018

Author’s response to reviews:

São Paulo, November 30th, 2018.

Dear Editor-in-Chief Dirk Krüger,

We are resubmitting a revised version of our manuscript (R1 BMRM-D-18-00088) entitled “Reliability and validity of a sedentary behavior questionnaire for South American children and adolescents: SAYCARE Study”.

We would like to thank the reviewers for the thoughtful and constructive comments. We have considered all the suggestions and have incorporated them into the revised manuscript. We believe that our manuscript is stronger as a result of these modifications and we hope you now consider the paper suitable for publication in BMC Medical Research Methodology.

Please, find below the reviewers’ comments in full, followed by our point-by-point responses (in bold). All changes made in the manuscript are highlighted in the track changes mode of MS Word in the
EDITORIAL COMMENT:
Your manuscript "Reliability and validity of a sedentary behavior questionnaire for South American pediatric population: SAYCARE Study." (BMRM-D-18-00088) has been assessed by our reviewers. They have raised a number of points which we believe would improve the manuscript and may allow a revised version to be published in BMC Medical Research Methodology.

Their reports, together with any other comments, are below. The reviewers have raised several important concerns and we look forward to your responses. Although all of the reviewers' comments are important, please note that both reviewers expressed concern that the methods and results do not support the conclusions you reach regarding feasibility, reliability, and validity. It is critical that you fully address this concern. Please also take a moment to check our website at https://bmrm.editorialmanager.com/ for any additional comments that were saved as attachments.

If you are able to fully address these points, we would encourage you to submit a revised manuscript to BMC Medical Research Methodology.

We would like to thank the Editor for these comments and the opportunity into improve our manuscript. We have read carefully your recommendations; as well as, the reviewers’ comments.

REVIEWERS' COMMENTS TO AUTHOR:

Dear reviewers, we would like to thank by time that you devoted to our work. We have read carefully your comments and recommendations. Please, find below our responses point-by-point.

REFEREE 1 (Wendy Yajun Huang)

1- This study investigated the reliability and validity (using accelerometry) of a sedentary behavior questionnaire for South American children and youth. The merit of this study is the under-researched target users of this questionnaire although there have been quite a number of sedentary behavior questionnaires available in the literature. As a result, stronger rationale is needed to justify how this questionnaire was developed taking into considerations of the cultural difference.

Thank you for your comment. We agree with your concern. However, previously to conduct this original manuscript about reliability and validity of a sedentary behavior questionnaire in pediatric population, we conducted a systematic review about this topic. In this review, we found 14 studies assessing the validity of sedentary questionnaires (against objective measures) and only one was...
addressed to South American population (Nascimento-Ferreira et al. Nutrición Hospitalaria, accepted to publish). Another systematic also found similar results (Hidding et al. Sports Medicine 2017; 47(4): 677-699). In addition, both systematic reviews, did not retrieve studies addressing questionnaire validity (against objective measure) based on multicenter approach. Thus, in the revised version, we have now provided a better rationale taking into considerations of the cultural difference in the INTRODUCTION section (page 4, line 94): “Although valid tools to evaluate sedentary behavior in South American pediatric population are scarce,[2-4] a recent systematic review showed only one questionnaire assessing psychometric properties in comparison with objective measure [4]. In this sense, in children and adolescents, no self-report or proxy-report sedentary behavior questionnaires are available that are both valid and reliable.[4] Thus, more high-quality research is required into the measurement properties of measurement instruments of sedentary behavior.[4, 6]

Additionally, prior literature suggests that ethnic and socioeconomic differences are observed in sedentary behavior trends.[7] One solution to this gap is to use standardized data from a multicenter approach.[8] In Europe, multicenter approaches are largely used to prepare questionnaires for epidemiological researches, regarding their design, feasibility, reliability and validity.[9-11] Conversely, in South American there is no sedentary behavior questionnaire with its reliability and validity tested in pediatric population with multicenter approach. The rationale behind the standardization using multicenter approach is based on transcultural adaptation focused in the equivalence of meaning and semantic of the questions among different languages and countries.[12] …”

2-Another major concern of this study is the limited domains of sedentary behavior measured in the questionnaire. It is not surprising to see the obvious difference between questionnaire-assessed and accelerometer-determined sedentary time (shown in table 3) because these two methods do not measure the same metric.

Thank you for your comment. We also agree with your concern. In our systematic review (Nascimento-Ferreira et al. Nutrición Hospitalaria, accepted to publish), we found questionnaires assessing 1 (Anderson et al. Child Dev 1985; 56: 1345-57) to 13 (Hardy et al. J Adolesc Health 2007; 40: 158-65) sedentary behaviors. And, the increased number of sedentary behaviors in the questionnaire was not related clearly with validity performance. One potential explanation could be the concomitant and intermittent natures of some sedentary activities are difficult to recall (Atkins et al. Int J Epidemiol 2012; 41(5):1460-1471). Thus, once that literature shows difficult in the adherence of educational programs (extrapolated here to questionnaire answering) in Latin-American pediatric population (MacArthur et al. Int J Obes Relat Metab Disord 2001, 25(8): 1262-1268), we opted by a simpler questionnaire and we included in the SAYCARE questionnaire the sedentary behaviors (risk factors) related to health outcomes (Bryant et al. Obes Rev 2007; 8(3):197-209); excluding the school time, for example (Beck et al. BMC Obesity 2016; 3:34). Additionally, the SAYCARE questionnaire was designed based on European questionnaires (e.g., Rey-Lopez et al. Eur J Public Health 2011; 22(3): 373–377) that also addressed up to 4-5 sedentary behaviors and the authors found interesting insights about epidemiology of sedentary behavior in European pediatric population (Garaulet et al. Int J Obes (Lond) 2011;35(10):1308-17). We have now added this information in the METHODS section (page 6, line 163): “… .The questions were designed to assess sedentary behaviors related to health outcomes in pediatric population.[3, 15] …”

Thus, sedentary behavior comprises a wide variety of activities, e.g watching television, quiet play, passive transport, and studying (Hidding et al. Sports Medicine 2017; 47(4): 677-699). In this sense, time spent in sedentary behaviors is not necessarily indicative of young people’s overall or total sedentary time, which also occurs in other contexts such as sitting during class time at school or during motorized transport (Cliff et al. obesity reviews 2016; 17: 330–344). However, sedentary time is also (and partially) composed by time spent in sedentary behaviors.
In our questionnaire, we are also interested in to assess the ability of four sedentary behaviors in to estimate the sedentary time by accelerometer. Based on the literature, accelerometers and inclinometers are acknowledged as both valid and reliable instruments for measuring “sedentary behavior” in children and adolescents; however, these measures are labor-intensive for researchers and are costly, and cannot provide information on the type and setting of sedentary behavior (Lubans et al. Obesity Reviews 2011; Atkins et al. Int J Epidemiol 2012; 41(5):1460-1471; 12:781–99; Hidding et al. Sports Medicine 2017; 47(4): 677-699). Or, at least, accelerometer remains the most widely applied instrument and serves as a reference method to validate sedentary behavior questionnaires (Hidding et al. Sports Medicine 2017; 47(4): 677-699). Unfortunately, no questionnaires assessing total sedentary behavior or other constructs of sedentary behavior with both a positive evidence rating for reliability and validity were available. Hence, we have no conclusive recommendation about the best available sedentary behavior self-report or proxy-report questionnaire in children and adolescents (Hidding et al. Sports Medicine 2017; 47(4): 677-699). Thus, efforts in the direction of studying the validity of sedentary behavior questionnaire are needed.

3-Abstract: it should be made clear in the abstract that for children, their parent responded to the questionnaire. Consequently, the results and discussion should be presented for parent-report and youth-report questionnaires separately.

Thank you for your suggestion. We have now provided this information. Please, see ABSTRACT section (page 3, line 67): “Methods: Children and adolescents from seven South American cities were involved in the test-retest reliability (n=161) and concurrent validity (n=187) studies. The SAYCARE sedentary behavior questionnaire was administered twice with two-week interval and the behaviors were parent-reported for children and self-reported for adolescents. Questions included time spent watching television, using a computer, playing console games, passive playing (only in children) and studying over the past week. Accelerometer was used for at least 3 days, including at least one weekend day. We compared values of sedentary time, using accelerometers, by quartiles of reported sedentary behavior time and their sum.”

4-Page 4, line 93: what do the authors mean when saying "a cross-cultural sedentary behavior questionnaire"? As said, more justifications should be provided here, e.g. what might be the cultural difference? How this difference might affect the development of a sedentary behavior questionnaire?

Thank you for your comments. We have now provided this information. Please, see METHODS section (page 6, line 163): “The SAYCARE sedentary behavior questionnaire was developed from questionnaires used in European multicenter studies. [9, 10] The questions were designed to assess screen time. [3, 15] In addition, questions about behaviors related to health outcomes in Brazilian pediatric population were included. [5] Based on varying cultural backgrounds, a cross-cultural questionnaire was adapted for all research centers (cities) in two versions/languages, Portuguese (for Brazilian cities) and Spanish (other cities), following the instructions for a tool design and development. [1, 5] In this sense, we performed a cross-cultural adaptation (with translation and back translation) from European Spanish to Brazilian Portuguese and South American Spanish languages in two steps i) meaning and ii) semantic equivalence which might be directly related with questions comprehension. [12]”

5-Page 4, power calculation: although the authors have provided detailed power calculation, they haven't considered the parents' sample and youth sample separately. It is inappropriate to combine parent-report and youth-report questionnaire in data analyses and interpretation of the data.

Thank you for your comments and suggestions. However, we analyzed parent-report and youth-report in separately over the manuscript. In this sense, we have provided now sample power (in separately) for parent-report and youth-report questionnaire. Please, see DISCUSSION section: (page 11, line 308):
“Our study has some limitations. First, although the sample was robust in size and diversity, the sample was not equally distributed among cities and school types. Second, our sample size was not calculated in separately for children (parent-report) and adolescents (self-report) although the analyzes were done separately. However, in post-hoc analysis the sample size from children (N = 55; β < 1%) and adolescents (N = 106; β < 1%) remained significant in power. …”

6-Page 5, line 133: it is not until this part that the readers understand that parents respond to the questionnaire for children. It should be mentioned in the abstract. Thank you for your suggestion. We have now provided this information. Please, see ABSTRACT section (page 3, line 67): “…The SAYCARE sedentary behavior questionnaire was administered twice with two-week interval and the behaviors were parent-reported for children and self-reported for adolescents. …”. Please, see METHODS section (page 5, line 152): “…For children, the questionnaire was parent-reported and for adolescents was self-reported. …”. Please, see RESULTS section (page 8, line 234): “…For total days, test-retest reliability of sedentary behavior time (min/day) for parent-report in children yielded a rho of 0.70; whereas, yielded a rho of 0.50 for self-report in adolescents. In addition, the reliability of quartile agreement and percentage of parent-report for children who followed the recommendations was k ≥ 0.40 (Table 2). …”

7-Page 5, lines 145-150: why only limited number of sedentary behaviors is included in the questionnaire? Any questions reflect the cross-cultural differences? Thank you for your comments. The cross-cultural differences were reflected in the language adaptation and the addition of sedentary behaviors other than screen time in according with a systematic review addressing Brazilian pediatric population (Guerra et al. Rev Saude Publica 2016; 50:9). The questionnaire was initially retrieved in European Spanish and then it was translated in two different languages (Brazilian Portuguese and South American Spanish) and back translated to European Spanish to ensure consistency across all language versions. Regarding limited number of sedentary behaviors included, we provided a detailed explanation in the QUESTION 2. Briefly, we opted to assess key behaviors related to health outcomes in pediatric population. Additionally, behavior with potential possibility of intervention and change in a future study (e.g., cohort study).

8-Page 6, calculation of weekly sedentary behavior: how to handle missing data? What happens if the respondent only responded to partial questions, e.g. only have data on weekdays? Thank you for your comments. To avoid increased questionnaire underestimation, we included only questionnaires with complete information. Please, see METHODS section (page 6, line 150): “Participants were excluded if they were pregnant or if their questionnaires were incomplete…” Regarding to second question, our questionnaire allows the researcher to analyze results in separately for weekdays and weekend days in absence of information. Whereas, for total days, the sedentary behavior should be restricted to individuals with complete information, due the variation of sedentary time during segments of the week (Ramirez-Rico et al. Eur J Sport Sci. 2014;14(3):287-93). We have now included this explanation. Please, see DISCUSSION section (page 11, line 319): “…Another concern may regard questionnaire response, for total days in future inferential researches, the sedentary behavior should be restricted to individuals with complete data, due the variation of sedentary time during segments of the week.[30] In the presence of missing data, we strongly encourage the measure of sedentary behavior for week and weekend days in separately…”

9-Page 6, lines 166-167: provide reference to support the decision of using 20 minutes of 0 counts as non-wear time. Thank you for your suggestion. We have now provided this reference. Please, see METHODS section (page 7, line 189): “… In addition, periods with 0 (zero) counts per minute (cpm) for more than 20
minutes were excluded as periods of non-use.[10]. …”

10-Page 6, line 168: how the accelerometer data was corrected by wearing time? Thank you for your suggestion. We have now provided this information. Please, see METHODS section (page 7, line 193): “… In this sense, sedentary time was expressed amount of time accumulated below 100 cpm during periods when the accelerometer was worn based on proportion of monitor-wearing time. [10, 19] …”

11-Page 6, lines 157-159 and lines 174-175: the sedentary behavior guidelines (no more than 2 hours per day) refer to screen-based behaviors only. Not appropriate to use this cut off for total sedentary behaviors measured by the questionnaire and accelerometer. Thank you for your comments. In line with your comments, the sedentary behaviors guidelines indicate: “The final guideline recommendations state that for health benefits, children (aged 5–11 years) and youth (aged 12–17 years) should minimize the time that they spend being sedentary each day. This may be achieved by (i) limiting recreational screen time to no more than 2 h per day — lower levels are associated with additional health benefits; and (ii) limiting sedentary (motorized) transport, extended sitting time, and time spent indoors throughout the day (Tremblay et al. Appl Physiol Nutr Metab 2011; 36: 59–64). In this sense, several studies have included other behaviors than screen time in the guidelines. These behaviors are related to sitting time (out of school), as we can see in multicenter studies (Rey-Lopez et al. Eur J Public Health 2011; 22(3): 373–377). In our study, we have used this classification only for questionnaires, in order to compare with European data. We have now provided this information. Please, see DISCUSSION section (page 11, line 314): “…Additionally, although the SAYCARE sedentary behavior questionnaire recorded information about four sedentary activities (including passive playing in children and studying in adolescents), it did not gather all the types of sedentary behaviors that adolescent were able to do.[30] In this sense, in order to compare with European data, [10, 30] we assessed reported sedentary behaviors, other than screen time, and ranked participants who attend media time recommendations (≤ 120 min/day).[17]”

12-Page 7, line 197: what does 'the quartile agreement' mean? I do not see any results relevant to this. Thank you for your comments. However, due the large amount of difference between questionnaire and accelerometer, we hypothesized if at least in categorical data (quartile) these measures could agree. In this sense, we tested the reliability of quartile agreement (questionnaire vs questionnaire); and, after that we tested the validity of quartile agreement (questionnaire vs accelerometer). We have now provided this information. Please, see METHODS section (page 7, line 209): “…We assessed unweighted k-coefficients to identify participants who followed the recommendations and weighted (quadratic) k-coefficient to compare objectively measured sedentary time by groups (quartiles) of self-reported sedentary behavior time.”

13-Discussion should be organized for parent-report and self-report questionnaire separately. Interestingly, there were obvious differences between first and second administration of the questionnaire for adolescents, but not for parents/children (shown in table 2). Any explanation? Thank you for your comments and suggestions. We have now improved the discussion. Please, see DISCUSSION section (page 9, line 261):
“Reliability and validity of parent-reported sedentary behavior in children
Findings of the present study indicate that the SAYCARE questionnaire has an acceptable reliability to measure parent-reported sedentary behavior time in the South American children. The reliability performance was probably due to the fact that parents were asked to recall their children’s usual specific behaviors.[17, 26] We also found that reliability was better to identify children who followed
the recommendations, than to identify time spent in sedentary behavior. This finding was particularly interesting and can likely be explained by the fact that sedentary behaviors tend to be more stable when compared to active behaviors and therefore, this can facilitate the recall process.[27]

Conversely, in our work, there was a low correlation between questionnaires and accelerometers. Our results are not far from previous findings, as a prior systematic review found that the correlation between these methods ranged from -0.16 to 0.55.[4] The limited validity of the reported measures compared with the accelerometer was partially explained in the literature: the concomitant and intermittent natures of some sedentary activities are difficult to recall,[1] and metrics (outcome) assessed by questionnaires and accelerometers are “distinct” and “not equivalent”. [6] However, the accelerometer remains the most widely applied instrument and serves as a reference method to validate sedentary behavior questionnaires.[4]

Moreover, the SAYCARE sedentary behavior questionnaire showed systematically underreport of sedentary time. The underreporting of sedentary time is not uncommon. Wen et al. showed a mean difference of 175 min/day between an sedentary behavior questionnaire and accelerometer in children. [26] In line with our findings, the authors argue that the subjective tool seems particularly useful to rank children by activity levels, but similarly to other reported tools, the parent-report overestimated time spent in physical activity and underestimate sedentary time compared with the accelerometer.[26]

Reliability and validity of self-reported sedentary behavior in adolescents
In adolescents, despite an acceptable reliability for self-reported sedentary behavior time, the reliability of quartile classification was lower than the acceptable values. We also found no reliability to identify adolescents who followed the recommendations. Evidence suggests that reliability for self-reported information regarding behaviors (e.g., physical activity and sedentary behavior) is lower in young population groups. [9] In addition, sedentary behavior seems not to be settled at young ages and has been characterized as unstructured activities. [9] Moreover, interpretation or understanding of questions is a factor affecting questionnaire reliability and the quality of the reports;[9] memory also plays an important role.[28] The results of the present study are in line with previous systematic reviews that found an acceptable reliability of the questionnaire for time spent in sedentary behavior.[2-4]

Our findings indicate that the correlation between self-reported minutes of sedentary behaviors from the SAYCARE and accelerometer-sedentary time was not acceptable. The questionnaire had inadequate ability to identify adolescents according to their sedentary time. Our results are in line with previous systematic reviews. [4, 29] In addition, findings in the present study suggest an important underestimation of the sedentary time compared to accelerometer. Similarly, Affuso et al. found a mean difference of 295 min/day between a self-reported leisure-time sedentary behavior and accelerometer in adolescents.[30] A possible explanation for the higher accelerometer measurements when compared with questionnaires could be the elevated values of absence of movement computed as sedentary time in the accelerometers whereas the questionnaire was developed to measure the time spent on four sedentary activities associated with health outcomes.[31, 32]”

14-It seems not convincing that the authors conclude that the questionnaire is feasible.
Thank you for your comments. We have now improved information about SAYCARE (general) sample, in order to clarify the feasibility of SAYCARE sedentary behavior questionnaire. Please, see METHODS section (page 4, line 114): “The current study is part of the South American Youth Cardiovascular and Environmental (SAYCARE) study. A pilot feasibility multicenter study which collected data from 237 children (3-10 years) and 258 adolescents (11-18 years), totalizing 495 participants. The inclusion criteria were: i) age ranging from 3 to 18 years old, ii) to sign an informed written consent and iii) provide information about sex and age. A detailed information about SAYCARE study was published elsewhere [13]. …”
Please, see RESULTS section (page 8, line 223): “Altogether 495 participants met the general SAYCARE inclusion criteria (age range 3–18 years, signed an informed written consent and provided information about sex and age). Among these, 415 (at least 83.8%) completed sedentary behavior Q1. Thus, we have found incomplete (or missing in the Q1) information about sedentary behavior in 16.2% of the valid sample. For the reliability study, data from 161 participants were analyzed (Q2 response rate, 38.8%). For the validity analysis, data from 187 participants were analyzed (Q2 response rate, 45.1%), …”

Please, see DISCUSSION section (page 9, line 249): “The SAYCARE sedentary behavior questionnaire achieved complete information from 83.8% of the participants; whereas, a European multicenter study about this topic achieved 75% of complete information. [10] Conversely, we found important decreases in the response rate from Q1 to Q2 surveys. …”

Said that, regarding questionnaire first application (Q1), we concluded that our questionnaire is feasible due to our missing data of 16.2% be lesser than missing data from of European questionnaire. As well as, the (raw) missing data was within the expected in the methodology.

15- It seems not convincing that the authors conclude that the questionnaire is reliable
Thank you for your comments. However, we disagree. Differently from biological variables, behaviors are more difficult to find stable measures. For example, a recent systematic review (with meta-analysis) about physical showed pooled (rho) agreement between questionnaire and objective measures around 0.30 (Nascimento-Ferreira et al. Obes Rev 2018; 19(6):810-824). In addition, similar original studies addressing sedentary behavior inferred validity and/or reliability using these same cut points, rho ≥ 0.30 and k ≥ 0.40 (e.g., Ridgers et al. J Sci Med Sport 2012, 15(2):136-141; Hong et al. Hong et al. Int J Behav Nutr Phys Act 2012, 9:93).

15-How can the questionnaire be used without acceptable validity?
Thank you for your comments. Evidences suggest that in children and adolescents, no self-report or proxy-report sedentary behavior questionnaires are available that are both valid and reliable (Hidding et al. Sports Med 2017;47(4): 677-699). In this sense, poor performance in convergent validity (against accelerometer) could not invalidate a questionnaire. Mainly because, although accelerometer is the most common instrument used as reference, there is no consensus in the literature that the metric from both methods (questionnaire and accelerometer) converge satisfactorily, as we can see, for physical activity, for example (Kelly et al. Int J Behav Nutr Phys Act 2016, 13(1):32). Conversely, the questionnaire showed acceptable reliability, for the first time, in South American Pediatric Population in a multicenter approach.

REFEREE 2

1-This is a small scale study on a convenience sample. It can be considered as a pilot study that tests the reliability & validity of the sedentary behavior questionnaire.
Thank you for your comments. We have now improved this information. Please, see METHODS section (page 4, line 115): “…A pilot feasibility multicenter study which collected data from 237 children (3-10 years) and 258 adolescents. …”

2-The authors have reached the wrong conclusion. The tool has poor reliability & validity.
Thank you for your comments. However, we disagree with you regarding the reliability. Differently
from biological variables, behaviors are more difficult to find stable measures. For example, a recent systematic review (with meta-analysis) about physical showed pooled (rho) agreement between questionnaire and objective measures around 0.30 (Nascimento-Ferreira et al. Obes Rev 2018; 19(6):810-824). In addition, similar original studies addressing sedentary behavior inferred validity and/or reliability using these same cut points, rho ≥ 0.30 and k ≥ 0.40 (e.g., Ridgers et al. J Sci Med Sport 2012, 15(2):136-141; Hong et al. Int J Behav Nutr Phys Act 2012, 9:93).

3-What is the program/formula used for the sample size calculation?
Thank you for your comments. We used the Stata “sampsi_rho” command. A detailed explanation about this sample size calculation was published recently (Nascimento-Ferreira et al. Obes Rev 2018; 19(6):810-824). In this sense, we have now provided this information. Please, see METHODS section (page 5, line 130): “We adopted assumptions of Nascimento-Ferreira et al. [14] for sample size estimation. …”

4-The vague sampling method is not representative of the general population. Thank you for your comments. We agree with your point of view. However, our sample did not pretend to be representative of the population. Our sample size and sampling method were not generated to achieved representativeness and sample was selected by convenience. However, the representativeness is not helpful in studying heterogeneity of effects across subgroups as Rothamn KJ and cols. described few years ago (Rothman et al. Int J Epidemiol 2014; 43(2):633-4). The rationale behind our convenience sample is that the differences between private and public schools across different countries can provide a good idea of the socio-demographic characteristics present in South America (Moreno et al. Int J Obes (Lond) 2008, 32 Suppl 5:S4-11). A common practice for this type of design study (Gonzalez-Gil et al. Obes Rev 2014; 15 (Suppl. 3): 61–66), where, the scientific significance has more importance than statistical ones (Rothman et al. Int J Epidemiology 2013;42:1012–1014).

In this sense, a previous European multicenter study called “HELENA study” designed in 2007 to evaluate the cardiovascular environment of a representative sample of European adolescents (Moreno et al. Int J Obes (Lond) 2008, 32 Suppl 5:S4-11). This study can be considered as the precursor and reference of the SAYCARE Study. The experience acquired in the HELENA was an important support to establish the SAYCARE study design and protocol. In this sense, first, as pilot and feasibility study, our study was organized on a small scale in each center to check every step of the procedure, from sampling to data processing. Second, the objectives of our study were to assess: (a) the collaboration of the school staff in the context of the sampling and data collection procedure; (b) the acceptability of the different blocks of data collection and (c) the importance of the different causes of missing data (refusal rate, loss of questionnaires, outliers or missing values) (Moreno et al. Int J Obes (Lond) 2008, 32 Suppl 5:S4-11; Carvalho et al. Obesity (Silver Spring) 2018, 26 Suppl 1:S5-S13). Moreover, although a general framework was established for sampling and data collection, some adaptation will be necessary depending on the local special conditions of the different center for the cross-sectional study.

Thus, the sample of our study was designed to find significant agreement (reliability and validity) between measures among different socioeconomic conditions, and identify the difficulties of a multicenter study in the context of South America; rather than, represent the general population. Findings of the SAYCARE study data gave us the opportunity to understand and modify the sample size weight, especially those addressed directly to participants public school and low education level. Based on SAYCARE sample, it was possible to study psychometric properties (e.g., reliability and validity) of at least four questionnaires to assess: socioeconomic status (De Moraes et al. Obesity 2018; 26, S14–S22), environmental factors (De Moraes et al. Obesity 2018; 26, S14–S22), physical activity (Nascimento-Ferreira et al. Obesity 2018; 26: S23–S30) and food frequency (Saravia et al. Obesity (2018) 26, S31–S40).
5-What is the language of the questionnaire?
Thank you for your comments. We have now provided this information. Please, see METHODS section (page 6, line 165): “… Based on varying cultural backgrounds, a cross-cultural questionnaire was adapted for all research centers (cities) in two versions/languages, Portuguese (for Brazilian cities) and Spanish (other cities), following the instructions for a tool design and development. [1, 5] …”

6-The study has a very low response rate for the reliability study (which may invalidate its results).
Thank you for your comments and suggestions. Although Q2 showed low response rate, in post hoc analysis the sample remains significant. We have provided new sample power for parent-report and youth-report questionnaire. Thinking about this possible limitation, we performed the post-hoc sample size power analysis (describe here the parameters). Even this loses/refusals, our sample size keeps sufficient to test the reliability. We introduce these informations on the discussion section. Please, see DISCUSSION section: (page 11, line 308): “Our study has some limitations. First, although the sample was robust in size and diversity, the sample was not equally distributed among cities and school types. Second, our sample size was not calculated in separately for children (parent-report) and adolescents (self-report) although the analyzes were done separately. However, in post-hoc analysis the sample size from children (N = 55; β < 1%) and adolescents (N = 106; β < 1%) remained significant in power. …”

7-The manuscript needs more elaboration on data analysis.
Thank you for your comments. However, we studied exhaustively data analysis about agreement between methods/measures [e.g., Ludbrook. Clin Exp Pharmacol Physiol. 2002;29(7):527-536; Ludbrook. Clin Exp Pharmacol Physiol. 2010;37(2):143-149; Barth et al. BMJ. 2017;356: j14; Zaki et al. PLoS One. 2012;7(5):e37908; Bland JM, Altman DG.. Lancet. 1986;1(8476):307-310; Kottner et al. Int J Nurs Stud. 2011;48(6):661-671]. Thus, we performed data analysis step-by-step following: i) define an a priori hypothesis based on measurement or method of interest, ii) calculate the sample size, iii) specify the statistical method used as agreement estimate, iv) specify the power of analysis, v) specify whether the alternative is one-sided or two-sided, vi) examine the sample distribution of the variable of interest, vii) describe potential differences between the main sample or population and subsample(s): sensitivity analysis, ix) based on sample distribution, describe the data, x) report data central tendency, xi) report data dispersion, xii) assess the relationship between measures or methods (estimates of agreement and report estimates of disagreement), and xiii) provide post-hoc confirmation of the sample power.
Thus, we have now improved data analysis information. Please, see METHODS section: (page 7, line 203): “Analyses were conducted in Stata 14 software (StataCorp, College Station, TX, USA). Shapiro-Wilk was used to determine sample normality. Descriptive analyses of demographic and socioeconomic variables are presented. In the sensitivity analyses, we assessed differences between categorical variables using the Chi-square goodness-of-fit test. We measured the test–retest reliability, agreement between Q1 and Q2, using a Spearman correlation coefficient (rho) for continuous variables and Cohen’s kappa (k) agreement for categorical variables. We assessed unweighted k-coefficients to identify participants who followed the recommendations and weighted (quadratic) k-coefficient to compare objectively measured sedentary time by groups (quartiles) of self-reported sedentary behavior time.
To assess the validity, for agreement between Q1 and accelerometer, we used Spearman correlation coefficient and kappa agreement. A moderate Spearman correlation (rho ≥ 0.30)[21, 22] and kappa coefficient (k ≥ 0.40)[23] were considered acceptable agreements. Additionally, we used Bland-Altman analysis to assess disagreement between Q1 and accelerometer. The range of agreement was defined as
the mean bias ± 1.96 standard deviations (and 95% limits of agreement, LOA) and heteroscedasticity between the measures was checked using the Pitman (Trend) test.[24] Heteroscedasticity was examined to verify whether the absolute intermethods difference (bias) was associated with the magnitude of the mean of sedentary behavior measured (ie, intermethods mean).[24] We considered a P value < 0.05 statistically significant.”

8-There are some incomplete references, e.g., ref 4 & 5.
Thank you for your suggestions. We have fixed now these references.