Original Article

RBAFS Revista Brasileira de Atividade Física & Saúde
SOCIEDADE BRASILEIRA DE ATIVIDADE FÍSICA & SAÚDE

rbafs.org.br

# Reliability and concurrent validity of the Global Physical Activity Questionnaire in adults from a Brazilian rural area

## Teste-reteste e validade concorrente do Questionário Global de Atividade Física em adultos de uma população rural no Brasil

AUTHOR'S

Rafaela Costa Martins[1]
Cauane Blumenberg[1]
Inácio Crochemore Mohnsam da Silva[1]

1 Post-Graduate Program in Epidemiology, Universidade Federal de Pelotas, Pelotas, Rio Grande do Sul, Brazil.

CORRESPONDING

Rafaela C Martins
*rafamartins1@gmail.com*
Universidade Federal de Pelotas. Post-Graduate Program in Epidemiology.
Rua Marechal Deodoro, 1160, 3º piso.
Pelotas, Rio Grande do Sul, Brasil. CEP: 96020-220.

DOI

10.12820/rbafs.23e0017



ABSTRACT

The aim of this study was to test reliability of the Global Physical Activity Questionnaire (GPAQ), and its concurrent validity using accelerometers, in a Brazilian rural population. Fifty-five adults (18+ years) living in a rural Brazilian city were selected to wear an accelerometer for one week and to answer the GPAQ twice in a week period interval. Reliability was verified using Intraclass Correlation Coefficients and the concurrent validity was analyzed based on Bland and Altman's diagrams. We found relatively high reliability for the total physical activity (ICC = 0.78; 95%CI: 0.63; 0.87). The average difference between both GPAQ measures was 192 minutes (95%CI: -69.63; 453.75). Wider differences were found when participants reported > 400 minutes of physical activity per week. There was limited agreement between the GPAQ and accelerometer measures (total average difference = -1043 minutes; 95%CI: -1452.23; -634.09). Thus, the questionnaire presented limited concurrent validity and acceptable reliability, being able to be used in rural populations. However, there is still a need of new physical activity instruments focused on rural populations.

**Keywords**: Physical activity; Reliability and validity; Reproducibility of findings; Rural population; Accelerometry.

RESUMO

*O objetivo deste estudo foi realizar um teste-reteste do Questionário Global de Atividade Física (GPAQ) e sua validade concorrente com acelerômetros em uma população rural brasileira. Cinquenta e cinco adultos (18 anos ou mais) moradores da área rural da cidade de Arroio do Padre foram selecionados para responder o GPAQ em duas semanas subsequentes e usar um acelerômetro por uma semana. Para realizar o teste-reteste foi usado o coeficiente de correlação intraclasse (CCI) e a validade concorrente foi analisada usando gráficos Bland and Altman. O resultado do teste-reteste de atividade física total foi alto (CCI = 0,78; IC95%: 0,63; 0,87). A diferença média entre as duas medidas do GPAQ foi 192 minutos (IC95%: -69.63; 453.75). Maiores diferenças foram encontradas quando os participantes reportaram praticar mais de 400 minutos de atividade física por semana. A concordância foi limitada entre as medidas do GPAQ e do acelerômetro (diferença média total = -1043 minutos; IC95%: -1452.23; -634.09). Logo, o GPAQ apresentou limitada validade concorrente e aceitável teste-reteste, podendo ser usado em populações rurais. Contudo, a criação de um novo instrumento de atividade física focado na população rural é necessária.*

**Palavras-chave**: *Atividade física; Confiabilidade e validade; Reprodutibilidade dos testes; População rural; Acelerometria.*

## Introduction

Studies in the 50's started suggesting that physical activity was beneficial for health[1]. After half a century, the health-benefits of physical activity are now widely known[2,3,4], and over five millions deaths per year worldwide are attributable to physical inactivity[4]. Despite its importance for public health, the measurement of physical activity has always been challenging. Although there is a trend towards the objective assessment of physical activity, particularly using accelerometers[5], most studies worldwide still rely on questionnaires, particularly for surveillance purposes.

In order to assess physical activity among adult populations, two are the questionnaires mostly used: the International Physical Activity Questionnaire (IPAQ) and the Global Physical Activity Questionnaire (GPAQ). Although IPAQ was developed earlier being widely used worldwide[6], particularly in Latin America[7], GPAQ has been used as part of the World Health Organization's (WHO) Stepwise approach

towards more data availability worldwide on physical activity. GPAQ evaluates physical activity in low and middle-income countries (LMICs) and is recommended for use both in rural and in urban areas[8,9,10].

Despite this recommendation, few studies have evaluated the reliability and validity of GPAQ in rural areas. A Pubmed search combining the terms (validity OR reliability OR repeatability) AND (GPAQ) AND (rural) resulted in only three articles. One study, published in 2009, which used data from nine countries and found that slight higher reliability results were found in urban as compared to rural areas[11]. The other two are more recent studies and were conducted in India[12,13]. No studies were found on the reliability and validity of GPAQ in Brazil.

Besides that, as more than 50 million of the Brazilian population lives in rural areas, it is important to have an instrument that has a reliable measure for physical activity[14]. The purpose of this study was to test the 7-days reliability of the GPAQ in a rural Brazilian population, and to assess concurrent validity comparing GPAQ results to accelerometry-based physical activity.

## Methods

Adults (18+ years) living in the city of Arroio do Padre, Brazil, were selected. Arroio do Padre is a small rural city, with approximately 3,000 inhabitants. The sample size was calculated based on Walter and colleagues' article[15]. We used a significance level of 5% and a statistical power of 80%. The acceptable intraclass correlation coefficient (ICC) was set as 0.60 and was estimated to be 0.79. Also, we added 10% for possible dropouts and refusals. The required sample size was estimated in 50 individuals. The sampling process enrolled individuals of different age and sex categories in order to guarantee heterogeneity. Eligibility of the participants demanded living in Arroio do Padre, agreement to answer the questionnaire twice (seven days apart) and wear an accelerometer for one week. The study was conducted between September and October of 2015.

Reliability was assessed by comparing the results from the first and second applications of GPAQ, carried out in a seven-days interval. The GPAQ was applied in a face-to-face manner by a trained interviewer in the participant's household. If the participant did not answer one of the questions of the GPAQ, the interview was considered invalid. GPAQ takes into consideration the frequency, duration and intensity of physical activity, and collects data from three different domains (work/domestic, transport and leisure-time), as well as sedentary time. To assess total time spent in physical activity during the week, the duration (minutes per day) is multiplied by the frequency (number of days practiced per week) – vigorous activities are further multiplied by two – and then is totalized (the same can be made for each domain). In order to assess this information, the respondent has to recall about activities that lasted for more than ten minutes, which are performed during an usual week.

After the first interview, individuals were invited to use an accelerometer during the following week. The concurrent validity was assessed by comparing the results from accelerometry and the first GPAQ interview. Accelerometers (wGT3x+ ActiGraph) were used to objectively assess physical activity. The device was placed in the right side of the hip, the epoch was set at five seconds and the sample frequency was 30Hz. Participants received the accelerometer on Saturdays, and data were analyzed from Sundays onwards, up to the next Saturday. Individuals were advised to wear the accelerometer 24 hours per day for six consecutive days. The individuals that reported non-use of the accelerometer or that had 0 minutes of total physical activity during the week, were excluded from the analyses. The Sasaki et al[16] cut-points were applied for defining different intensities of physical activity. Moderate-intensity was set from 2690 to 6166 cpm, vigorous-intensity from 6167 to 9642, and very vigorous-intensity 9643+ cpm. Initialization, download and accelerometer data analyses were made in Actilife 6. For the analyses, we combined the time spent in moderate-, vigorous- and very vigorous-intensity physical activity, generating a moderate to vigorous physical activity (MVPA) score.

Reliability was evaluated using ICC and the Bland and Altman plots[17]. We present ICC results by sex, age and socioeconomic level, as well as stratified by domain. Spearman's correlation and Bland and Altman plots were used in the concurrent validity analyses. The Spearman correlation coefficient was interpreted according to the rule of thumb presented in Hinkle et al[18]. All analyses were performed in Stata 12.0.

The Ethics Committee of the Physical Education School of the Federal University of Pelotas approved the present study. Written informed consent was obtained from all participants prior to data collection.

## Results

Fifty-five individuals were invited to participate in the study, of whom 49 (89.1%) had valid answers to GPAQ and 44 (80.0%) had valid accelerometry measures. Those included in the analyses were similar to those excluded in terms of sex and age. Socio-demographic characteristics of the study participants are presented in Table 1. Most participants (55.1%) were men, around half were aged older than 40 years, and almost half (44.9%) were in the intermediate socioeconomic category.

**Table 1 –** Socio-demographic characteristics of the study participants and reliability analysis. Arroio do Padre, Brazil, 2015.

| Variable | p[a] | ICC (95% CI) |
|---|---|---|
| All | | |
| Sex | 49 (100.0) | 0.78 (0.63–0.87) |
| Male | 27 (55.1) | 0.67 (0.39–0.83) |
| Female | 22 (44.9) | 0.90 (0.78–0.96) |
| Age (years) | | |
| 18-40 | 24 (49.0) | 0.79 (0.58–0.90) |
| 41+ | 25 (51.0) | 0.75 (0.52–0.88) |
| Socioeconomic position | | |
| B (richest) | 15 (30.6) | 0.85 (0.62–0.95) |
| C | 22 (44.9) | 0.80 (0.58–0.91) |
| D and E (poorest) | 12 (24.5) | 0.67 (0.20–0.89) |

a = 49 participants with valid GPAQ information.

To measure GPAQ's reliability, we compared two administrations of the questionnaire with one-week interval. We found high reliability for the total minutes spent in physical activity (ICC = 0.78; 95%CI: 0.63 – 0.87), as shown in Table 1. Stratified analyses confirmed that the reliability was high in all age groups. Reliability was higher among women (ICC = 0.90; 95%CI: 0.78 – 0.96) than among men (ICC = 0.67; 95%CI: 0.39 – 0.83). Furthermore, ICC values tended to be higher among the richer as compared to poorer participants. We also analyzed the ICC according to physical activity domains, as presented in Figure 1. The lowest ICC value occurred in the transport-related domain (ICC = 0.34; 95%CI: 0.07 – 0.56), whereas the highest ICC (ICC = 0.76; 95%CI: 0.62 – 0.86) was observed in the work domain.

We also visually compared total minutes of physical activity between both GPAQ measures, from the first and second weeks through the Bland and Altman's plot present in Figure 2A. The average difference was 192 minutes (95%CI: 69.63; 453.75), representing on average higher values reported in the first interview compared to the second one. The distribution of the dots along the X-axis indicates random variability and wider differences when participants reported more physical activity practice (higher than 400 minutes per week).

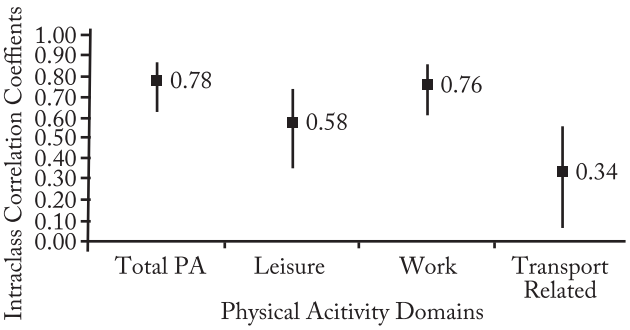


**Figure 1 –** Intraclass correlation coefficient for the total physical activity and divided by domains (Arroio do Padre, Brazil, 2015).

To analyze concurrent validity between the accelerometer and the GPAQ interview we used correlation test and Bland and Altman graph. We considered in these analyses the 43 individuals who wore the accelerometer during one week and provided valid answers for the GPAQ applied on the first week. There was a moderate correlation between accelerometry and GPAQ. Spearman's correlation coefficient comparing the accelerometer to the first week interview was 0.5. In Figure 2B, it is possible to perceive limited agreement between the accelerometer and the GPAQ from the first week, where the average difference between the measurements was -1043 minutes (95%CI: -1452.23; -634.09). This happens since the self-reported physical activity is systematically higher than the objectively measured physical activity. Wider differences are found among participants who reported 800 minutes or more (marked by the vertical line in Figure 2B).

## Discussion

This study measured reliability and concurrent validity of the GPAQ instrument in rural area. Reliability between both interviews seemed higher among women, younger adults, and there were a positive trend for socioeconomic status (the richer had higher reliability than the poorer). The reliability of GPAQ is apparently good, but the validity comparing estimates from accelerometry and GPAQ was limited.

Misra and colleagues compared reliability of GPAQ in India between peri-urban and rural populations with an interval of one month. They found similar ICC for work and leisure-time, compared to our study (0.67 vs. 0.76, and 0.68 vs. 0.58, respectively). However, ICC for

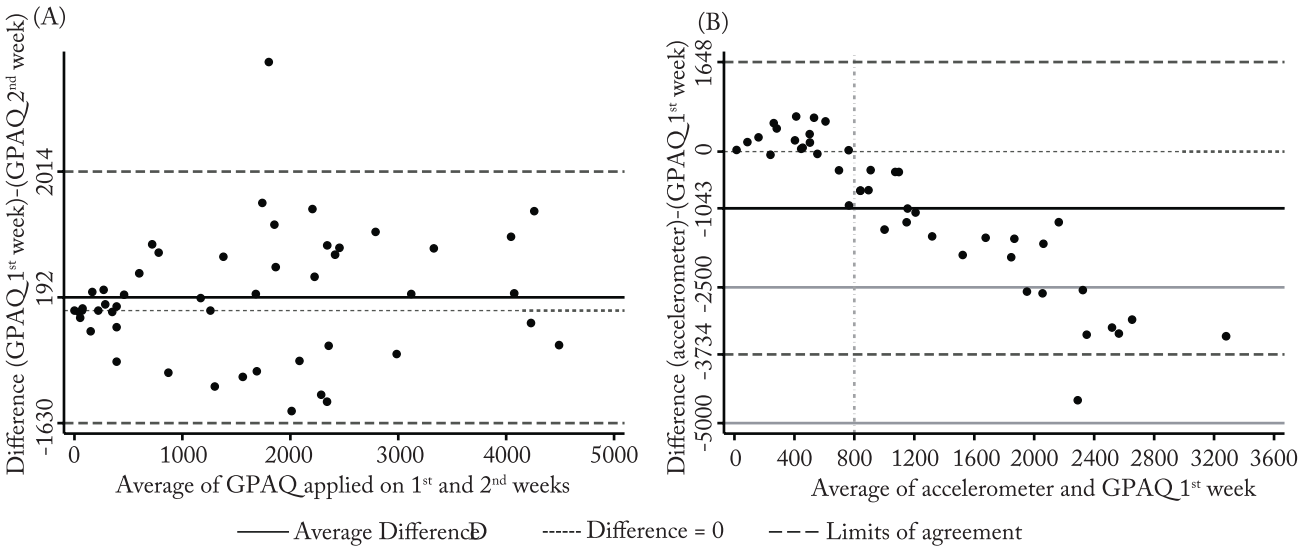


**Figure 2 –** Bland and Altman's plots analyzing the minutes per week agreement between (A) GPAQ's applied with a one-week interval, and (B) accelerometer and GPAQ applied on the first week (Arroio do Padre, Brazil, 2015).

transport domain was higher among their sample compared to ours (0.72 vs. 0.34)[13]. The difference between ICC values for transport domain may be due to the composition of the sample. In Misra's study, half of the sample was living in a peri-urban area, which could comprise better-educated individuals compared to an exclusively rural sample. Thus, the probability of misunderstanding the questionnaire could be greater in our study.

Bland-Altman's agreement between accelerometer and GPAQ had mean difference of -1043 minutes. Once again, this difference is mostly due to those participants reporting high physical activity levels. A study in Bangladesh found a similar result to ours, which shows a pattern of increased error with the increase of the physical activity average, given that GPAQ overestimates the accelerometer in rural residents[19]. Another study performed in Singapore with adults 21 years or older using the same accelerometer showed a similar Bland and Altman diagram in terms of dots distribution along the X-axis, although their confidence interval of the average difference was much shorter (-138.7 to 210.4). For that, there are three possible explanations: first, their sample was with students or working adults from different departments of an university, which mean that they probably have more years of study than our sample. The other explanation can be that their sample size was greater than ours (113 vs. 48 individuals)[20]. Finally, our population is rural and, as they work several hours per day, they tend to overestimate work-domain physical activity. It is important to notice that, although researchers still compare these two methods as being the same metric, they are not equivalent. Motion devices measures short time physical activity, detecting more specific details than self-reported data, which focused on more continuous physical activities practice[5].

Our study has some limitations. First, while the questionnaire measured seven days of physical activity, the accelerometer measured six days due to logistics limitations of the study. To minimize the limitation, we set the device to objective measure the individual's physical activity in five weekdays and in one weekend day to capture the different acceleration patterns during the days of the week. However, we also carried out sensitivity analyses based on the average time spent in physical activity per day and the results were in the same direction with a slightly lower difference compared to the total week time analysis. Second, the accelerometer measured the last week of physical activity while the questionnaire measured a regular week. During the interviews, the interviewees stated that they had a well-defined routine during the week. Thus, we believe that a regular week (as measured by the questionnaire) would not be different than the last week (as measured by the accelerometer). Third, different interviewers could perform the interviews of the first and second weeks. The interviewers were trained and standardized to perform the questionnaire, reducing the possibility of interviewer bias. However, the participants could provide more social desirable answers (overestimate the physical activity practice) depending on the characteristics of the interviewer applying the

questionnaire. But, the high ICC values of our reliability analyses reduce this possibility.

Even though important results are shown in this study and the fieldwork experience has highlighted many problems about the questionnaire. GPAQ has difficult components to be understood in populations with low educational levels. There are details and specificities that might confuse the participants such as the total time that need to be reported, the intensity differences and the need to differentiate domains of physical activity. The questionnaire often seemed not well comprehensive in both applications, and this may be the main source of inconsistences. Despite that, GPAQ was more reliable when considering the work-domain and for the ones that reported less than 400 minutes of physical activity, in other words, the error is greater for the very active individuals. Thus, in terms of public health, this questionnaire could be used in large-scale studies to investigate work-domain physical activity or the prevalence of physical activity according to the percentage of people that achieve the recommendations of physical activity. This is feasible as the recommendations consider a cutoff value of 150 minutes per week, and even though the GPAQ physical activity levels are overestimated from the 400 minutes on, the individuals would still be correctly classified as actives.

Our study identified a good reliability between the two applications of the GPAQ questionnaire, and a limited agreement comparing GPAQ and accelerometer measures. The questionnaire could be used to assess total physical activity among rural populations if supported by other instrument[21], but would be better suited to identify active and inactive individuals. Also, it could be used to study leisure- and work-related physical activity, but should be used with care when considering the transport domain. However, it is important to highlight the need for a new questionnaire that is easier, simpler and more practical to be applied in rural areas.

## Conflicts of interest

The authors declared no conflict of interest.

## Author's contribution

Martins RC and Blumenberg C, conducted the fieldwork, did the analysis and wrote the manuscript. Silva ICM, helped in the interpretation of the data and revised the article.

## Acknowledgements

The authors thank Prof. Pedro Curi Hallal, from Federal University of Pelotas, for his assistance on designing the fieldwork.

# References

1. Morris JN, Heady JA, Raffle PA, Roberts CG, Parks JW. Coronary heart-disease and physical activity of work. Lancet. 1953;262(6796):1111–20.
2. Paffenbarger RS Jr, Laughlin ME, Gima AS, Black RA. Work activity of longshoremen as related to death from coronary heart disease and stroke. N Engl J Med. 1970;282(20):1109–14.
3. Warburton DER. Health benefits of physical activity: the evidence. CMAJ. 2006;174(6):801-9.
4. Lee IM, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT, et al. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. Lancet. 2012;380(9838):219–29.
5. Troiano RP, McClain JJ, Brychta RJ, Chen KY. Evolution of accelerometer methods for physical activity research. Br J Sports Med. 2014;48(13):1019–23.
6. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc. 2003;35(8):1381–95.
7. Hallal PC, Gomez LF, Parra DC, Lobelo F, Mosquera J, Florindo AA, et al. Lessons learned after 10 years of IPAQ use in Brazil and Colombia. J Phys Act Health. 2010;7Suppl2:S259–64.
8. Guthold R, Louazani SA, Riley LM, Cowan MJ, Bovet P, Damasceno A, et al. Physical activity in 22 African countries: results from the World Health Organization STEPwise approach to chronic disease risk factor surveillance. Am J Prev Med. 2011;41(1):52–60.
9. Bui TV, Blizzard CL, Luong KN, Truong Nle V, Tran BQ, Otahal P, et al. Physical Activity in Vietnam: Estimates and Measurement Issues. PloS One. 2015;10(10):1-14.
10. Moniruzzaman M, Mostafa Zaman M, Islalm MS, Ahasan HA, Kabir H, Yasmin R. Physical activity levels in Bangladeshi adults: results from STEPS survey 2010. Public Health. 2016;137:131–8.
11. Bull FC, Maslin TS, Armstrong T. Global physical activity questionnaire (GPAQ): nine country reliability and validity study. J Phys Act Health. 2009;6(6):790–804.
12. Anjana RM, Sudha V, Lakshmipriya N, Subhashini S, Pradeepa R, Geetha L, et al. Reliability and validity of a new physical activity questionnaire for India. Int J Behav Nutr Phys Act. 2015;12(40):1-12.
13. Misra P, Upadhyay RP, Krishnan A, Sharma N, Kapoor SK. A community based study to test the reliability and validity of physical activity measurement techniques. Int J Prev Med. 2014;5(8):952–9.
14. IBGE. Classificação e caracterização dos espaços rurais e urbanos do Brasil – uma primeira aproximação. Rio de Janeiro: IBGE, 2017. Acesso em: 25 de novembro de 2017.
15. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. Stat Med. 1998;17(1):101–10.
16. Sasaki JE, John D, Freedson PS. Validation and comparison of ActiGraph activity monitors. J Sci Med Sport. 2011;14(5):411–16.

**17.** Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res. 1999;8(2):135–60.

**18.** Hinkle DE, Wiersma W, Jurs SG. Applied statistics for the behavioral sciences. Boston: Houghton Mifflin, 2003.

**19.** Mumu SJ, Ali L, Barnett A, Merom D. Validity of the global physical activity questionnaire (GPAQ) in Bangladesh. BMC Public Health. 2017;17(650):1-10.

**20.** Chu AH, Ng SH, Koh D, Müller-Riemenschneider F. Reliability and Validity of the Self- and Interviewer-Administered Versions of the Global Physical Activity Questionnaire (GPAQ). PloS One. 2015;10(9):1-18.

**21.** Kirunda BE. Body weight and physical activity of adults in rural Uganda. [PhD thesis]. Bergen: University of Bergen; 2017.

Recebido: 25/10/2017
Aprovado: 13/08/2018

**Quote this article as:**

*Martins RC, Blumenberg C, Silva ICM. Reliability and concurrent validity of the Global Physical Activity Questionnaire in adults from a Brazilian rural area. Rev Bras Ati Fis Saúde. 2018;23:e0017. DOI: 10.12820/rbafs.23e0017*