

Validity and reliability of a telephone survey of physical activity in Brazil

Validade e reprodutibilidade de inquérito telefônico de atividade física no Brasil

Alexandra Dias Moreira¹, Rafael Moreira Claro¹,
Mariana Santos Felisbino-Mendes¹, Gustavo Velasquez-Melendez¹

ABSTRACT: *Introduction:* Physical activity in Brazil is assessed by the Telephone-based Risk Factor Surveillance System for Chronic Diseases (VIGITEL). *Objectives:* To evaluate the reliability and concurrent validity of VIGITEL physical activity indicators. *Methods:* For the reliability study, 305 individuals responded to VIGITEL interviews, which were repeated with intervals of 7–15 days, in Belo Horizonte, 2013. The evaluated indicators included “sufficiently active on leisure time,” “active in transportation,” “inactive in four domains of physical activity (leisure, work, transportation, and housework),” and “watching TV for long periods.” Kappa coefficient (k) was used to measure agreement between both interviews. For concurrent validity assessment, the same subjects also responded to the Global Physical Activity Questionnaire (GPAQ), used as reference method for comparison of VIGITEL indicators. Comparison was assessed by measures of sensitivity, specificity, positive (PPV), and negative predictive values (NPV). *Results:* Reliability study showed substantial agreement for active individuals in leisure time ($k = 0.70$) and inactive individuals ($k = 0.64$). The agreement of watching TV for long periods was moderate ($k = 0.56$) and the activeness in transportation showed fair agreement ($k = 0.35$). In comparison to the reference method, sensitivity ranged from 54.8 to 67.7 in frequency of inactive, active in leisure time and work. Transportation domain was represented by 11.9 of sensitivity. Specificity ranged from 72.0 to 91.2 among four domains of physical activity and inactive. *Conclusion:* Physical activity questionnaire used by the surveillance system seems to be reliable in all domains, except in transportation questions. VIGITEL was considered comparable to GPAQ in most aspects of physical activity evaluation.

Keywords: Motor activity. Reproducibility of results. Validity of tests. Questionnaires. Efficiency. Epidemiology.

¹School of Nursing of the Maternal Child and Public Health Department of *Universidade Federal de Minas Gerais* – Belo Horizonte (MG), Brazil.

¹School of Nursing, Department of Nutrition of *Universidade Federal de Minas Gerais* – Belo Horizonte (MG), Brazil.

Corresponding author: Gustavo Velasquez-Melendez. Escola de Enfermagem do Departamento de Enfermagem Materno-Infantil e Saúde Pública da Universidade Federal de Minas Gerais. Avenida Alfredo Balena, 190, Santa Efigênia. CEP: 30130-100. Belo Horizonte, MG, Brasil. E-mail: jguveme@ufmg.br

Conflict of interests: nothing to declare – **Financial support:** Ministry of Health (Agreement 188).

RESUMO: *Introdução:* A atividade física no Brasil é avaliada pelo Sistema de Vigilância de Fatores de Risco e Proteção para Doenças Crônicas por Inquérito Telefônico (VIGITEL). *Objetivo:* Avaliar a reprodutibilidade e a validade dos indicadores de atividade física do VIGITEL. *Métodos:* Para o estudo de reprodutibilidade, 305 indivíduos em Belo Horizonte responderam a entrevistas do VIGITEL repetidas com intervalos de 7 a 15 dias, em 2013. Os indicadores avaliados foram “suficientemente ativos no lazer”, “ativos no deslocamento”, “inativos em quatro domínios da atividade física” (lazer, trabalho, transporte e atividades domésticas) e “assistir TV por longos períodos”. O Coeficiente Kappa (k) foi utilizado para medir concordância entre as entrevistas. Para a validade, os participantes responderam também o Questionário Global de Atividade Física (GPAQ), método de referência para a comparação dos indicadores VIGITEL. A comparação foi avaliada por sensibilidade, especificidade, valores preditivos positivos (VPP) e negativos (VPN). *Resultados:* A reprodutibilidade mostrou concordância substancial para os indivíduos ativos no lazer (k = 0,70) e inativos (k = 0,64). A concordância do hábito de assistir TV foi moderada (k = 0,56) e o deslocamento mostrou concordância regular (k = 0,35). No estudo de validade, a sensibilidade variou de 54,8 a 67,7 na frequência de inativos e ativos no lazer, respectivamente. O deslocamento foi representado por 11,9 de sensibilidade. A especificidade variou de 72,0 a 91,2 nos quatro domínios da atividade física e inativos. *Conclusão:* O questionário de atividade física utilizado pelo Sistema de Vigilância parece ser confiável em todos os domínios, exceto nas questões de deslocamento. O VIGITEL foi comparável ao GPAQ na maioria dos aspectos da atividade física. *Palavras-chave:* Atividade motora. Reprodutibilidade dos testes. Validade dos testes. Questionários. Eficiência. Epidemiologia.

INTRODUCTION

Physical inactivity stands out as one of the main modifiable risk factors for the occurrence of Noncommunicable diseases (NCDs)¹. That said, high frequencies of insufficient physical activity are observed in both developed and developing countries². In Brazil, approximately 1 in 2 adults (49.4%) does not reach the recommended level of physical activity (PA)².

The World Health Organization (WHO) defines as healthy the weekly practice of at least 150 minutes of activity of moderate intensity or the equivalent, such as 75 minutes of intense activities or even an equivalent combination of both³. The volume and frequency of physical activity of individuals and populations is often measured using self-reported questionnaires^{4,5}. In this regard, the WHO developed the Global Physical Activity Questionnaire (GPAQ) in 2002 in order to allow a wide assessment of PA in several domains that could be used internationally^{6,7}. Its development took into account cultural differences as well as sub-populations, considering people with low socioeconomic status and illiterate individuals. This instrument was translated into several languages and presented satisfactory reliability in several studies carried out up to the present moment^{5,7}. This tool has been validated for the adult Brazilian population⁵, and its use is present in several studies at the international level⁶. It is often applied through face-to-face mediation of an interviewer or even through telephone interviews⁸.

In Brazil, official and systematic PA surveillance data are available since 2006. These data were collected by the Surveillance System of Risk and Protective Factors for Chronic Diseases by Telephone Survey (VIGITEL) and is based on non-face-to-face interviews of self-reported data. However, the nature of VIGITEL (a questionnaire carried out by telephone that addresses all major risk and protective factors for NCDs), prevents the full use of instruments such as the GPAQ because of the need to use simple and short questions with objective and direct responses⁹. Thus, the VIGITEL assessment of PA is made through a set of questions adapted from surveys already in existence at the time of conception (such as the Behavioral Risk Factor Surveillance System of the Centers for Disease Control and Prevention – BRFSS/CDC), and questionnaires used globally in the area (such as the GPAQ and the International Physical Activity Questionnaire – IPAQ). This created the need for validation studies of this questionnaire and analyses of its reliability. In a study carried out in the city of São Paulo in 2005, the validity and reliability of the VIGITEL physical activity indicators were investigated using the information obtained in three 24-hour recalls (R24H)¹⁰. On that occasion, satisfactory results were obtained for both validity and reliability of the indicators. However, it has subsequently been found that at least four days of R24H are required to estimate one week of physical activity (three days during the week and one day on the weekend). Furthermore, little progress has been made since these findings, so important issues remain unknown such as the comparability of VIGITEL with international questionnaires.

Therefore, the objective of this study is to verify the reliability and validity of the VIGITEL PA indicators using more advanced and accurate methods.

METHODS

POPULATION STUDY AND DATA COLLECTION

This is a cross-sectional epidemiological study with a diagnostic approach carried out in the city of Belo Horizonte, Minas Gerais. For the sample calculation, the expected kappa value for the study was 0.9, with an expected 50% positive proportion for first and second observation, 5% absolute precision, and 95% confidence level. A sample of 300 individuals was defined using Stata 9.1 software. Initially, a random subsample of the individuals studied by VIGITEL in the city of Belo Horizonte between March and April was defined and established as initial contact (baseline) alongside study participants (sample replicates 1, 2, 3, and 4; further details on the sampling procedure of VIGITEL are published in the system's annual report⁹). Thus, 418 individuals were included in order to allow the desired minimum sample to be reached even with total refusals close to 30%.

A total of 305 people accepted to participate in the study and were included in the subsequent data collection steps for the reliability and validity study. Participants were contacted again by telephone 7–15 days after they had answered the original VIGITEL interview and

were asked to respond again to the block of PA questions in this survey (reliability study) and version 2 of the GPAQ questionnaire, the reference method for the validity study. The interviews for this study were guaranteed in such a way that the second interviewer was always different from the one who had conducted the initial interview.

The GPAQ was applied according to recommendations of the Global Physical Activity Questionnaire Analyses Guide set by the WHO¹². Interview simulations were performed so that possible errors were identified prior to the start of the study. Moreover, during data collection, periodic quality control was performed through recordings of randomly chosen interviews.

ORGANIZATION OF DATA AND PREPARATION OF INDICATORS OF INTEREST

The VIGITEL questionnaire includes information on PA carried out in four domains: leisure time, occupational, commuting, and household activities. For the reliability study all the VIGITEL indicators were used for the comparison between the second and initial interview of the system⁹.

Subjects who practiced at least 150 minutes of moderately intense PA or at least 75 weekly minutes of intense PA were considered sufficiently active in their leisure time. The activities classified as moderately intense consisted of walking, treadmill walking, bodybuilding, water aerobics, general gymnastics, swimming, martial arts, cycling, and volleyball. The physical activities considered intense were running, running on treadmills, aerobics, soccer, basketball, and tennis⁹. It is worth mentioning that in VIGITEL there are no questions about the individual perception of the intensity of PA according to the modalities considered⁹.

Individuals who usually commuted to work or school via walking or cycling and spent 30 minutes or more daily on the round-trip route were considered active on the commuting PA⁹.

Individuals who were considered physically inactive were those who did not engage in physical activity in their leisure time during the least three months and who did not make intense physical efforts at work, nor engaged in PA in their daily commute to work, which is to say they did not walk or cycle for a minimum of 10 minutes per trip daily, and were not responsible for the heavy cleaning of their houses⁹.

In addition, the study demonstrated the frequency of adults who have the habit of watching television for three hours or more per day. This indicator associates the habit of watching television with sedentary behavior, which in turn is related to the risk of developing chronic pathologies such as obesity, type II diabetes, cardiovascular diseases, and metabolic syndrome, regardless of the recommended practice of physical activity⁹.

The GPAQ includes 16 questions with information about physical activities held in three domains: leisure time (sports, recreational exercise), commuting (walking or cycling), and occupational (paid or voluntary work, and domestic activities), in addition to information regarding sedentary behavior (sitting time, reading, watching TV, talking with friends, traveling). PAs are evaluated according to frequency and duration in the last week and classified

as moderate or vigorous. These questions allowed for the elaboration of three of the four PA indicators used by the VIGITEL system, considering the duration, frequency, and intensity of activities: “physical activity in leisure time” (sufficiently active are individuals who practice at least 150 minutes of moderately intense PA per week or at least 75 minutes per week of intense PA), “commuting physical activity” (sufficiently active are individuals who commuted from one place to another by bicycle or walking and who spend 30 minutes or more on daily commute), and “inactivity” (inactive are individuals who did not engage in physical activity in their leisure time and who did not perform physical exertion at work, and did not commute by foot or bicycle). The “television habit” indicator has not been elaborated, since the question about sedentary behavior in the GPAQ is related to several moments in which the individual remains seated during the week, and not exclusively to the habit of watching television.

DATA ANALYSIS

For the reliability study, the differences between the proportions of active individuals during leisure, commuting, inactive individuals, and individuals with a habit of watching TV for 3 hours or more in the original and repeated interviews were tested using the McNemar test, considering significance $<5\%$.

The Kappa statistic was employed to determine the agreement between the results of the initial interview (VIGITEL) and the repeated interview. The reference values were 0.80 as almost perfect agreement; between 0.61 and 0.80 for substantial agreement; between 0.41 and 0.60 for moderate agreement; between 0.21 and 0.40 fair agreement; and below 0.21 for slight agreement¹³. The agreement between all four PA indicators present in VIGITEL was analyzed using this procedure.

The Bland-Altman method was then used to investigate the relationship between measurement errors and actual values. As the real value is not known, the mean among the measurements was assumed to be the best estimate of reality¹⁴. In this analysis, we considered the total number of minutes spent in one week of leisure time PA, and the graph was constructed from the difference between the repeated and original interview in minutes.

The GPAQ was used as a reference method for the concurrent validation of the PA question block of VIGITEL system. Leisure-time, commuting PA, and inactivity indicators elaborated from the second interview of VIGITEL were compared to those indicators obtained through the interview with the GPAQ. The comparative analyses between the questionnaires using the GPAQ as reference were performed by calculating sensitivity (S), specificity (E), positive predictive value (PPV), and negative predictive value (NPV), in addition to the area under receiver operating characteristic (ROC) curve.

The calculations obtained through the GPAQ were in accordance with the WHO's Global Physical Activity Questionnaire Analyses Guide¹². Three individuals with a period longer

than 16 hours per day of physical activity were excluded. The final sample for the validity study was, therefore, 302 individuals.

The present study was evaluated and approved by the Research Ethics Committee of the *Universidade Federal de Minas Gerais*, under opinion number 325.242 (CAAE: 15689713.3.00005149) and there is no conflict of interest. Informed consent was obtained at the time of telephone contact, and in the original VIGITEL interviews. Participants were assured of secrecy about the information collected and offered freedom of withdrawal at any stage of the research.

RESULTS

The sample obtained was composed of 305 individuals for the reliability study. Of the participants, 58% were women and 42% were men. The mean age for both studies was 49.7 years (standard deviation – SD = 18.2), and approximately half of the respondents reported being legally married. For the validity study, the final sample consisted of 302 individuals whose characteristics were similar to those of the participants in reliability study.

Regarding the comparison between original and repeated interview of VIGITEL, frequency of PA was 37.0% and 32.8%, respectively. Kappa coefficient (k) was 0.70, indicating substantial agreement. Regarding the “commuting PA” indicator, we found frequencies of physically active individuals of 11.5% and 10.5% for original and repeated interviews, respectively, with a k of 0.35, indicating fair agreement. In addition, 14.7% of sample participants were considered physically inactive in the original interview and 16.4% in the repeated interview. There was substantial agreement for the inactivity indicator, with a value for k of 0.64. Regarding the “habit of watching television”, we found a proportion of people with this habit of 30.2% and 35.1%, respectively, for original and repeated interviews, with moderate agreement. No statistically significant difference was found in the comparison of proportions between the interviews after McNemar’s test (Table 1).

Differences between the total minutes of leisure-time PA in the two interviews (repeated interview – original interview) in relation to their mean are generally close to the mean (Graph 1). There were no sub or overestimation trends of the self-reported variables for this indicator.

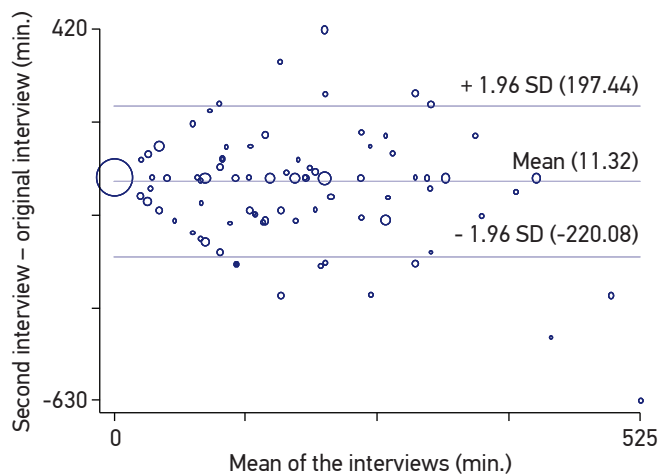
Regarding the validity analyses (Table 2), there were no significant differences in the frequencies obtained through VIGITEL (32.8%) compared to the GPAQ (30.8%) ($p = 0.538$). Satisfactory sensitivity and specificity values were obtained (67.7% and 82.8%, respectively), in addition to an area under ROC curve of 0.75, a PPV of 63.3%, and VPN of 85.2%.

Considering the commuting PA indicator, there was a significant difference between the frequencies obtained by the VIGITEL questionnaire and the GPAQ ($p < 0.001$). Only 9.9% of the studied population was considered physically active in this indicator by the VIGITEL system, whereas 36.1% was considered physically active by the GPAQ, evidencing a large underestimation of this measure by the first questionnaire for both sexes. Consequently,

Table 1. Reliability of indicators of physical activity and sedentary lifestyle in adults. Data obtained by telephone interview. Belo Horizonte, MG, 2013. Sample of 305 individuals.

Indicator	Original Telephone Interview	Repeated Telephone Interview	p-value*	k (95%CI)
	n (%)	n (%)		
Activity during leisure time				
Male	58 (45.3)	53 (41.1)	0.424	0.60 (0.46 – 0.74)
Female	55 (31.1)	47 (26.5)	0.076	0.78 (0.67 – 0.88)
Total	113 (37.0)	100 (32.8)	0.059	0.70 (0.62 – 0.78)
Activity during commuting				
Male	8 (6.2)	9 (7.0)	1.000	0.31 (0.01 – 0.61)
Female	27 (15.2)	23 (12.3)	0.571	0.35 (0.16 – 0.54)
Total	35 (11.5)	32 (10.5)	0.749	0.35 (0.19 – 0.51)
Inactive				
Male	21 (16.4)	20 (15.6)	1.000	0.56 (0.37 – 0.76)
Female	24 (13.6)	30 (16.9)	0.179	0.69 (0.55 – 0.84)
Total	45 (14.7)	50 (16.4)	0.458	0.64 (0.52 – 0.76)
Habit of watching television				
Male	95 (25.8)	39 (30.5)	0.263	0.61 (0.46 – 0.76)
Female	59 (33.3)	68 (38.4)	0.199	0.52 (0.39 – 0.65)
Total	154 (30.2)	107 (35.1)	0.067	0.56 (0.46 – 0.66)

*p-values for McNemar test; k: Kappa coefficient; 95%CI: 95% confidence interval.



Graph 1. Differences in time of physical activity in leisure time obtained in the repeated and original interview¹. Belo Horizonte, MG, 2013.

the sensitivity value was only 11.9%. The specificity value was satisfactory, reaching 91.2%, with PPV of 43.3% and NPV of 64.7%. The area under the ROC curve was 0.52.

Finally, for the physically inactive indicator in the four PA domains, there was an over-estimation of the VIGITEL system in relation to the GPAQ ($p = 0.007$), with frequencies of 16.6% and 10.3%, respectively, obtained for the total sample of the study. This made the sensitivity slightly lower (54.8%) in relation to the leisure-time domain of PA. The specificity was high (87.8%) and the area under the ROC curve was 0.71 with PPV of 34.0% and NPV of 94.4%.

Table 2. Frequency, sensitivity, specificity, and positive and negative predictive value of physical activity indicators in leisure, physical activity in commuting and inactivity. Belo Horizonte, MG, 2013. Sample of 302 individuals.

Indicator	Second VIGITEL Interview		GPAQ		S	E	AUC (95%CI)	PPV	NPV
	n (%)	95%CI	n (%)	95%CI					
Physical Activity during leisure time									
Men	53 (42.1)	33.6 – 51.1	51 (40.5)	31.3 – 49.0	78.0	81.3	0.80 (0.72 – 0.87)	73.6	84.7
Women	46 (26.1)	19.5 – 32.5	42 (23.9)	17.9 – 30.7	55.8	83.6	0.70 (0.62 – 0.78)	52.2	85.5
Total	99 (32.8)	27.4 – 38.1	93 (30.8)	25.5 – 36.0	67.7	82.8	0.75 (0.70 – 0.81)	63.6	85.2
Physical activity during commuting*									
Men	8 (6.4)	2.04 – 10.7	46 (36.5)	28.2 – 45.3	8.7	94.9	0.52 (0.47 – 0.57)	50.0	64.1
Women	20 (12.5)	7.52 – 17.3	63 (35.8)	28.4 – 43.7	14.3	88.6	0.51 (0.46 – 0.58)	40.9	65.2
Total	30 (9.9)	6.54 – 13.3	109 (36.1)	30.6 – 41.5	11.9	91.2	0.52 (0.48 – 0.55)	43.3	64.7
Inactivity*									
Men	20 (16.0)	9.5 – 22.5	13 (10.3)	4.97 – 15.8	30.8	85.7	0.58 (0.45 – 0.72)	20.0	91.4
Women	30 (17.0)	11.4 – 22.5	18 (10.2)	4.97 – 15.8	72.2	89.3	0.81 (0.70 – 0.92)	43.3	96.6
Total	50 (16.5)	12.3 – 20.8	31 (10.3)	5.67 – 14.7	54.8	87.8	0.71 (0.62 – 0.80)	34.0	94.4

*p-value for McNemar test (<0.05); n: sample number; 95%CI: 95% confidence interval; GPAQ: Global Questionnaire of Physical Activity; S: sensitivity; E: specificity; AUC: area under ROC curve; PPV: positive predictive value; NPV: negative predictive value.

DISCUSSION

The consolidation of VIGITEL as a central instrument for the monitoring of risk and protective factors for chronic noncommunicable diseases in the country¹⁵ made it possible to carry out studies on the reliability and validity of the system generated data. Therefore, the results of the present study expose the quality of information regarding the practice of PA as assessed by VIGITEL through comparison with the GPAQ, a questionnaire validated and recommended by the WHO. With respect to the reliability of the indicators, the results were acceptable both in the comparison of the population frequencies (near frequencies between the indicators) and in the intraindividual analysis (k varying from moderate-to-substantial agreement) for the indicators investigated – “inactivity,” “habit of seeing television,” and “PA in leisure time”. Similarly, the comparison of the VIGITEL data with those of the GPAQ reference method presented acceptable results for leisure-time PA and inactivity; however, for commuting, a worse performance was observed in VIGITEL. The indicator that obtained the best performance in the reliability between the interviews was the PA during leisure time. This is an important finding, since this is the most important domain of practice for the prevention of NCDs and reduction of mortality in general¹⁶⁻¹⁸. Similar results were found in other studies^{5,10,19}. The indicators “habit of watching television” and “inactivity” had close approximations to that of other reliability studies^{5,10,20}. Our findings expand the understanding about the theme identified in these studies, since it is based on an unpublished methodology and addresses not only the practice indicator of PA, but also its dimensions (such as duration of practice).

The reliability presented by most indicators suggests adequate standardization of the interviews and that the participants understand the questions well, providing similar answers at different moments of measurement. These findings favor the implementation of public policies based on population data provided by the VIGITEL PA questionnaire, as well as other PA studies based on the use of this instrument, with the exception of the “commuting PA” indicator. This situation can be partially explained by the complexity of the commutes in the state capitals or even by the fragility of the questions used by VIGITEL to measure this indicator, which is observed in studies with similar results^{5,19,21}. It is known that the section of the VIGITEL questionnaire aimed at measuring PA in commuting was changed in 2009, aiming at the best performance of this indicator (without its validity being tested after this fact). Although it is not possible to compare our results with those obtained before this change, the need for a new revision of the questionnaire remains explicit.

When compared to the GPAQ, VIGITEL showed acceptable performance in most analyses. High specificities and negative predictive values were observed for the PA indicators in leisure time and in commuting, demonstrating the effectiveness of the system in detecting individuals who do not practice PA in these domains. Furthermore, the sensitivity value of this study for leisure-time PA was satisfactory, similar or superior to that of other validation studies^{10,22}, demonstrating that VIGITEL performs adequately when compared to other instruments.

At the other extreme, a low sensitivity value and large difference in frequency were obtained in the PA domain of commuting, with the VIGITEL questionnaire being underestimated when compared with the GPAQ. This unsatisfactory result can be explained by the differences in the composition of the questions to scale the commuting indicator. While the VIGITEL system considers only commute to/from school and work, the GPAQ considers all trips made in one day. Therefore, given the underestimation of the real prevalence of physically active individuals owing to commuting in the population, monitoring may be inefficient in this regard. In addition, this study also showed low reliability related to this indicator, reinforcing the weaknesses of the VIGITEL questions on this domain.

The main limitations of the validity study are the differences in the questions related to the domains of PA between VIGITEL and the GPAQ, which hindered the comparability. Nevertheless, most of the performance tests were satisfactory and similar to those found in many validation studies in the literature. Moreover, the validation of a questionnaire which uses as a reference another questionnaire that is also based on self-reporting can lead to common errors in both surveys, since both are subjective measures. However, GPAQ is known as an internationally validated questionnaire for use in developing countries and is used by many researchers to evaluate PA, and is therefore a reference in this study. Objective measures could have been used as a gold standard, such as instruments named accelerometers; however, the use of this technology also presents some limitations of applicability and is very costly.

CONCLUSION

The use of the VIGITEL system by telephone surveys is adequate to measure and monitor trends relating to indicators of physical and sedentary activity in leisure time. However, unsatisfactory results were obtained regarding the domain of commuting. The need still persists for other methods of evaluation with more accurate reference measures to confirm these results.

REFERENCES

1. Institute of Health Metrics and Evaluation (IHME). GBD Profile: Brazil, 2010. p. 1-4. Disponível em: https://www.healthdata.org/sites/default/files/files/country_profiles/GBD/ihme_gbd_country_report_brazil.pdf (Acessado em 06 de março de 2015).
2. World Health Organization (WHO). Prevalence of insufficient physical activity. Geneva: WHO, 2008.
3. World Health Organization (WHO). Global recommendations on physical activity for health. Geneva, Switzerland: WHO, 2010.
4. Matsudo S, Araújo T, Matsudo V, Andrade D, Andrade E, Oliveira LC, et al. Questionário internacional de atividade física (IPAQ): estudo de validade e reprodutibilidade no Brasil. *Rev Bras Ativ Fís Saúde* 2001; 6(2): 05-18.
5. 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.
6. Dumith SC. Physical activity in Brazil: a systematic review. *Cad Saúde Pública*. 2009; 25(Suppl 3): S415-26.

7. Armstrong T, Bull F. Development of the World Health Organization Global Physical Activity Questionnaire (GPAQ). *J Public Health* 2006; 14: 66-70.
 8. Macías R, Garrido-Muñoz M, Tejero-González CM, Lucia A, López-Adán E, Rodríguez-Romo G. Prevalence of leisure-time sedentary behaviour and sociodemographic correlates: a cross-sectional study in Spanish adults. *BMC Public Health* 2014; 14: 972.
 9. Brasil. Ministério da Saúde. Secretaria de Vigilância em Saúde. Secretaria de Gestão Estratégica e Participativa. VIGITEL Brasil 2011. Vigilância de fatores de risco e proteção para doenças crônicas por inquérito telefônico. Brasília: Ministério da Saúde, 2012.
 10. Monteiro CA, Florindo AA, Claro RM, Moura EC. Validade de indicadores de atividade física e sedentarismo obtidos por inquérito telefônico. *Rev Saúde Pública* 2008; 42(4): 575-81.
 11. Osti RFI, Garcia LMT, Florindo AA. Validation of the 24-hour physical activity recall in elderly adults. *Rev Bras Cineantropom Desempenho Hum* 2014; 16(1): 15-26.
 12. World Health Organization (WHO). Global Physical Activity Questionnaire (GPAQ). Analysis Guide. Disponível em: http://www.who.int/chp/steps/resources/GPAQ_Analysis_Guide.pdf (Acessado em 20 de maio de 2015).
 13. Lands JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33(1): 159-74.
 14. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res* 1999; 8(2): 135-60.
 15. Brasil. Ministério da Saúde. Secretaria de Vigilância em Saúde. Departamento de Análise de Situação de Saúde. Plano de ações estratégicas para o enfrentamento das doenças crônicas não transmissíveis (DCNT) no Brasil 2011-2022. Brasília: Ministério da Saúde, 2011.
 16. Wagnacker DS, Pitanga FJG. Atividade física comparada a atividade física no tempo livre como fator de proteção à hipertensão arterial sistêmica. *Rev Bras Ci e Mov* 2009; 17(1): 1-19.
 17. Perez LG, Pratt M, Simoes EJ, de Moura L, Malta DC. Association between leisure-time physical activity and self-reported hypertension among Brazilian adults, 2008. *Prev Chronic Dis* 2013; 10: E172.
 18. Zhao G, Li C, Ford ES, Fulton JE, Carlson SA, Okoro CA, et al. Leisure-time aerobic physical activity, muscle-strengthening activity and mortality risks among US adults: the NHANES linked mortality study. *Br J Sports Med* 2014; 48(3): 244-9.
 19. Hallal PC, Simoes E, Reichert FF, Azevedo MR, Ramos LR, Pratt M, et al. Validity and reliability of the telephone-administered international physical activity questionnaire in Brazil. *J Phys Act Health* 2010; 7(3): 402-9.
 20. Pettee KK, Ham SA, Macera CA, Ainsworth BE. The reliability of a survey question on television viewing and associations with health risk factors in US adults. *Obesity* 2009; 17(3): 487-93.
 21. Matthews CE, Shu XO, Yang G, Jin F, Ainsworth BE, Liu D, et al. Reproducibility and validity of the Shanghai Women's Health Study Physical Activity Questionnaire. *Am J Epidemiol* 2003; 158(11): 1114-22.
 22. Medina C, Barquera S, Janssen I. Validity and reliability of the International Physical Activity Questionnaire among adults in Mexico. *Rev Panam Salud Publica* 2013; 34(1): 21-8.
- Received on: 09/11/2015
 Final version presented on: 01/04/2016
 Accepted on: 08/09/2016