

# Reprodutibilidade (teste-reteste) do questionário internacional de atividade física (QIAF-Versão 6): um estudo piloto com adultos no Brasil

Mauro V.G. Barros<sup>1,2</sup>  
Markus V. Nahas<sup>1</sup>

**Resumo:** As medidas da atividade física incluem dois tipos principais: aquelas que utilizam informação fornecida pelas pessoas (questionários, entrevistas ou diários) e aquelas que utilizam indicadores fisiológicos (consumo de oxigênio, frequência cardíaca) ou sensores de movimento, que registram objetivamente certas características das atividades durante um período determinado. Os questionários têm sido usados extensivamente por sua aplicabilidade para grandes grupos, seu baixo custo e por permitirem coletar informações precisas sobre o *tipo* de atividades e o *contexto* em que acontecem. Entretanto, existem limitações quando se usam questionários, principalmente pela imprecisão das respostas e pela falta de um processo de construção e validação adequado, na maioria dos casos. Este estudo teve como objetivo a determinação da fidedignidade (teste-reteste) do Questionário Internacional de Atividade Física (QIAF-Versão 6), proposto pelo Grupo Internacional para Consenso em Medidas da Atividade Física, constituído sob a chancela da Organização Mundial da Saúde (OMS). A amostra incluiu 42 sujeitos (22 homens e 20 mulheres) com idade média de 34,7 anos (22 a 66; DP=8,8 anos), funcionários de diversos setores administrativos do SESI-SC. A estimativa do gasto calórico semanal (semana típica) para todas as atividades foi obtida a partir dos dados da Classificação do Custo Energético das Atividades Físicas Humanas, desenvolvida por Ainsworth, Haskell, Leon, et al. (1993) e o cálculo proposto por Kriska e Caspersen (1997). A análise dos dados foi feita através da ANOVA - medidas repetidas, coeficiente de correlação intraclassa - R (Thomas & Nelson, 1996; Safrit & Wood, 1989) e coeficiente Kappa (Coeficiente de Concordância para Escalas Nominais - Cohen, 1960). A análise foi realizada utilizando os programas Excell e SPSS 9.0 para windows. Os resultados sugerem que esta versão em português do QIAF-6 tem boa fidedignidade, confirmando-se os resultados de outros estudos de que a reprodutibilidade das medidas de atividades vigorosas (6 METs ou mais) é superior à reprodutibilidade das medidas de atividades moderadas (3 a 6 METs). Os coeficientes de correlação intraclassa podem ser considerados moderados a altos (R = 0,60 a 0,92) para a maioria dos itens, exceto para atividades moderadas no jardim e no lazer, caminhada e ciclismo como transporte. Os coeficientes Kappa para consistência de classificação (teste-reteste) podem ser considerados moderados para a maioria dos itens, enquanto o percentual de concordância na

classificação teste-reteste em três grupos foi modesta (56%) e em dois grupos pode ser considerada muito boa (93%).

**Abstract:** Measures of physical activity include two major types: those that use information given by the subjects (questionnaires, interviews, diaries), and those that utilize physiological markers or motion sensors for direct measurement of the activities during a period of time. Questionnaires have been extensively used for large groups because of the low cost and for practical reasons, but their validity and reliability are sometimes questionable. This pilot study intended to verify the reability of the International Physical Activity Questionnaire - IPAQ (WHO - version 6), under development and evaluation by different research groups around the world. The sample included 42 subjects, 22 males and 20 females (ages 22 to 66; M=34.7 years; SD=8.8), to whom a Portuguese version of the IPAQ-6 (usual week) was applied twice, a week apart. Subjects were office employees, with high school or college degrees. An estimate of the weekly caloric expenditure (usual week), was obtained utilizing the *Classification by Energy Cost of Human Physical Activities* (Ainsworth, Haskell, Leon, et al., 1993), and Kriska and Caspersen (1997). Data analysis was performed by ANOVA - repeated measures, the Intra-Class Correlation (R), and the Kappa statistic (Coefficient of Agreement for Nominal Scales - Cohen, 1960). Analysis was done using Excell and SPSS 9.0 for windows softwares. Small groups (10-15) were formed for the administration of the questionnaire, with the presence of the same instructor in all occasions. It took the subjects approximately 20 minutes to answer all questions. Significant F ( $p < 0.05$ ) was detected for Total Energy Expenditure in Sports, Recreation and Leisure time PA ( $F=5.94$ ;  $p=0.02$ ), mainly because of the moderate PA component in this item. Reliability (test-retest) was calculated for occupational physical activity; household and garden/yard work physical activity; transportation physical activity; recreation, sport, and leisure-time physical activity. To determine reliability for vigorous and moderate PA, information from leisure activities was grouped, and the values for R were 0.48 (moderate PA) and 0.92 (vigorous PA). Reliability of the overall caloric expenditure was 0.86. A coefficient of 0.8 or better is, usually, accepted as a good indicator of stability. Large interindividual variance was observed in all items of the questionnaire. Subjects were also separated in tertiles (active, moderately active, and low active/inactive) and two groups (high/low) according to the

1 - Núcleo de Pesquisa em Atividade Física & Saúde - UFSC.  
2 - Universidade de Pernambuco - UPE.

first application, and the consistency of the classification (agreement) in the second application was observed (one week later). Group classification is often of epidemiology significance for describing population segments that may benefit or respond in unique ways to specific interventions, and categorical measurement can at times be more accurate than continuum. Percentage agreement was higher for the two-way grouping, as expected. The classification had excellent agreement ( $Kappa > 0.75$ ) only for cycling for transportation; fair to good agreement (0.40-0.75) for moderate and vigorous garden/yard activities, vigorous leisure activity, and vigorous activity at work; and poor agreement for other variables. Results suggest that this Portuguese version of the IPAQ-6 is fairly reliable, confirming the findings from other studies which indicate better reliability for self-reported vigorous than moderate physical activity. Kappa coefficient of agreement values were in the fair to good range, and the overall accuracy of agreement (%) when the values were grouped can be considered modest for tertiles (56% correct classification) and excellent for two groups (high/low active – 93% correct classification).

## Introdução

As medidas da atividade física incluem dois tipos principais: aquelas que utilizam informação fornecida pelas pessoas (questionários, entrevistas ou diários) e aquelas que utilizam indicadores fisiológicos (consumo de oxigênio, frequência cardíaca) ou sensores de movimento, que registram objetivamente certas características das atividades durante um período determinado. Mais informações sobre medidas da atividade física em geral podem ser encontradas num artigo de Nahas (1995), onde as principais medidas da atividade física e sua utilização em grupos específicos são discutidas.

Os questionários têm sido usados extensivamente por sua aplicabilidade para grandes grupos, seu baixo custo e por permitirem coletar informações precisas sobre o *tipo* de atividades e o *contexto* em que acontecem (o que não acontece com os sensores de movimento e monitores de frequência cardíaca). Entretanto, existem limitações quando se usam questionários, principalmente pela imprecisão das respostas e pela falta de um processo de construção e validação adequado, na maioria dos casos.

Dishman e Steinhart (1988) citam como requerimentos mínimos para um instrumento de coleta de informações sobre atividades físicas realizadas a fidedignidade e validade das medidas, a fácil administração e a não interferência com os padrões habituais de comportamento. Para que um instrumento seja considerado válido, é preciso que ele seja fidedigno, quer dizer, se aplicado repetidas vezes num período de tempo em que o nível de atividades físicas não tenha mudado, os resultados devem ser semelhantes. A fidedignidade (ou reprodutibilidade) de um instrumento pode ser *absoluta* ou *relativa* (Safrit & Wood, 1989). Ela é absoluta quando se compara diversas aplicações de uma medida num mesmo grupo, considerando-se a significância da diferença absoluta (entre os escores). A fidedignidade relativa é a mais importante neste processo de validação de instrumentos de medida e geralmente é determinada por algum coeficiente de correlação. Neste caso, a fidedignidade estimada pode ser de *consistência interna*

(grau de similaridade dos escores em aplicações num mesmo dia), de *estabilidade* (reprodutibilidade dos resultados em aplicações em dias diferentes), ou de *objetividade* (grau de concordância entre resultados de um teste aplicado a um mesmo grupo em ocasiões diferentes, por diferentes pessoas).

Uma boa medida deve refletir as várias dimensões da atividade física habitual: frequência, duração, intensidade e tipo de atividade, possibilitando uma estimativa do gasto calórico total e em atividades moderadas ou vigorosas. Essas características num instrumento possibilitam seu uso em estudos epidemiológicos que pretendam investigar a relação entre atividade física e doenças específicas, mortalidade, qualidade de vida ou autonomia na velhice.

Este estudo teve como objetivo a determinação da fidedignidade (coeficiente de estabilidade, teste-reteste) do Questionário Internacional de Atividade Física (QIAF-Versão 6), proposto pelo Grupo Internacional para Consenso em Medidas da Atividade Física, constituído sob a chancela da Organização Mundial da Saúde (OMS), com representantes de 25 países, inclusive o Brasil.

## Procedimentos de Coleta e Análise dos dados

O Questionário Internacional de Atividade Física (QIAF) está sendo desenvolvido para servir como um instrumento válido para comparações dos níveis de atividade física entre diversas populações e contextos culturais e sociais. Foram propostas diversas versões do instrumento: curta e longa; em forma de entrevista por telefone e questionário auto-aplicado; considerando como período de referência a última semana ou uma semana típica (usual). O QIAF está sendo desenvolvido para estimar o gasto energético total (kcal) e o tempo gasto em atividades de diferentes intensidades, em quatro domínios: trabalho, transporte, recreação e lazer, e atividades domésticas (em casa e no jardim). O desenvolvimento do questionário prevê uma série de estudos para determinar a validade e fidedignidade dos escores obtidos nas diversas traduções e formas.

O QIAF, em sua versão 6 (na forma longa, semana usual/típica), foi traduzido para o português pelos autores e submetido à apreciação de dois outros pesquisadores fluentes em inglês e português. Este questionário foi aplicado em duas oportunidades, num espaço de sete dias, para os sujeitos deste estudo.

A amostra incluiu 42 sujeitos (22 homens e 20 mulheres) com idade média de 34,7 anos (22 a 66; DP=8,8 anos), funcionários de diversos setores administrativos do SESI-SC. Essas pessoas foram convidadas a participar neste estudo piloto (amostra de conveniência) e foram informadas que seus nomes não seriam divulgados, devendo lembrar-se do número do questionário respondido para que se pudesse fornecer os resultados da análise individual das informações (gasto calórico semanal em atividades físicas). O nível educacional do grupo pode ser considerado bom, incluindo, na maioria, graduados em nível médio e superior.

A estimativa do gasto calórico semanal (semana típica) para todas as atividades foi obtida a partir dos dados da Classificação do Custo Energético das Atividades Físicas

Humanas, desenvolvida por Ainsworth, Haskell, Leon, et al. (1993) e o cálculo proposto por Kriska e Caspersen (1997).

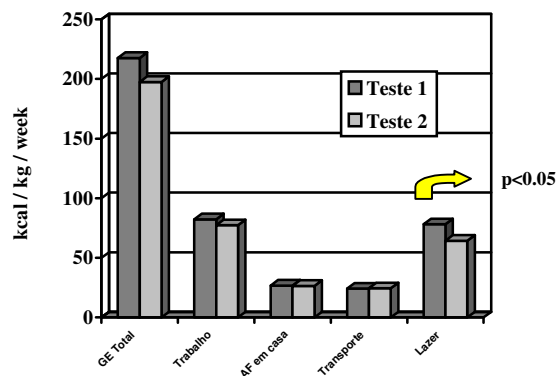
A análise dos dados foi feita através da ANOVA - medidas repetidas, coeficiente de correlação intraclass - R (Thomas & Nelson, 1996; Safrit & Wood, 1989) e coeficiente Kappa (Coeficiente de Concordância para Escalas Nominais - Cohen, 1960). A análise foi realizada utilizando os programas Excell e SPSS 9.0 para Windows.

Uma aplicação preliminar deste questionário mostrou que a administração em pequenos grupos (10-15 por grupo), com um instrutor presente, foi a maneira mais eficiente (menor tempo e respostas mais completas).

## Resultados & Discussão

Neste estudo piloto, a versão 6 do QIAF foi utilizada com referência às atividades físicas de uma semana usual (típica). Pequenos grupos foram formados (10-15) para administração do questionário, com a presença de um instrutor. Os sujeitos demoraram em torno de 20 minutos para responder a todas as questões. Apesar de 42 sujeitos terem participado do estudo, em alguns itens o número considerado para análise foi menor, por falta de informação.

A estimativa do gasto calórico para cada componente (trabalho, transporte, atividade doméstica e lazer) e o gasto energético total pode não refletir a medida real, mas isto em nada afetou a análise de reprodutibilidade, uma vez que o cálculo foi feito de maneira idêntica para as duas aplicações (erro sistemático).



**Figura 1** – Gasto Energético (GE) total e dos principais componentes do QIAF-6, nas duas aplicações (Teste 1, Teste 2)

A Tabela 1 inclui as médias, desvios-padrão, valores máximos e mínimos para os principais itens do questionário (gasto energético em kcal/kg/semana). Um F (ANOVA medidas repetidas) o detectado apenas para os valores de gasto energético em esportes, recreação e lazer (F=5,94; p=0,02), principalmente pela diferença no componente atividade física moderada. Observou-se também uma grande variação interindividual, como se pode verificar na Tabela 1 pelos altos valores do desvio padrão.

**Tabela 1** – Gasto Energético (kcal/kg/sem) nas Diversas Atividades, Valores de F e p (ANOVA – repeated measures) para T<sub>1</sub> (teste) e T<sub>2</sub> (reteste)

Variável	T	Média	DP	Max	Min	F	p
<i>1. AF no Trabalho</i>							
Gasto EnergéticoTotal	1	82,44	31,43	201,00	50,00	2,55	p=0,12
N = 36	2	77,41	27,48	192,00	45,00		
AF Vigorosa	1	8,73	20,17	84,00	0,00	0,93	p=0,34
N = 36	2	6,12	18,14	84,00	0,00		
AF Moderada	1	18,10	14,46	50,00	0,00	0,04	p=0,83
N = 36	2	18,70	18,73	80,00	0,00		
<i>2. AF em casa</i>							
Gasto EnergéticoTotal	1	26,76	31,87	139,83	0,00	0,00	p=0,96
N = 38	2	26,48	37,42	181,00	0,00		
<i>3. AF no transporte</i>							
Gasto EnergéticoTotal	1	24,18	17,17	85,00	2,63	0,00	p=0,93
	2	24,46	19,47	77,88	2,63		
<i>4. Recreação, esporte e lazer</i>							
Gasto EnergéticoTotal	1	78,24	42,32	245,00	26,67	5,94	p=0,02*
N = 31	2	64,33	27,44	135,25	26,67		
AF Vigorosa	1	13,03	18,87	72,00	0,00	0,00	p=0,99
N = 40	2	13,05	17,38	63,00	0,00		
AF Moderada (sem caminhada)	1	12,52	22,88	120,00	0,00	4,04	p=0,052
N=39	2	5,52	8,99	40,00	0,00		
AF Moderada (+ caminhada)	1	17,77	25,65	127,00	0,00	3,94	p=0,05
N = 39	2	10,35	13,49	61,25	0,00		
Caminhada	1	5,3	7,87	28,58	0,00	0,29	p=0,59
N = 40	2	4,7	7,58	28,58	0,00		
<i>5. Gasto EnergéticoTotal</i>							
	1	217,57	83,55	464,00	102,75	3,86	p=0,06
N = 27	2	197,48	76,11	392,50	107,00		

A fidedignidade (teste-reteste) dos escores foi determinada pelo coeficiente de correlação intraclass (R), calculado para todos os itens do questionário (Tabela 2). Os valores de R foram médios ou altos para quase todos os itens, exceto atividades moderadas no lazer, o que coincide com as informações de outros estudos, como o de Godin & Shephard (1985).

As atividades de lazer foram agrupadas para determinar a fidedignidade das atividades físicas moderadas (R = 0,48) e vigorosas (R = 0,92). O coeficiente de correlação intraclass para o gasto energético total foi igual a 0,86. Um coeficiente igual ou superior a 0,80 é considerado um bom indicador de estabilidade (teste-reteste).

**Tabela 2:** Coeficientes de Correlação Intraclass (R) para o QIAF-6 (N=42)

Variável	R
1. Atividade Física no Trabalho	0,88
2. Atividade Física em casa/jardim	0,67
3. Atividade Física no Transporte	0,68
4. AF Recreação, esporte, lazer	0,71
AF Vigorosa	0,92
AF Moderada (sem caminhada)	0,36
AF Moderada (+ caminhada)	0,52
Caminhada no lazer	0,75
5. Gasto energético total	0,86

Para determinar a consistência na classificação dos sujeitos em grupos nas duas aplicações do questionário, a amostra foi dividida em três (ativo, moderadamente ativo e pouco ativo), e dois grupos (superior e inferior) de acordo com os resultados da primeira aplicação. A classificação em grupos é interessante em estudos epidemiológicos para descrever os segmentos populacionais que podem mais se beneficiar ou responder de maneira peculiar numa intervenção. A Tabela 3 apresenta os valores dos coeficientes Kappa e os valores percentuais de concordância na classificação dos sujeitos na segunda aplicação em relação à primeira. Os percentuais foram maiores para a concordância em dois grupos do que em três, como era esperado. Os valores de Kappa (grau de concordância) foram excelentes ( $k > 0,75$ ) apenas para o uso da bicicleta como meio de transporte; médio para bom (0,40-0,75) para atividades moderadas e vigorosas no jardim, atividade vigorosa no lazer e no trabalho, e baixos (índices de concordância  $k < 0,40$ ) para as demais variáveis.

**Tabela 3:** Concordância de Classificação (%) e Coeficiente Kappa\* para o Teste-Reteste

Variável	% concordância		Kappa
	Terços	50/50	
<b>1. Atividade Física no trabalho</b>	<b>56%</b>	<b>72%</b>	<b>0,33</b>
AF Vigorosa	72%		0,58
AF Moderada	53%		0,29
<b>2. Atividade Física em casa/jardim</b>	<b>50%</b>	<b>74%</b>	<b>0,25</b>
AF Vigorosa	71%		0,56
AF Moderada	76%		0,64
<b>3. Transporte</b>	<b>61%</b>	<b>68%</b>	<b>0,41</b>
Caminhada	45%		0,17
Bicicleta	95%		0,92
<b>4. Recreação, esporte e lazer</b>	<b>55%</b>	<b>69%</b>	<b>0,32</b>
AF Vigorosa	70%		0,55
AF Moderada	62%		0,42
Caminhada - lazer	65%		0,47
<b>5. Gasto Energético Total</b>	<b>56%</b>	<b>93%</b>	<b>0,39</b>

\* Coeficiente de Concordância para Escalas Nominais (Cohen, 1960)  
 Kappa – Nível de Concordância: > 0,75 – excelente  
 0,40 - 0,75 – Médio para bom  
 < 0,40 - baixo

## Conclusões

Os resultados sugerem que esta versão em português do QIAF-6 tem boa fidedignidade, confirmando os resultados de outros estudos de que a reprodutibilidade das medidas de atividades vigorosas (6 METs ou mais) é superior à reprodutibilidade das medidas de atividades moderadas (3 a 6 METs). Os coeficientes de correlação intraclasse podem ser considerados moderados a altos ( $R = 0,60$  a  $0,92$ ) para a maioria dos itens, exceto para atividades moderadas no jardim e no lazer, caminhada e ciclismo como transporte. Os coeficientes Kappa para consistência de classificação (teste-reteste) podem ser considerados

moderados para a maioria dos itens, enquanto o percentual de concordância na classificação teste-reteste em três grupos foi modesta (56%) e em dois grupos pode ser considerada muito boa (93%). Estudos com pessoas de nível educacional e sócio-econômico diferentes deste são necessários, uma vez que a validade e a fidedignidade de instrumentos de medida são específicas a cada grupo populacional. Também são necessários estudos de validação (de conteúdo, de critério e de constructo) da versão brasileira do QIAF-6, pois um instrumento pode ser fidedigno, mas não necessariamente válido.

## Referências Bibliográficas

1. A *Collection of Physical Activity Questionnaires for Health-Related Research* (1997). *MSSE*, 29(6-S).
2. AINSWORTH B, HASKELL W, LEON AS, et al. (1993). *Classification by energy cost of human physical activity*.
3. COHEN J (1960). "A coefficient of agreement for nominal scales". *Education and Psychology Measures*, 20(1), 37-46.
4. DISHMAN R & SUNHARD M (1988). "Reliability and concurrent validity for a 7-d re-call of physical activity in college students". *Medicine and Science in Sports and Exercise*, 20(1), 14-25.
5. GODIN G & SHEPHARD RJ (1985). "A simple method to assess exercise behavior in the community". *Canadian Journal of Applied Sports Science*, 10(3), 141-146.
6. KRISKA, AM & CASPERSEN, CJ (1997). "Introduction to collection of physical activity questionnaires". *Medicine & Science in Sports and Exercise*, 29(6), S5-S9.
7. MONTOYE HJ; KEMPER CG; SARIS WHM & WASHBURN RA. (1996). *Measuring physical activity and energy expenditure*. Champaign. IL: Human Kinetics.
8. NAHAS, MV (1995). "Revisão de métodos para determinação da atividade física em diferentes grupos populacionais". *Revista Brasileira de Atividade Física e Saúde*, 1(4), 27-37.
9. SAFRIT MJ & WOOD TR (1989). *Measurement concepts in physical education and exercise science*. Champaign. IL: Human Kinetics.
10. THOMAS JR & NELSON JK (1996). *Research methods in physical activity*. Champaign. IL: Human Kinetics.