

Addressing overreporting on the International Physical Activity Questionnaire (IPAQ) telephone survey with a population sample

Randy Rzewnicki^{1,*}, Yves Vanden Auweele¹ and Ilse De Bourdeaudhuij²

¹Faculty of Physical Education and Physical Therapy, Katholieke Universiteit Leuven, Tervuursevest 101, B3000 Leuven, Belgium; ²Department of Movement and Sport Sciences, Ghent University, Ghent, Belgium

Submitted 17 July 2002; Accepted 10 October 2002

Abstract

Objective: To examine a possible problem of overreporting and to describe the degree of error with the International Physical Activity Questionnaire (IPAQ) short telephone protocol.

Design: Cross-sectional study, using two different physical activity (PA) self-report protocols.

Setting: Telephone interviews about PA in Belgium.

Subjects: Fifty adults who had previously been interviewed with IPAQ in a national survey.

Results: Seventy-five per cent reported less PA with the modified procedure than with the IPAQ. Twenty-three of the 50 individuals were found to have reported some amounts of PA with the IPAQ (either walking, or vigorous or moderate PA) when they should have reported none. In total, based on their revised reports of PA, 50% fewer persons met PA recommendations than was the case with IPAQ. The overreporting could not be related to types of error-prone individuals.

Conclusions: Overreporting of PA in population samples is a serious problem that could be reduced by implementing procedure changes without changing the IPAQ items themselves.

Keywords

IPAQ

Overreporting

Telephone survey

Physical activity monitoring

The need for the development and implementation of an international standard for measuring physical activity^{1,2} has been responded to. Long and short versions of the International Physical Activity Questionnaire (IPAQ) were developed, with distinct purposes. The short IPAQ was recently developed for use by public health officials to monitor entire populations, and its reliability and validity have been tested in over a dozen countries^{3,4}. As with many self-reported surveys, there are some concerns that the IPAQ has a tendency towards overreporting of physical activity (PA).

Self-report of PA is subject to a variety of factors, most notably overreporting of time and intensity⁵. Some guidelines for interviewers explicitly warn about this, and provide relevant instructions^{6,7}.

One reason for overreporting PA may be the wish to present oneself in a positive light by giving socially desirable responses, which is the tendency to describe oneself in favourable terms⁸. It is an important factor in responding to surveys that affects the validity of self-reports, because it results in underreporting of socially undesirable characteristics or overreporting of socially desirable behaviour⁹.

The tendency to make socially desirable reports varies across social groups that hold differing norms and values

regarding the desirability of the behaviours or traits under investigation. Those with higher education and income may tend to overreport PA since their peers value a healthy lifestyle more than individuals in lower classes¹⁰. Less educated and older males may have a propensity to overreport physical activity from the distant past, whereas younger and higher educated males are more likely to underreport¹¹.

We suspected overreporting in our examination of a recent survey using the short (last 7 days) telephone interview IPAQ (IPAQ-S7T). An RDD (random-digit dialling) national sample of Belgian adults ($n = 610$, 335 females), with characteristics generally similar to the population, completed an IPAQ-S7T between December 2001 and May 2002. The interviewers were adult employees of a marketing research company with a wide range of experience, up to several years. They were supervised for adherence to the protocol by a company supervisor and the research project leader. The questionnaire was administered via a Computer Assisted Telephone Interview (CATI) system. On the computer screen CATI displayed the telephone script described by the IPAQ standardisation group³, which contains explicit definitions of vigorous and moderate intensity PA with examples, and minimum duration. Reminders to interviewers about

*Corresponding author. Email randy.rzewnicki@flok.kuleuven.ac.be

the 10-minute minimum duration criterion appear repeatedly on the CATI screen, but their mention to respondents is left to the interviewers' discretion.

Answers to the IPAQ questions about time spent sitting and walking were summed with reports of moderate and vigorous PA. Five per cent of the national sample ($n = 30$) accounted for more than 126 h week⁻¹ or 18 h day⁻¹, including 20 of the 30 who reported more than 126 h of walking, moderate and vigorous PA. These reports may be considered not credible, if only because they allow, at most, 6 h of sleep per day.

This 5% of reports with extremely high (not credible or even impossible) levels of activity may indicate problems with the IPAQ protocol, which may not be limited only to these extreme cases. It is likely that the problems that allowed reporting of such extreme data by this 5% also affect the reports made by people with lower reported amounts of PA, although perhaps to a lesser degree.

Answers to the IPAQ questions about moderate and vigorous PA were used to assess compliance with health-related recommendations. The criteria for moderate-intensity activities are accumulating 30 min per day on five or more days per week. For vigorous activity the criteria are 20 min or more per day on at least three days per week. Respondents meeting either or both of these criteria* were classified as meeting recommended levels of PA, consistent with guidelines appearing in the US Surgeon General's report^{12,13}.

Seventy per cent of our Belgian sample was found to meet PA recommendations. This contrasts with the recent report¹⁴ of 68% of Belgians who do not meet PA recommendations. However, there are major differences between the IPAQ and the other measures used. Most important is the limited range of leisure-time PA (LTPA) versus the wide range of domains in the IPAQ, which add transportation activity, household and family chores, and work activity to LTPA and sport. This makes it very difficult to compare results between studies. It is possible that only a small proportion of the population meets the recommendations when only LTPA is taken into account, but that almost 70% of people meet the standards if transportation, work activity and housework and gardening are added.

Despite some uncertainty about whether or not there is a problem with overreporting in our data, it seems that at least 5% of the respondents have overreported their amount of PA.

The IPAQ protocol asks respondents to report an average time per day each activity is performed. Here problems might arise. If PA is reported for more than a

single day, the subject must report an average time per day. It is likely that subjects will tend to overreport this mean time per day, by reporting the day they executed the activity most. Also, the fact that respondents have to calculate mean hours per day over many activities, including work, transport, leisure, household and gardening activities, increases the chance of exaggeration.

The interviewers conducting the original data collection and their supervisors were experienced with interviews, but not specially trained to handle issues related to PA surveying, such as overreporting. The many records showing possible overreporting indicated that inadequate attention was given to this problem while interviewing. As argued before, this might also be a problem in (many) other records where extreme scores are not found. Providing interviewers with a briefing, a script and regular supervision seems inadequate to exclude overreporting. Specific training and guidelines for interviewers focusing on overreporting, such as developed for the Minnesota physical activity interview⁶, could be an important addition.

In summary, the 5% of reports that are not trustworthy raise questions about the degree to which other reports are affected. The unexpectedly high proportion of the sample meeting PA recommendations implied that problems of overreporting might affect more than 5%.

We suspect that features of the IPAQ increase the chance of overreporting. We propose that the training and supervision of the interviewers should be considered more closely to handle this.

At the current stage of development, the degree to which IPAQ is subject to overreporting is unknown. We proposed to check if specially trained and supervised interviewers using the IPAQ procedures would again produce overreporting. Therefore we re-interviewed 50 respondents from the initial sample with the same instrument and the same instructions to interviewers. A second objective was to describe features of overreporting and estimate its impact on PA recommendations. To achieve this, after the respondents completed the IPAQ, they were asked to explain their responses, and to detail their PA for the last 7 days.

Material and methods

Subjects

Respondents were randomly selected from the national sample of Belgian adults who had previously been interviewed with the IPAQ (see Introduction). They were among the 90% who had given explicit consent to be contacted again. Their PA response patterns at that time, called IPAQ-1, ranged from very low to very high. Those reporting above median amounts of PA were over-sampled in order to examine the issue at hand. Sixty-six people were contacted. Fifteen refused and 51 completed the interview, giving a response rate of 77%. One case was

*The PA recommendations criteria are understood to refer to leisure time (LTPA), while the IPAQ additionally measures PA at work, at home and for transport. Lacking more appropriate accepted criteria, we apply the moderate and vigorous LTPA recommendations to the results from IPAQ data strictly for the purposes of this study.

unable; thus the present paper reports data from 50 subjects.

Interviewers

Two female interviewers (university graduates, native speakers, aged 30–40 years) were given 6–8 hours of training, and practice interviews. The training focused on the potential problem of overreporting, with the aim of obtaining accurate reports. During the data collection phase, they were supervised and given feedback by the first author (R.R.).

Measures

The IPAQ-S7T was administered under the exact same conditions as the initial protocol. We call it IPAQ-2 to distinguish it from the initial national survey. After giving informed consent, the nine IPAQ items, which ask for reports of walking and moderate and vigorous PA in the last 7 days, were posed in the same order and with the same instruction set. There were no procedural variations from the initial protocol. The interviewers did not use their specialised training to prevent overreporting in this phase of the interview.

A probe protocol was used to systematically collect data about the responses provided to the IPAQ. During the interview, immediately after answering the final IPAQ-S7T questions, respondents were asked to explain their responses, and give more exact, more complete and more detailed reports for the last 7 days. This is called the IPAQ-Probe protocol. A sample question posed by interviewers: 'You said that you did vigorous physical activity on 2 days for an average of 2 hours. Can you please tell me about that activity'. Probe questions included 'What was the time at start and finish?' and 'How was your breathing affected?' Attention was given to the explicit and novel criteria used by the IPAQ: i.e. breathing intensity for moderate and vigorous PA, and minimum duration of 10 min for individual bouts of PA. When either the respondent or the interviewer determined that IPAQ criteria were not met (e.g. time, intensity), no PA was recorded for that category.

Data handling

Following standard IPAQ procedures, the number of days was multiplied by the number of hours reported for each component. If the usual pace of walking was reported as slow, walking reports were excluded (from calculations related to meeting PA recommendations). In order to provide comparable data from the IPAQ-Probe reports, the number of minutes reported for each activity was summed in their respective intensity component categories.

Statistics

All physical activity scores were strongly skewed, and standard deviations (SD) were as much as double the mean scores, as seen in the original dataset and previous

research with PA self-reports¹⁵. The assumptions of statistical tests we intended to conduct were severely violated. Therefore a logarithmic transformation (log) was used to approximate the normal distribution^{13,16}.

Descriptive statistics reported include frequency counts, mean, SD and median. Pearson product-moment correlations were used to test for association between the measures. To test for differences between the measures, paired comparisons *t*-tests were used.

To determine if erroneous reporting of PA was related to types of individuals, we used two procedures. Chi-square tests compared those who overreported with those who did not on personal and sociodemographic characteristics (i.e. body mass index (BMI), income, sex, etc.). Hierarchical cluster analysis of the difference scores between IPAQ-2 and IPAQ-Probe on walking, moderate PA and vigorous PA was used to make three groups. Chi-square tests then compared the same set of characteristics as above. All tests were conducted using SAS version 8¹⁷, and the level of statistical significance was set at $P < 0.05$.

Results

Profile of the respondents

Descriptive data on study participants are shown in Table 1. The study population included equivalent numbers of males and females, French and Dutch speakers, as well as people from all age groups, education, employment and income categories. Half of the sample is categorised as having normal BMI (i.e. BMI = 19.5–24.99 kg m⁻², data not shown), and 47% is either overweight or obese (BMI > 24.99 kg m⁻²).

Reports of PA made with IPAQ-2 are significantly larger than with IPAQ-Probe on all three components and on the total of reported PA. The means and medians of the IPAQ-Probe reports are less than half the size of the IPAQ-2 reports in all but one case (Table 2).

Table 1 Characteristics of the sample ($n = 50$)

Characteristic	%
Sex	
Female	52
Age (years)	
18–24	12
25–34	8
34–44	30
45–54	14
55–64	16
65+	20
Work status	
Paid work	50
Unemployed	4
Income	
High (> 75th percentile)	17.8
Low (< 25th percentile)	35.6
Education	
< 10 years	32.7
> 12 years	32.7

Table 2 Mean (SD) and median duration (min week⁻¹) of PA reported with the IPAQ-2 and IPAQ-Probe protocols

	IPAQ-2	IPAQ-Probe	IPAQ-2 vs. IPAQ-Probe (t-test)
Vigorous PA	212.8 (364.8)	64.68 (155.2)	4.08**
Median	10	0	
Moderate PA	163.9 (304.7)	121.3 (287.2)	2.70*
Median	52	0	
Walking	520.9 (700.0)	173.6 (366.7)	2.95*
Median	270	30	
Total PA	682.6 (819.6)	282.0 (451.9)	4.58***
Median	382.5	112.5	

SD – standard deviation; PA – physical activity; IPAQ – International Physical Activity Questionnaire.
 *, *P* < 0.01; **, *P* < 0.001; ***, *P* < 0.0001.

Results of the paired *t*-tests of the log-transformed data comparing the IPAQ-2 and IPAQ-Probe protocols appear in Table 2. PA levels reported with IPAQ-2 were significantly higher than those reported with IPAQ-Probe.

On each IPAQ component, 14 to 20% of the respondents who had reported some PA with the IPAQ-2 reported no PA to the relevant questions with the IPAQ-Probe protocol. This was the case for 10 of the 25 reports of vigorous PA, eight of 30 reports of moderate PA, and eight of 26 reports of walking, in total 23 different individuals. These cases are called ‘zeros’.

Table 3 displays the frequencies of responses that differed between the protocols. Thirty-seven subjects (74%) reported less total PA with IPAQ-Probe than they had with IPAQ-2 (overreporters). Five respondents reported more total PA with IPAQ-Probe than with IPAQ-2 (underreporters), and eight subjects reported exactly the same amount of PA on the two protocols (‘stable’).

The plots in Fig. 1 show the log-transformed data from IPAQ-2 plotted against those from IPAQ-Probe. Each plot shows the same pattern of two distinct data clouds. One data cloud indicates the degree of variation between the two reports, which appears to be small. The collection of data points at the bottom of each plot represents the ‘zeros’, whose reports from IPAQ-2 were not confirmed with IPAQ-Probe.

Table 3 Proportion of respondents (%) whose responses on PA differed between IPAQ-2 and IPAQ-Probe protocols

	Underreporters*	Stable†	Overreporters‡	Zeros§
Vigorous PA	2	56	42	20
Moderate PA	44	12	44	16
Walking	8	24	68	16
Total PA	10	16	74	14

PA – physical activity; IPAQ – International Physical Activity Questionnaire.
 * Underreporters – more PA reported with IPAQ-Probe than with IPAQ-2.
 † Stable – (no change) reports at IPAQ-2 are confirmed with IPAQ-Probe.
 ‡ Overreporters – less PA reported with IPAQ-Probe than with IPAQ-2.
 § Zeros – reports of PA with IPAQ-2 that were found to be zero with IPAQ-Probe.

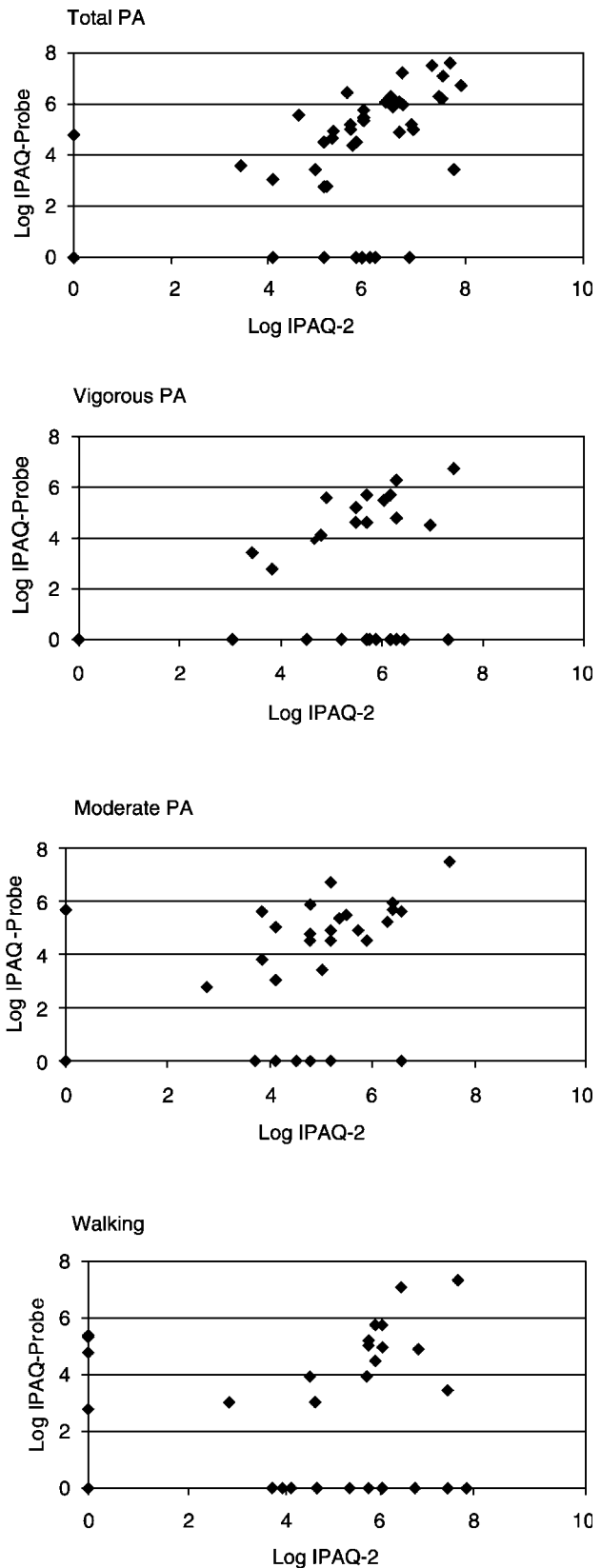


Fig. 1 Scatter plots of log-transformed data with IPAQ-2 and IPAQ-Probe. IPAQ – International Physical Activity Questionnaire

Significant correlations for each of the PA components ranging from 0.52 to 0.79 ($P < 0.0001$) were observed between IPAQ-2 and IPAQ-Probe, indicating a moderate to strong relationship.

To identify potential error-prone individuals a visual examination was conducted of personal and demographic features (age, gender, language, work status, income, education, BMI) of two subgroups of respondents, those identified as 'overreporters' ($n = 37$) vs. the 'stable' and 'underreporters' ($n = 13$). The goal was to identify who were the most or least likely to make erroneous reports. No differences in the make-up of these groups were observed.

We thereafter compared the 'zeros' (i.e. those who made erroneous reports on at least one IPAQ-2 component, $n = 23$) with those who did not make such gross overreports to try and identify them. Examination of these two groups' personal and sociodemographic characteristics with chi-square tests produced no significant differences.

Personal and sociodemographic characteristics of the groups produced by the cluster analysis were examined. Chi-square tests did not produce significant differences.

Comparison of IPAQ-2 and IPAQ-Probe data with recommended PA norms

Figure 2 shows the percentage of the sample who met PA recommendations according to measurement protocols. Fifty-six per cent of the sample responding to the IPAQ-2 met at least one of the recommendations for vigorous or moderate PA. Based on IPAQ-Probe, 28% of the sample was categorised as meeting the recommendations.

Qualitative data explaining problems in asking or answering the IPAQ items

Qualitative responses to open-ended questions about the IPAQ were collected from interviewers and respondents.

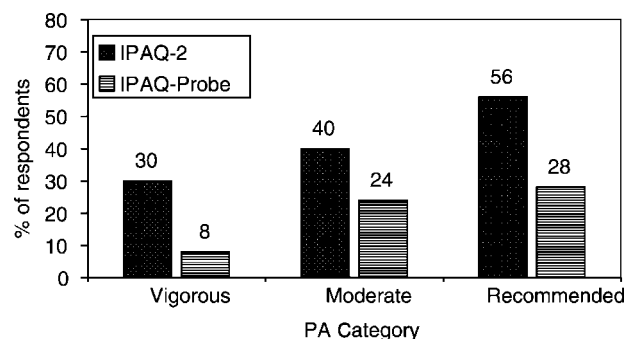


Fig. 2 Percentage of respondents ($n = 50$) meeting recommendations for physical activity (PA) with interview protocols IPAQ-2 and IPAQ-Probe. The criteria for moderate-intensity activities are accumulating 30 minutes per day on five or more days per week. For vigorous activity, the criteria are 20 minutes or more per day on at least three days per week. Respondents meeting either or both of these criteria were classified as meeting recommended levels of PA, consistent with guidelines appearing in the US Surgeon General's report¹². IPAQ – International Physical Activity Questionnaire

Interviewers said the biggest problems were the time criterion (i.e. minimum of 10 continuous minutes), the intensity criterion (breathing) and making the sums or estimates for the last 7 days. These were problems particularly for walking and moderate PA. Respondents usually did not consider the 10-minute minimum overtly and specifically. Reports of walking to the shop, for example, often were revealed to be less than 10 min. Some people said that they were walking and/or doing moderate PA 'all day', 'all the time' or 'non-stop'. When asked to quantify it precisely, reports of 8–18 h were provided. Evaluation of intensity was a frequent problem, as many people used criteria other than modification of breathing to evaluate it. Respondents referred instead to sweating, feeling hot, increased heart beat and even disliking the activity.

Discussion

Large proportions of the sample overreported PA with the IPAQ, and many subjects reported PA that did not qualify for reporting with the IPAQ. While 74% reported less PA with the IPAQ-Probe protocol, one quarter did not overreport their amount of PA, including 10% who reported more PA with the IPAQ-Probe protocol. Therefore it appears that the probe protocol itself is not responsible for the reports of less PA.

Visual inspection of PA reported with the IPAQ and probe protocols (see Tables 2 and 3) shows there are serious differences related to the reported PA on the different components (vigorous, moderate and walking). Over 40% of the subjects overreported vigorous and moderate PA, and over-two thirds overreported walking. While the majority confirmed that their IPAQ report of vigorous PA was accurate, this was the case for only one quarter of the walking reports and one in eight of the moderate PA reports. Forty-four per cent underreported moderate PA with the IPAQ. This implies that many people from the general population do not understand the IPAQ consistently.

The plots in Fig. 1 show the problem succinctly. The degree of variation between the IPAQ and Probe reports appears to be small for some of the people, some of the time. The collection of data points at the bottom of each plot indicates that a critical feature of the overreporting issue is respondents reporting PA that should not be reported. Over 25% of reports of moderate PA, over 30% of reports of walking and 40% of reports of vigorous PA should have been zeros. In these cases, for these items, none of the PA reported was relevant.

The correlations between IPAQ-2 and IPAQ-Probe are moderate to high, in part because the reference period, the interviewer and moment of the interview are exactly the same. The correlations would be much higher if the 'zeros' were prevented.

The use of a more specific IPAQ-Probe protocol administered by the skilled interviewers produced large and significant decreases in PA reports. These reports are statistically and realistically smaller than the PA that was reported with the IPAQ-2.

In addition to PA, overreporting has been found to be significantly correlated with health reports and health-related behaviours^{18,19}. We attempted to identify the people responsible for inaccuracies in reporting their PA behaviours after considering evidence of systematic (non-random) reporting bias on self-reports of diet and weight, and such bias for PA shown by high-income groups¹⁰, as well as older and less educated males¹¹. However, we could not isolate personal or sociodemographic characteristics related to PA self-reports. Given the relatively small samples in this study, we suggest further examination of this issue may be more fruitful. Identification of types of error-prone individuals might allow calculation of a correction factor, or focus the attention of interviewers during data collection.

Responses to IPAQ questions about moderate and vigorous physical activity were used to assess compliance with health-related recommendations. On the basis of the strict IPAQ-Probe protocol, exactly half as many meet these recommendations as would meet it with the data from IPAQ-2. Less than one-third of those who would be classified as meeting the vigorous recommendations using the IPAQ would meet them using the probe protocol. The current IPAQ protocol thus appears to risk overestimating the prevalence of those who perform recommended PA levels and underestimating the prevalence of insufficient PA.

The repeated measures method has some disadvantages, such as learning and memory effects, which may have affected our study. The magnitude and significance of our results suggest, nonetheless, that the phenomenon we discuss is real and serious. Further investigations of this issue might profitably make use of alternative approaches and methods, such as split samples.

The problem of overreporting that we document could be due to features of the IPAQ, such as asking for average times and best estimates of frequencies, asking about PA at work, transport and at home, and using a criterion of perceived intensity of breathing. Or perhaps it is due to the interviewers using a novel measurement tool, with a novel criterion and novel PA modes. Or maybe the respondents may have had difficulty with some features, novel or standard. Our specially trained and supervised interviewers delivering the standard IPAQ protocol obtained extreme reports of PA including up to 28 h of vigorous PA and 56 h of walking per week. When the IPAQ-Probe protocol was strictly applied, significant reductions in reported PA were noted, and many reports of PA were eliminated as they were too brief or insufficiently intense. Thus, it is not only general interviewer training that makes the difference. A focus on overreporting of PA is needed. Interviewers must use a

protocol that attends to issues of overreporting. Properly trained and supervised university students might well provide more accurate responses to PA questionnaires than skilled marketing researchers lacking relevant knowledge.

Conclusion

Using the standard IPAQ script, our specially trained interviewers obtained many very high reports of PA. When they used a probe protocol and probing skills, they obtained reports of PA that were significantly lower. A large proportion of the sample had reported PA that should not have been included. This resulted in an overestimation of the number of people who meet the PA recommendations by a factor of 2. Thus, the results of overreporting with IPAQ have a considerable impact on calculations of such public health estimates.

The standard IPAQ-S7T protocol apparently led respondents to overreport PA, affecting reports of moderate and vigorous PA and walking, and the sum of all three components, regardless of interviewer training. The identification of those who overreport might be useful. Prevention of such gross overreporting is preferable.

Recognising the importance and value of an international survey instrument for physical activity, we make the following recommendations to users of the IPAQ.

1. Interviewers administering the IPAQ should have adequate training, regular supervision and feedback. Interviewers and supervisors should be familiar with overreporting issues, and how to promote accuracy and precision in reporting. Supervisors should monitor for overreporting during interviews.
2. Where possible a CATI system should be programmed to detect extreme reports. CATI should prompt interviewers to probe such responses. The programming could involve specific responses, such as a maximum number of hours per day or days per week that one reported doing vigorous PA. It could also direct probing if the product of days and hours exceeded set levels; and to probe if total reported PA time, perhaps including time reported sitting, was questionable.
3. For all interviews, IPAQ items, or specific parts such as intensity and time criteria, could be repeated or phrased in question form, to confirm they are comprehended and used by respondents.
4. Users should consider instituting a validation protocol, such as the one described in this report.
5. Reports of IPAQ results should include extreme scores, outliers and disposition thereof.

Acknowledgements

The authors would like to express their gratitude for assistance to R. Stroobants, M. Sjöström, P. Godin,

M. Lejeune, P. Van deVliet, M. Hulens, A. Vuileman, T. Stahl, P. Lariviere and G. Lambear for their invaluable assistance with this research.

This research was funded by the KU Leuven, Faculty of Physical Education and the Research Unit of Sports and Exercise Psychology.

References

- 1 Booth ML. Assessment of physical activity: an international perspective. *Res. Q. Exerc. Sport* 2000; **71**: 114–20.
- 2 Booth M, Jones DA, Ainsworth BE. Methods of assessing the repeatability and validity of self-report measures of physical activity participation: a discussion paper. August 1999.
- 3 Craig CL, Russell SJ. Reliability and validity of measures of adult physical activity patterns. In Mini-symposium: Can public health surveillance of physical activity be standardized internationally? *Med. Sci. Sports Exerc.* 1999; **31**: 389.
- 4 Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, *et al.* and the IPAQ Consensus Group and the IPAQ Reliability and Validity Study Group. International Physical Activity Questionnaire (IPAQ): 12-country reliability and validity. *Med. Sci. Sport Exerc.* 2002; in press.
- 5 Montoye HJ, Kemper HGC, Saris WMH, Washburn RA. *Measuring PA and Energy Expenditure*. Champaign, IL: Human Kinetics, 1996.
- 6 De Backer G, Kornitzer M, Sobolski J, Dramaix M, Degre S, de Marneffe M, *et al.* Physical activity and physical fitness level of Belgian males aged 40–55 years. *Cardiology* 1981; **67**: 110–28.
- 7 Taylor HL, Jacobs DR, Schucker B, Knudsen J, Leon AS, Debacker G. Questionnaire for the assessment of leisure time physical activities. *J. Chron. Dis.* 1978; **31**: 741–55.
- 8 Edwards AL. *The Social Desirability Variable in Personality Assessment and Research*. New York: Dryden, 1957.
- 9 Nunnally JC. *Psychometric Theory*, 2nd ed. New York: McGraw-Hill, 1978.
- 10 Droomers M, Schrijvers CTM, van de Mheem H, Mackenbach JP. Educational differences in leisure time physical inactivity: a descriptive and explanatory study. *Soc. Sci. Med. Health* 1998; **47**(11): 1665–76.
- 11 Falkner KL, Trevisan M, Zielezny J, Freudenheim W, Winkelstein W, Fisher RP. Relative validity of recall of physical activity in the distant past [abstract]. *Am. J. Epidemiol.* 1994; **139**: S17.
- 12 US Department of Health and Human Services. *Physical Activity and Health: A Report of the Surgeon General*. Atlanta, GA: Centers for Disease Control and Prevention, 1996.
- 13 Jones DA, Ainsworth BE, Croft JB, Macera CA, Lloyd EE, Yusuf HR, *et al.* Moderate leisure time physical activity: who is meeting the public health recommendations? A national cross-sectional study. *Arch. Fam. Med.* 1998; **7**: 285–9.
- 14 Margetts BM, Rogers E, Widhal K, Remaut de Winter A-M, Zunft H-JF. Relationships between attitudes to health, body weight, and physical activity and level of physical activity in a nationally representative sample in the European Union. *Public Health Nutr.* 1999; **2**(1a): 97–103.
- 15 Rzewnicki R, De Bourdeaudhuij I, Stahl T, Welshman J, Vanden Auweele Y, Ziemainz H, *et al.* How methods affect measures of physical activity in 8 European countries: The Eupass Experience. Submitted for publication.
- 16 De Bourdeaudhuij I, Sallis J. Relative contribution of psychological determinants to the explanation of physical activity in three population-based adult samples. *Prev. Med.* 2002; **34**: 279–88.
- 17 SAS Institute, Inc. *SAS Procedures, Version 8*. Cary, NC: SAS Institute Inc., 2000.
- 18 Kristiansen CM, Harding CM. The social desirability of preventive health behavior. *Public Health Rep.* 1984; **99**: 384–8.
- 19 Mathers C, Vos T, Stevenson C. The burden of disease and injury in Australia. *Bull. World Health Org.* 2001; **79**: 1076–84.