VALIDATION OF THE GREEK VERSION OF THE INTERNATIONAL PHYSICAL ACTIVITY QUESTIONNAIRE – SHORT FORM IN ADULTS 18–65 YEARS OLD

Nikos Dimitriadis[1, 2], Maria Partsafa[1], Nikoleta-Maria Bokou[1], Alexandros Dimitriadis[2], Katy Taveira[2], Demosthenes Panagiotakos[1]
[1]Department of Nutrition and Dietetics, School of Health Sciences and Education, Harokopio University of Athens, Athens, Greece
[2]Myathlete Exercise Physiology Hub, Glyfada, Greece

SUMMARY

Objectives: Various tools and methods have been proposed to measure physical activity, with the International Physical Activity Questionnaire (IPAQ) being one of the most popular self-reported tools. The study's aim was to test the validity of IPAQ-SF among Greek male and female adults, aged 18–65 years.

Methods: A total of 134 adults (98 males) were randomly selected and participated in the study. The mean age of participants was 37.9 years (SD = 11.6), with 73.1% being male. Translation and back-translation was performed to ensure cultural adaptation. The validity of IPAQ-SF was ascertained using fitness measures (maximum oxygen uptake, VO_2 max, velocity of VO_2 max, maximum heart rate, and heartbeat at 1 min).

Results: Based on the IPAQ-SF, participants engaged in vigorous physical activity for an average of 3 days per week (mean = 68.8 minutes/day), moderate activity for 2.6 days per week (mean = 56.2 minutes/day), and walking for 3.7 days per week (mean = 24.4 minutes/day). The median total MET-minutes/week was 2,751. IPAQ-SF demonstrated moderate validity as modest positive correlations were observed between MET-minutes/week and VO_2 max (rho = 0.3096, $p < 0.001$), and vVO_2 max (rho = 0.2511, $p = 0.008$). IPAQ-SF demonstrated higher validity among females and individuals with BMI ≥ 25 kg/m^2, with correlations reaching 0.5201 ($p < 0.001$) for VO_2 max in participants with overweight/obesity.

Conclusion: The present validation study suggests IPAQ-SF utility for monitoring physical activity levels in populations with overweight or obesity, while further refinement may be necessary to improve its applicability across all demographic groups.

Key words: physical activity, tools, IPAQ, validity, adults, Greece

Address for correspondence: D. Panagiotakos, Department of Nutrition and Dietetics, School of Health Sciences and Education, Harokopio University, 70 El. Venizelou Street, Kallithea, Athens 17676, Greece. E-mail: dbpanag@hua.gr

https://doi.org/10.21101/cejph.a8446

INTRODUCTION

It is now very well understood and appreciated that physical activity is essential for maintaining and improving overall health, contributing to physical, mental, and emotional wellbeing. Regular engagement in physical activity can substantially prevent several chronic conditions, mainly cardiovascular disease, which account for approximately one half of disability adjusted life years lost worldwide, as well as enhance quality of life and promote longevity (1–3).

Evaluating physical activity is crucial for several reasons, ranging from promoting individual health to informing public health policy. Evaluating physical activity at an individual level can identify groups at risk due to sedentary behaviors, helping to implement targeted interventions to reduce inactivity, and consequently the burden of related health conditions. Physical activity evaluation at a population level helps track trends and identify areas where interventions are needed to promote a more active lifestyle. A variety of tools and methods have been proposed to measure physical activity, ranging from simple self-reported questionnaires to advanced wearable technology (e.g., pedometers, acetometers, fitness trackers), GPS technology, heart rate monitoring, and others (4). The choice of tool for evaluating physical activity depends on the context, including the population being studied, the type of activity, and the resources available. Wearable technology and mobile apps offer convenient and accessible options for everyday tracking, while laboratory and field tests provide more accurate assessments for clinical and research purposes. One of the most popular self-reported tools is the International Physical Activity Questionnaire, in Short and Long Form (IPAQ-SF/IPAQ-LF). This tool was proposed in the late 1990s as a self-reported measure of physical activity suitable for assessing population levels of physical activity across countries (5). Several studies validating the IPAQ-SF have been conducted in various populations and age groups. While some limitations exist, such as overestimation of activity, the tool is considered as a valuable resource for assessing physical activity patterns in the population and has been recommended as a reliable and cost-effective method to assess physical activity levels (6, 7). In Greece, IPAQ-SF tool has been validated in 218 health science students, aged 19–29 years. IPAQ-Gr was found to have acceptable reliability properties and high repeatability for total and vigorous physical activities, and good for moderate and walking physical activities (8). However, this tool has been extensively used in adults older than 30 years old and of various social domains and physical activity levels, without any indications about its reliability and validity. Therefore, the aim of this study was to test the validity of IPAQ-SF among Greek male and female adults, of all age groups.

MATERIALS AND METHODS

Setting

Participants were recruited from the general population, on a feasibility sampling basis from their neighbourhoods or public places and offices, with a special effort avoiding any selection bias. Participants who had a recent injury or other exercise-related problems were excluded from the sampling.

Sample

The validation study included 134 adults, aged 18–65 years; recruitment was performed within this age group, as proposed by the IPAQ inventors (5). Mean (standard deviation) age of the 94 participating males was 39.2 (11.1) years, and of the 35 females 34.4 (12.5).

Bioethics

All participants were informed about the study's objectives, data collection methods, and their rights regarding data privacy, in compliance with relevant ethical guidelines and data protection laws. After understanding all the procedures followed on and accepting participating in the study, participants were asked to sign an informed consent.

International Physical Activity Questionnaire

The International Physical Activity Questionnaires was developed in 1998 in Geneva, and comprises a set of 4 physical activity assessment questionnaires. The purpose of the questionnaires is to provide common instruments that can be used to obtain internationally comparable data on health-related physical activity. Translation from English is encouraged to facilitate the worldwide use of IPAQ*. This validation study utilized the IPAQ-SF (7-item) version, which is the most commonly used tool. The tool provides information on the time spent walking, in vigorous- and moderate-intensity activity and in sedentary activity, during the previous 7 days of the week. Participants were instructed to refer to all domains of physical activity (Table 1).

Scoring IPAQ

After performing data cleaning for missing and out-of-range values, the data from the IPAQ-SF questionnaire were analysed according to the physical activities recorded and time spent sitting per week. The collected data were used to estimate total weekly physical activity expenditure by weighting the reported minutes per week within each activity category by a metabolic equivalent (MET) energy estimate – 1 MET is defined as the amount of oxygen consumed while sitting at rest, equivalent to about 3.5 ml of oxygen per kg of body weight per minute (ml/kg/min) – of each category of activity (moderate, vigorous), as it has been proposed by Craig et al. (5). Both total volume and number of days/sessions are included in the IPAQ analysis algorithms. In particular, median values of MET-minutes/week were computed for walking, moderate- and vigorous-intensity activities using the following formulas: walking = 3.3 * (walking minutes) * (walking days); moderate = 4.0 * (moderate-intensity activity minutes) * (moderate days); vigorous = 8.0 * (vigorous-intensity activity minutes) * (vigorous-intensity days) (5). Then, a combined total physical activity MET-min/week was computed as the sum of (walking) + (moderate) + (vigorous) MET-min/week scores. The sitting question was not included as part of any summary score of physical activity. Moreover, participants were classified as inactive, minimally active and health enhancing physical activity (HEPA), using the number of days and duration in minutes engaged in moderate and vigorous physical activities, according to the guidelines of IPAQ scoring (5).

Validation Protocol

Regarding reliability, a translation and back-translation procedure was performed to ensure cultural adaptation. The original and

Table 1. *Short form of the International Physical Activity Questionnaire used in this validation study*

Item	
1	During the last 7 days, on how many days did you do vigorous physical activities like heavy lifting, digging, aerobics, or fast bicycling?
2	How much time did you usually spend doing vigorous physical activities on one of those days?
3	During the last 7 days, on how many days did you do moderate physical activities like carrying light loads, bicycling at a regular pace, or doubles tennis? Do not include walking.
4	How much time did you usually spend doing moderate physical activities on one of those days?
5	During the last 7 days, how many days did you walk for at least 10 minutes at a time?
6	How much time did you usually spend walking on one of those days?
7	During the last 7 days, how much time did you spend sitting on a weekday?

Source: Craig et al. (5)

*www.ipaq.ki.se

the Greek version of IPAQ-SF are presented in Electronic Supplementary Material. Regarding repeatability of IPAQ-SF, this has been tested in a previous study (8). Validity was assessed using fitness measures, in the same laboratory, and by experienced personnel, because physical activity levels have been strongly associated with cardiorespiratory fitness, and fitness measures presented in detail below have shown consistent reliability in a variety of studies (9). In particular, maximum oxygen uptake (VO_2 max), velocity of VO_2 max and heartbeat at 1 min at rest, and maximum power output were measured and used as a reference standard to compare the IPAQ-SF outcome, i.e., MET-minutes/week. A higher correlation indicated higher ability of the tool to assess true fitness levels (validity).

Treadmill Protocol and Gas Analysis

Incremental exercise test was performed on a motor-driven treadmill Woodway Pro, under the supervision of an experienced exercise-physiologist. Treadmill speed was calibrated while the subject ran at different running speeds by counting the time it takes for the completion of 20 treadmill revolutions with the aid of a measuring tape and a digital stopwatch, as recommended by the British Association of Sports Sciences (10). The incremental exercise VO_2 max test consisted of 5 minutes warm up at a speed which had been previously determined during familiarization with treadmill running and usually corresponded to a perceived exertion of less than 9 on the 15 point Borg Scale (11). On completion of the warmup subjects were left free for five minutes to perform the usual pre-exercise stretching routine. The VO_2 max test protocol consisted of 3-minute running stages with 1 km/h increments per stage while the treadmill elevation was kept horizontal. This protocol has been validated by other studies (12). Subjects ran until volitional exhaustion. VO_2 was measured using an open circuit system. The subjects breathed through a two-way, high velocity, low resistance no-rebreathing valve Rudolph #2700 mask. A dry gas meter and gases analysers (VacuMed O_2 17620 and CO_2 17620) were used to measure volume of expired air, oxygen and carbon dioxide. Barometric pressure and room temperature were measured. Strong verbal encouragement was given during the final stages of the VO_2 max test. The highest VO_2 value obtained during the incremental test was recorded as the subject's VO_2 max, as suggested by others (12, 13). The velocity (speed), at which the VO_2 max (vVO_2 max) was achieved, was also recorded, directly from the treadmill speed. Moreover, maximum power output (in watt) was measured using a cycle ergometer following standard protocol and we used the same procedures as in treadmill testing, as described above. Max power output, VO_2 max and vVO_2 max are direct measures of aerobic fitness and endurance (2). Moreover, using a heart rate monitor, and with the participant remaining seated in a relaxed state for at least 5 minutes, the resting heart rate (RHR) was recorded for 1 minute. Generally, a lower RHR is associated with a more efficient heart function, better overall cardiovascular health and fitness level. Further technical specification and performance properties of the activity fitness measures have been described elsewhere (2).

Other Measurements

In addition to the fitness testing measurements, basic socio-demographic characteristics, i.e., age and sex, potential engagement and duration in any type of physical activity, medical history, and anthropometric measurements – weight (kg), height (cm), and body fat were also measured. Body mass index (BMI) (kg/m^2) was then calculated and used for body weight status classification (normal BMI <25 kg/m^2, overweight 25–29.9 kg/m^2, and obesity >29.9 kg/m^2). Participants were also asked about their engagement in any type of physical activities during the past year.

Statistical Analysis

Continuous variables are presented as mean (standard deviation), and median and interquartile range (when skewed), whereas categorical variables are presented as frequencies (and percentages). As the IPAQ-SF data were not normally distributed, as evaluated using P-P plots, the nonparametric Spearman rho correlation coefficient was calculated as a primary measure of agreement between IPAQ-SF data and fitness measures (criterion validity). The categorical data was analysed by calculating the percent of the agreement (correct classification) to determine how many participants were classified in the same category by the method being compared, i.e., IPAQ-SF and tertiles of fitness measures. Cronbach's alpha coefficient was calculated to measure internal consistency of IPAQ tool (values above 0.7 are considered as acceptable reliability). Stata, version 17, statistical software was used for all calculations (Stata Corp., College Station, Texas, USA).

RESULTS

In Table 2 basic characteristics of the participants are presented. Most of the participants had normal BMI (61.24%), followed by individuals with overweight (31.01%) and obesity (5.43%). Moreover, almost 80% of the participants were at least minimally active (Table 2).

The analysis of the IPAQ-SF tool responses revealed that participants were engaged in vigorous activities on an average of 3 days per week, spending approximately 68.8 minutes per session. Moderate activities were reported for 2.6 days per week, with an average duration of 56.2 minutes per session. Walking was performed for an average of 3.7 days per week, for about 24.4 minutes each time. Overall, the median MET-minutes/week score was 2,751.0, with a significant difference between males and females ($p<0.001$) (Table 3).

The validation analysis (Table 4) revealed that the MET-minutes/week score derived from IPAQ-SF was moderately correlated with fitness measures, particularly VO_2 max (rho = 0.3096, $p<0.001$), vVO_2 max (rho = 0.2511, $p = 0.008$) and maximum power output (rho = 0.3493, $p = 0.185$). Correlations were stronger in females (VO_2 max: rho = 0.4359) compared to males (VO_2 max: rho = 0.2737). Among individuals with overweight/obesity, the correlation with VO_2 max (rho = 0.5201, $p<0.001$), vVO_2 max (rho = 0.4997, $p<0.001$) and maximum power output (rho = 1.000, $p<0.001$) was particularly significant, demonstrating better validity in this subgroup. Resting heart rate, heartbeat at 1 minute and maximum power output showed no significant correlations across the groups ($p > 0.100$). Internal consistency was good, as this was evaluated through Cronbach's alpha coefficient (0.779 for males and 0.909 for females).

Table 2. Participants' socio-demographic and anthropometric characteristics

Participants	134
Age (years), mean (SD)	37.94 (11.68)
Males, n (%)	98 (73.13)
Body mass index (kg/m^2), mean (SD)	24.31 (3.46)
Body weight status, n (%)	
Normal weight	79 (61.24)
Overweight	40 (31.01)
Obesity	7 (5.43)
Engagement in physical activities, n (%), yes	133 (99.25)

DISCUSSION

The present study aimed to evaluate the validity of the International Physical Activity Questionnaire-Short Form in a sample of Greek adults. To do that fitness indices were measured as the gold-standard, in an exercise physiology lab, in 134 men and women from the general population aged 18–65 years. Testing the validity of a physical activity tool is crucial, as it ensures the tool accurately measures what it is intended to measure. Validity is fundamental to the credibility and utility of the data collected, particularly in the research and clinical settings.

Our analysis revealed moderate validity of the IPAQ-SF as compared to fitness indices, which aligns with previous studies using IPAQ-SF in various populations and using the same methodology for validation (8, 14–17). This was supported by the fact that the MET-minutes/week score derived from IPAQ-SF was moderately correlated with fitness measures (Table 4). This could be supported due to the inherent limitations of self-reported physical activity questionnaire, including recall bias, social desirability bias, and difficulties in accurately estimating intensity and duration of activities. Moreover, the IPAQ-SF captures habitual physical activity across various domains (leisure, work, transportation, domestic), whereas fitness indices generally reflect physiological adaptations to regular activity, which may not correspond precisely to self-reported behaviour. Differences in individual interpretation of activity intensity and variability in energy expenditure across body compositions and age groups could also contribute to the observed moderate correlations. Greater validity was observed among females and overweight/obese participants. The validity of IPAQ varies significantly across different studies, with some studies showing high validity, while others failed to prove. A key similarity between our findings and previous studies is the issue of overreporting. Multiple studies have highlighted the tendency for physical activity to be overreported in self-administered physical activity questionnaires, like the IPAQ. This emphasizes the broader issue of self-reporting bias, where individuals may overestimate their physical activity due to questionnaire misinterpretation, recall bias, or social desirability bias. These discrepancies in reporting highlight the need to refine the IPAQ-SF, particularly to differentiate between actual sedentary behaviour and falsely reported activity levels. The observation that in females IPAQ-SF had higher validity may attribute to the fact that females are more health-conscious and motivated to monitor and report their physical activity levels accurately. Obese people also showed higher validity in our study as compared to normal weight, they may be highly motivated to track their physical activity as part of weight management or health improvement efforts. This motivation can translate into more careful and accurate responses on questionnaires. Nevertheless, incorporating objective measures, such as accelerometers, in conjunction with self-reports could help bridge the gap between perceived and actual physical activity among various groups of people.

In addition, validation studies often differ in several methodological aspects, which can influence their outcomes and interpretations. These differences include variations in study design, sample sizes, referent populations, including cultural context, the reference standard used, as well as data collec-

Table 3. Participants responses to the IPAQ-SF items

Participants	Mean (SD)	Median (Q1, Q3)
During the last 7 days, on how many days did you do vigorous physical activities like heavy lifting, digging, aerobics, or fast bicycling?	3.22 (1.77)	3.00 (2, 4)
How much time in minutes did you usually spend doing vigorous physical activities on one of those days?	68.80 (38.22)	60 (54, 90)
During the last 7 days, on how many days did you do moderate physical activities like carrying light loads, bicycling at a regular pace, or doubles tennis? Do not include walking.	2.60 (1.75)	3.00 (2, 3)
How much time in minutes did you usually spend doing moderate physical activities on one of those days?	56.28 (35.52)	60.10 (40, 75)
During the last 7 days, how many days did you walk for at least 10 minutes at a time?	3.75 (3.05)	3.50 (0, 7)
How much time in minutes did you usually spend walking on one of those days?	24.49 (25.17)	20.00 (0, 30)
During the last 7 days, how much time in minutes did you spend sitting on a weekday?	42.40 (18.55)	45.00 (28, 56)
MET-minutes/week		
All	3,091.23 (1,737.51)	2,751.00 (1,878, 3,831)
Males	3,146.87 (1,766.34)	2,751.00 (1,897, 3,746)
Females	2,939.77 (1,671.20)	2,718.00 (1,683, 4,000)

Results are presented as mean (standard deviation), median, first (Q1) and third (Q3) quartiles.

tion methods. These methodological differences are critical to consider when comparing validation studies, as they can significantly impact the interpretation and application of the results. A commonly observed methodological issue across the IPAQ validation studies lies in the demographic and sample characteristics of the recruited samples. For example, Rangul et al., involved adolescents aged 13–18 years (15), Fogelholm et al. and Kurtze et al. targeted younger male adults aged 21–43 and 20–39 years, respectively, despite the fact that IPAQ has been designed for people 18 to 65 years old (14, 17). Age differences likely influenced the variations in the strength of correlations between IPAQ-reported activity and fitness measures. For example, adolescents may show greater variability in self-reported data due to differing activity patterns, as reflected by the lower intraclass correlation coefficients observed in the Rangul et al. study, compared to adults (15).

Strengths and Limitations

Our study has several strengths that enhance its value. The comprehensive cultural adaptation of the IPAQ-SF through translation and back-translation ensures the tool is culturally relevant to the Greek population, improving its local applicability. The inclusion of a wide range of ages and both sexes allows for a more diverse assessment of the IPAQ-SF's validity across different demographics. Validating the tool against objective fitness measures like VO_2 max and vVO_2 max strengthens its credibility. Even though our study has some noteworthy strengths, there are also few limitations to point out. Indeed, while the sample size of 134 participants was sufficient in terms of achieving good statistical power (>87%), its generalizability to the broader Greek population, particularly among underrepresented groups (e.g., female sex, older adults), may be limited. In addition, this study focused only on the criterion validity of the Greek version of the IPAQ-SF, providing important evidence of its concurrent relationship with objective fitness measures. While other forms of validity (e.g., construct and content validity) were beyond the scope of the present study, future research could explore these aspects to further strengthen the overall validation of IPAQ.

CONCLUSIONS

In conclusion, while the IPAQ-SF remains a widely used tool for assessing physical activity, its limitations in accurately capturing true activity levels, particularly in populations prone to overreporting, highlight the need for refinement. Our findings, in line with those from previous studies, indicate that the questionnaire shows moderate validity when compared to objective fitness measures, but issues of overestimation persist. The tendency for self-report bias underscores the importance of incorporating objective methods, such as accelerometers, in future research to enhance the precision of physical activity measurements. Tailoring the IPAQ-SF to specific demographic and cultural contexts, as well as improving its capacity to differentiate between activity intensities, will be essential for improving its overall validity and reliability. These steps are crucial for the more accurate evaluation of physical activity and the development of more effective public health strategies.

Electronic Supplementary Materials

This article contains supplementary material available at https://doi.org/10.21101/cejph.a8446

Conflicts of Interest

None declared

REFERENCES

1. World Health Organization. Global action plan on physical activity 2018–2030: more active people for a healthier world [Internet]. Geneva: WHO; 2018 [cited 2024 Oct 14]. Available from: https://www.who.int/publications/i/item/9789241514187.
2. Garber CE, Blissmer B, Deschenes MR, Franklin BA, Lamonte MJ, Lee IM, et al. Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: guidance for prescribing exercise. Med Sci Sports Exerc. 2011 Jul;43(7):1334-59.
3. Lippi G, Sanchis-Gomar F. An estimation of the worldwide epidemiologic burden of physical inactivity-related ischemic heart disease. Cardiovasc Drugs Ther. 2020 Feb;34(1):133-7.
4. Ainsworth B, Cahalin L, Buman M, Ross R. The current state of physical activity assessment tools. Prog Cardiovasc Dis. 2015;57(4):387-95.
5. 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 Aug;35(8):1381-95.
6. Meh K, Jurak G, Sorić M, Rocha P, Sember V. Validity and reliability of IPAQ-SF and GPAQ for assessing sedentary behaviour in adults in the European Union: a systematic review and meta-analysis. Int J Environ Res Public Health. 2021 Apr 26;18(9):4602. doi: 10.3390/ijerph18094602.
7. Sember V, Meh K, Sorić M, Starc G, Rocha P, Jurak G. Validity and reliability of International Physical Activity Questionnaires for adults across EU countries: systematic review and meta analysis. Int J Environ Res Public Health. 2020 Sep 30;17(19):7161. doi: 10.3390/ijerph17197161.
8. Papathanasiou G, Georgoudis G, Papandreou M, Spyropoulos P, Georgakopoulos D, Kalfakakou V, et al. Reliability measures of the short International Physical Activity Questionnaire (IPAQ) in Greek young adults. Hellenic J Cardiol. 2009 Jul-Aug;50(4):283-94.
9. American College of Sports Medicine; Riebe D, Ehrman JK, Liguori G, Magal M, editors. ACSM's guidelines for exercise testing and prescription. 10th ed. Philadelphia: Wolters Kluwer; 2018.
10. Hale T, Armstrong N, Hardman A, Jakeman P, Sharp C, Winter E. Position statement on the physiological assessment of the elite competitor. 2nd ed. Leeds: White Line Press, 1988.
11. Borg G, Ottoson D, editors. Perception of exertion in physical exercise. Proceedings of an International Symposium held at The Wenner-Gren Center, Stockholm; 1985 Oct 3-5; Stockholm, Sweden. London: Macmillan Press Ltd; 1986.
12. Jones AM, Doust JH. A 1% treadmill grade most accurately reflects the energetic cost of outdoor running. J Sports Sci. 1996 Aug;14(4):321-7.
13. Coyle EF. Integration of the physiological factors determining endurance performance ability. Exerc Sport Sci Rev. 1995;23:25-63.
14. Fogelholm M, Malmberg J, Suni J, Santtila M, Kyröläinen H, Mäntysaari M, et al. International Physical Activity Questionnaire: validity against fitness. Med Sci Sports Exerc. 2006 Apr;38(4):753-60.
15. Rangul V, Holmen TL, Kurtze N, Cuypers K, Midthjell K. Reliability and validity of two frequently used self-administered physical activity questionnaires in adolescents. BMC Med Res Methodol. 2008 Jul 15;8:47. doi: 10.1186/1471-2288-8-47.
16. Mäder U, Martin BW, Schutz Y, Marti B. Validity of four short physical activity questionnaires in middle-aged persons. Med Sci Sports Exerc. 2006 Jul;38(7):1255-66.
17. Kurtze N, Rangul V, Hustvedt BE. Reliability and validity of the international physical activity questionnaire in the Nord-Trøndelag health study (HUNT) population of men. BMC Med Res Methodol. 2008 Oct 9;8:63. doi: 10.1186/1471-2288-8-63.

Received October 21, 2024
Accepted in revised form September 25, 2025