Reliability of the Serbian version of the International Physical Activity Questionnaire for older adults

Zoran Milanović1
Saša Pantelić1
Nebojša Trajković1
Bojan Jorgić1
Goran Sporiš2
Milovan Bratić1

1Faculty of Sport and Physical Education, University of Niš, Niš, Serbia; 2Faculty of Kinesiology, University of Zagreb, Zagreb, Croatia

Abstract: The purpose of this study was to determine the test–retest reliability of the International Physical Activity Questionnaire (IPAQ) for older adults in Serbia. Six hundred and sixty older adults (352 men, 53%; 308 women, 47%; mean age 67.65 ± 5.76 years) participated in the study. To examine test–retest reliability, the participants were asked to complete the IPAQ on two occasions 2 weeks apart. Moderate reliability was observed between the repeated IPAQ, with intraclass correlation coefficients ranging from 0.53 to 0.91. The least reliability was established in leisure time activity (0.53) and the most reliability in the transport domain (0.91). Men and women had similar intraclass correlation coefficients for total physical activity (0.71 versus 0.74, respectively), while the biggest difference was obtained for housework in men (0.68) and in women (0.90). Our study shows that the long version of the IPAQ is a reliable instrument for assessing physical activity levels in older adults and that it may be useful for generating internationally comparable data.

Keywords: questionnaire, elderly, IPAQ, physical activity

Introduction
Physical activity is defined as “any bodily movement produced by the skeletal muscle that results in energy expenditure”,1 and is considered to be one of the most important public health problems of modern society.2–3 Even though physical inactivity is not the only lifestyle or behavior factor associated with the development of chronic disease, it has received much interest in recent years.4 It is known that physical inactivity is linked with a number of chronic diseases, including coronary artery disease, stroke, hypertension, colon cancer, breast cancer, type 2 diabetes, and osteoporosis.5 In addition, the level of physical activity is often used as a parameter for monitoring and evaluating public health and is almost always associated with health status.1 This monitoring is especially important for older people over 60 years in programs aimed at preventing disease and inactivity and reducing mortality rates. It has been confirmed that older people are less physically active than other age groups in the population.6,7 Lower physical activity levels contribute to deterioration of health status, thus creating a serious public health issue given that the number of older adults continues to increase.

Physical activity patterns can be determined in a number of ways, including direct calorimetry, ingestion of doubly labeled water, and use of motion sensors, accelerometers, heart rate monitors, or oxygen consumption meters, as well as direct observation of gross movements by a trained observer or assessment of food intake.8 The validity
of both accelerometers and pedometers has been confirmed, and these devices have been used successfully for determination of physical activity levels in older adults. However, questionnaires are much cheaper than any direct method, and can show tendencies and habits regarding physical activity in older adults. This primarily refers to the type of physical activity that the elderly are involved in (e.g., jogging, walking, aerobics, cycling), as well as the frequency and duration of activity. The questionnaire method can provide additional information on what type of exercise they would prefer to do in their leisure time. There are quite a number of questionnaires available for measurement of physical activity, but only a few conform to international questionnaire standards when translated into different languages, and enable comparison of results between countries and people.

Several such questionnaires have addressed physical activity in people older than 60 years. It is essential to ensure that each localized version of the International Physical Activity Questionnaire (IPAQ) is reliable enough for the country it has been adapted for. Recall of physical activity patterns is a complex cognitive process that can generate errors with regard to interpretation of questions or cultural differences concerning activities and terminology.

In 1998–1999, an international consensus group developed the IPAQ as an instrument for cross-national monitoring of physical activity and inactivity. Eight different forms of this questionnaire have been developed to date and, according to the authors’ knowledge, 16 countries have confirmed its reliability and validity, either in the long form or short form. Its reliability was first confirmed in participants aged 18–65 years and then for those aged older than 65 years. Craig et al showed that the IPAQ was adequately reliable for different countries, sociocultural settings, and age groups, and reported intraclass correlation coefficients (ICCs) of 0.76 and 0.81 for the long and short versions, respectively. One study has indicated possible over-reporting on the IPAQ using the short form (last 7 days) telephone interview. The Chinese long and short versions of the IPAQ appear to have acceptable reliability when compared with other instruments used to assess physical activity in epidemiological studies, although the short version has been reported to underestimate the energy expenditure of total and moderate physical activity. Hallal and Victora found that the short version of the IPAQ may systematically underestimate physical activity levels in older adults. In another study, Rutten et al suggested that more research is needed to improve the quality of the IPAQ, especially for telephone interviewing.

Compared with other questionnaires that evaluate leisure physical activity levels, the IPAQ questionnaire seems to be much more comprehensive because its questions cover more activities from various aspects of life, such as work-related, transport, housework/gardening, and leisure activity. Further, it gives information about the frequency, duration, and intensity of exercise. The results regarding the development of the IPAQ are promising, but there is a need for further research to determine its reliability in more countries and in other age groups.

The aim of this study was to determine the test–retest reliability of the IPAQ in older adults in Serbia. There are limited data on physical activity levels in older Serbians, so we felt a need to develop and use an IPAQ translated into the Serbian language. In addition, we needed a clearer picture of physical activity levels in older adults that could be compared with those obtained in other countries. Future studies should address the reliability and validity of the IPAQ in different age groups to determine if it can be used without age restriction. We assumed that the long version of the IPAQ would be reliable for Serbian participants older than 60 years.

Materials and methods

Participants

Of 1,288 participants aged older than 60 years selected for testing, 700 were randomly selected to complete the IPAQ again to determine its reliability. All participants were from South-East Serbia, one of the five statistical regions in Serbia. Forty participants were excluded from further analysis because of data inconsistency, illness, or absence from home. The IPAQ recommendations on data cleaning and processing were followed for inconsistency of data. The remaining 660 participants, of whom 352 (53%) were men and 308 (47%) were women, were included in the final analysis. This sex ratio reflects that for older men and women in the Republic of Serbia. Participants were recruited via flyers handed out in senior centers and other community locations for older adults. A recruitment announcement was also circulated by email and e-newsletter to university staff with a request to pass the study information on to eligible friends and relatives. The general descriptive and demographic parameters of the participants are shown in Table 1. There were no statistically significant differences between the participants included in the study and those excluded from it. The study inclusion criteria were: being a native speaker of the Serbian language, being able and willing to provide self-reported information, age 60–80 years, and
being physically independent (able to walk 20 feet without assistance or rest). Volunteers who were recovering from an acute illness and those who had hearing and/or visual impairment were excluded.

Participation in this research was voluntary and participants could withdraw at any time. The protocol of the study, designed in accordance with the tenets of the revised Declaration of Helsinki, was approved by the ethics committee of the Faculty of Sport and Physical Education, University of Niš. The participants were given specific instructions for filling out the questionnaire as well as an outline of the potential benefits of such research for their age group. Testing of the participants was conducted from October through December 2011. Project team members, completed a 3-day training course before they started collecting data. This training was designed to enable the investigators to explain the possible ways of describing differences between walking, moderate activity, and vigorous activity to the participants. During training, we also provided information regarding the differences between work-related and housework or gardening domains. This training was done in accordance with IPAQ guidelines. A standard interview was conducted with potential participants individually or in small groups, either in their homes or in activity centers for older persons, after which data collection was started. Participants who could not complete the questionnaire by themselves were helped by skilled persons selected for the purpose of this research.

**Serbian version of the IPAQ**

We used the long version of the IPAQ for self-evaluation, which canvasses four types of physical activity, ie, work-related, transport, housework/gardening, and leisure activity. It also contains questions about sedentary habits. A Serbian version of the IPAQ (a long self-administered version canvassing physical activity in the past 7 days) for self-evaluation of physical activity was prepared by SP and MA and approved by the study group for implementation. The IPAQ was translated from English by two independent experts who were familiar with this type of questionnaire. The accuracy of the questionnaire was checked, and certain differences in translation were removed. The Serbian version of the IPAQ was then back-translated into English by a third independent person to establish the quality of translation.

**Procedures**

Data were collected from the participants twice, with an interval of 2 weeks between the first and second IPAQ test, which is considered a sufficient time period to eliminate the influence of the first test responses on the results of the retest.26 The participants were asked to report their physical activity in the previous 7-day period. In each of the four test domains, the participants recorded both the number of days and amount of time spent each day performing vigorous or moderate intensity activities and walking. Calculations for each item were done separately (vigorous activity, moderate activity, and walking) according to the instructions for use of the IPAQ (International Physical Activity Questionnaire, 2005) in order to determine the amount of time the participants spent performing the defined types of physical activity. A metabolic equivalent task (MET) value was calculated for each item, which is the ratio of the metabolic rate (measured in milliliters of oxygen per minute per kilogram body weight) to the metabolic rate associated with sitting. The values are rounded to the nearest whole number of METs.

Vigorous activities are defined as activities in which participants breathe much harder than normal, eg, lifting heavy items, digging, heavy construction work, or climbing stairs. Moderate physical activities are those in which a person is

---

**Table 1** General descriptive and demographic parameters of the subjects expressed as the mean (± standard deviation) or number (%)*

<table>
<thead>
<tr>
<th></th>
<th>Men n=352</th>
<th>Women n=308</th>
<th>Total n=660</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>67.53±5.75</td>
<td>67.80±5.77</td>
<td>67.65±5.76</td>
</tr>
<tr>
<td>Body height (cm)</td>
<td>176.09±7.50</td>
<td>165.42±24.17</td>
<td>171.08±18.23</td>
</tr>
<tr>
<td>Body mass (kg)</td>
<td>81.36±12.25*</td>
<td>69.57±11.18*</td>
<td>75.82±13.14</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>26.16±3.32</td>
<td>25.57±4.00</td>
<td>25.88±3.66</td>
</tr>
<tr>
<td>Marital status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>264 (75%)</td>
<td>209 (68%)</td>
<td>473 (72%)</td>
</tr>
<tr>
<td>Unmarried</td>
<td>88 (25%)</td>
<td>99 (32%)</td>
<td>187 (28%)</td>
</tr>
<tr>
<td>Residence</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>218 (62%)</td>
<td>160 (52%)</td>
<td>378 (57%)</td>
</tr>
<tr>
<td>Rural</td>
<td>134 (38%)</td>
<td>148 (48%)</td>
<td>282 (43%)</td>
</tr>
<tr>
<td>Employment status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employed</td>
<td>113 (32%)</td>
<td>28 (9%)</td>
<td>141 (21%)</td>
</tr>
<tr>
<td>Not employed</td>
<td>239 (68%)</td>
<td>280 (91%)</td>
<td>519 (79%)</td>
</tr>
<tr>
<td>Smoker status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smoker</td>
<td>144 (41%)</td>
<td>83 (27%)</td>
<td>227 (34%)</td>
</tr>
<tr>
<td>Non-smoker</td>
<td>208 (59%)</td>
<td>225 (73%)</td>
<td>433 (66%)</td>
</tr>
</tbody>
</table>

*Comparison between females and males (P<0.05); |data are presented as the mean ± standard deviation for quantitative variables, and as the number and percentage of the participants in different categories for categorical variables.

Abbreviation: BMI, body mass index.
breathing somewhat harder than normal and may include tasks such as carrying light loads. Walking is not considered to be a moderate physical activity. Vigorous and moderate activities are those lasting for at least 10 minutes continuously.

Statistical analysis
The data obtained from the male and female participants were analyzed separately using Statistical Package for Social Sciences SPSS version 18.0 software (SPSS Inc., Chicago, IL, USA). Data for demographic and lifestyle characteristics are shown as the mean ± standard deviation, ie, as a percentage of participants in different categories. Physical activity levels on test–retest are shown as the median ± quartile range, where 95% confidence limits for medians were calculated according to Bonett and Price.28 The test–retest reliability of the IPAQ was assessed by the ICC, (one-way random-effects model) with the 95% confidence interval (CI).

Results
Men and women in the study had similar mean body mass index values (26.16±3.32 kg/m² and 25.57±4.00 kg/m², respectively) and were of similar mean age, but were significantly different in their body mass (P<0.05, Table 1). The sex distribution in terms of marital status was similar, with 75% of men and 68% of women being married (P>0.05). More participants were from urban areas than from rural areas (62% versus 38% for men; 52% versus 48% for women), but the sex differences were the same (10%) for both the urban and the rural area without statistically significant differences. More men than women were employed (32% versus 9%, respectively). Forty-one percent of the men were smokers, while only 27% of the women smoked.

Table 2 shows that both men and women were physically active according to their total physical activity level (2,253±202 MET·min/week on the first day of testing and 2,228±216 MET·min/week for day 14). Total walking activity was 18% higher in men than in women, but the difference was not statistically significant (P>0.05). Total physical activity levels were similar between the first and second tests (4,509±162 MET·min/week versus 4,470±209 MET·min/week, respectively), which confirms the objectivity of the data obtained.

The ICC for all the domains examined between days 1 and 14 ranged between 0.53 and 0.91, which was statistically significant (P<0.05). The least reliability was established for leisure activity and the greatest for the transport domain (Table 2). Males and females had similar ICC values for total physical activity (0.71 versus 0.74), with the biggest difference found for housework (0.68 in males and 0.90 in females). Based on the results and the 95% CIs obtained, we would not expect reliability results lower than 0.42 or greater than 0.94 for men aged 60–80 years, or values lower than 0.53 or greater than 0.96 for women in the same age group.

Table 2 Test–retest reliability for the Serbian IPAQ (MET·min/week) on days 1 and 14 for men and women

<table>
<thead>
<tr>
<th>IPAQ (MET·min/week)</th>
<th>Men (n=352)</th>
<th>Women (n=308)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Test</td>
<td>Retest</td>
</tr>
<tr>
<td>Work-related</td>
<td>0.00±0.00</td>
<td>0.00±0.00</td>
</tr>
<tr>
<td></td>
<td>(0.00–0.00)</td>
<td>(0.00–0.00)</td>
</tr>
<tr>
<td>Transport</td>
<td>297.00±43.02</td>
<td>214.50±32.23</td>
</tr>
<tr>
<td></td>
<td>(231.00–396.00)</td>
<td>(198.00–297.00)</td>
</tr>
<tr>
<td>Housework/gardening</td>
<td>845.00±137.88</td>
<td>742.50±105.49</td>
</tr>
<tr>
<td></td>
<td>(660.00–1,140.00)</td>
<td>(580.00–945.00)</td>
</tr>
<tr>
<td>Leisure activity</td>
<td>198.00±69.64</td>
<td>66.00±48.78</td>
</tr>
<tr>
<td></td>
<td>(99.00–396.00)</td>
<td>(0.00–156.75)</td>
</tr>
<tr>
<td>Total walking</td>
<td>495.00±58.89</td>
<td>363.00±59.71</td>
</tr>
<tr>
<td></td>
<td>(396.00–627.00)</td>
<td>(297.00–495.00)</td>
</tr>
<tr>
<td>Total moderate</td>
<td>1,440.00±201.12</td>
<td>1,260.00±120.49</td>
</tr>
<tr>
<td></td>
<td>(1,140.00–1,800.00)</td>
<td>(960.00–1,440.00)</td>
</tr>
<tr>
<td>Total vigorous</td>
<td>0.00±8.00</td>
<td>0.00±8.20</td>
</tr>
<tr>
<td></td>
<td>(0.00–0.00)</td>
<td>(0.00–0.00)</td>
</tr>
<tr>
<td>Total PA</td>
<td>2,556.00±240.45</td>
<td>1,953.00±185.04</td>
</tr>
<tr>
<td></td>
<td>(2,110.61–3,126.64)</td>
<td>(1,709.00–2,383.50)</td>
</tr>
</tbody>
</table>

Note: 95% CI for median.
Abbreviations: ICC, intraclass correlation coefficient; CI, confidence interval; PA, physical activity; IPAQ, International Physical Activity Questionnaire-Serbian version; MET, metabolic equivalent task.
However, it should be emphasized that these recommendations apply only if the methodology presented in this paper is being used.

Discussion

In studies using the IPAQ in older adults, the ICCs were 0.41 in the Japanese population and 0.37 in the US population. In our study, the ICCs were higher (0.73) compared with the abovementioned studies, which leads us to the conclusion that the reliability of the Serbian version of the IPAQ is acceptable. Moreover, if we consider each domain separately, we obtained moderate to high reliability of the Serbian version of the IPAQ in older adults. Craig et al concluded that the long version of the IPAQ has high test–retest reliability. Compared with the long version, the short version has shown less reliability, especially in older adults. The reason for this is the absence of light physical activity such as walking in the short version, and walking has been considered to be the main activity of older adults.

Previous research on this subject has demonstrated reliability values higher than 0.70. The highest correlation coefficient (0.87) was found in an investigation of the Greek (participant age range 19–29 years) and Japanese (mean age 33.8 years) populations, and in the Chinese population (0.84; mean age 65.2 years), while the reliability in 12 further countries was about 0.76. There are studies in which the reliability coefficient was lower, ranging from 0.30 to 0.62 in Norway to 0.54 in Switzerland, while the lowest coefficient was found in research involving eight European countries (0.14–0.58).

Based on the findings of the research reported to date, it can be stated that there is a great variability in results concerning the reliability of the IPAQ. It has been demonstrated that the interval between the first and second tests significantly affects the reliability of the results. A period shorter than 8 days is associated with a higher reliability coefficient, whereas a period of 2–3 weeks significantly reduces the reliability. Given that the current research had a longer interval between the two tests, a moderate test–retest reliability coefficient was expected.

Our study yielded some significant results with regard to the work-related activities of men and women older than 60 years. Although their test scores were identical (zero), we obtained unequal ICC values (0.64 in males and 0.85 in females, respectively). The explanation for this could lie in the unequal work status of men and women older than 60 years. According to legislation in the Republic of Serbia, women retire at 60 years and men at 65 years, which is reflected in our finding that only 9% of women in this study were employed compared with 32% of men. Given that the majority of women (91%) did not answer the questions on work-related activity because they were not employed, it was expected that their correlation coefficient would have been much higher than for men.

Objective measures vary in different questionnaires and therefore in their primary outcomes. In addition, objective measures are not appropriate for large numbers of participants, especially if the testing is being done at different locations. Bearing in mind that people over 60 years of age are mostly bound to their place of residence and that attending testing premises may be an issue, the authors believe that the IPAQ and its methods and protocol are extremely valuable and practical, and meet the need for each country to monitor physical activity levels in the population. This is very important because the IPAQ evaluates physical activity and the results obtained can be compared with similar research performed in the future or in other countries. Compared with other methods used to assess physical activity (doubly labeled water, respiratory chambers, heart rate monitoring), the IPAQ is much cheaper and allows study of larger populations. Our study also shows that the IPAQ could be used in both developing and developed countries. It is essential to ensure that each localized version of the IPAQ is reliable for the country in which it has been adapted, given that recall of physical activity is a complex cognitive process that can generate errors from interpretation of questions as well as cultural differences in activity and terminology.

A limitation of this study is that the criterion validity of the Serbian version of the IPAQ was not determined to enable a comparison between pedometer-measured steps and the data (MET·min/week) from the IPAQ. Future studies should focus on determination of the reliability and validity of the IPAQ for different cultures, nationalities, and sociocultural settings. In previous research on reliability and validity, the most widely studied sector of the population has been middle-aged adults. Therefore, the current study is especially important because it extends the use of the IPAQ to older adults. However, to the best of the authors’ knowledge, the reliability and validity of this questionnaire is still not widely explored in the adolescent population, and this could be an area of future research. Further studies are also needed to establish the reliability of the Serbian IPAQ when used in other age groups.

In conclusion, the Serbian long version of the IPAQ was used to measure physical activity levels in older adults (>60 years). The IPAQ is acceptable for evaluation.
of different types of physical activity, their intensity, and total physical activity levels in older adults. Our study demonstrates that the IPAQ has adequate reliability for assessment of physical activity levels and that it may be a useful instrument for generating internationally comparable data.

Acknowledgment
This research was carried out as part of a project that was financially supported by the Ministry of Science of the Republic of Serbia, entitled “Physical activity and the fitness component of older adults” (number 179056), approved in 2010 and carried out by the Faculty of Sports and Physical Education at the University of Nis, Serbia.

Disclosure
The authors declare that they have no conflict of interest in relation to this paper.

References