J. Phys. Ther. Sci.
27: 3215–3221, 2015

Original Article

The Turkish version of the pregnancy physical activity questionnaire: cross-cultural adaptation, reliability, and validity

Ozge Celiker Tosun[1], Ulas Solmaz[2], Atalay Ekin[2]*, Gokhan Tosun[2], Ebru Kaya Mutlu[3], Emre Okyay[4], Mehmet Adiyeke[2], Cenk Gezer[2], Emre Mat[2], Mehtap Malkoc[1]

[1] *School of Physiotherapy, Dokuz Eylul University, Turkey*
[2] *Department of Obstetrics and Gynecology, Tepecik Training and Research Hospital: Izmir, Turkey*
[3] *Department of Physiotherapy and Rehabilitation, Faculty of Health Sciences, Istanbul University, Turkey*
[4] *Department of Obstetrics and Gynecology, Dokuz Eylul University, Turkey*

Abstract. [Purpose] The aim of this study was to translate the Pregnancy Physical Activity Questionnaire, adapt it for use with Turkish subjects and determine its reliability and validity. [Subjects and Methods] The Pregnancy Physical Activity Questionnaire was translated into Turkish and administered twice at 7–14-day intervals to pregnant women to assess the test-retest reliability. Cronbach's α was used for internal consistency, and the inter-rater correlation coefficient was used to calculate the test-retest reliability. The Turkish Short Form 36 Health Survey (SF-36) and the International Physical Activity Questionnaire were used to estimate validity. [Results] The internal consistency during the first and third trimesters of pregnancy was excellent, with Cronbach's α values of 0.93 and 0.95, respectively. The mean interval between the two assessments was 11.1 ± 2.1 days. The correlation coefficient between the total activity measured by the Turkish version of the Pregnancy Physical Activity Questionnaire and the International Physical Activity Questionnaire estimates of the total metabolic equivalent were fair to poor during the first, second, and third trimesters of pregnancy (r = 0.17, r = 0.17, r = 0.21, respectively). The Turkish version of the Pregnancy Physical Activity Questionnaire showed fair correlations with the Short Form 36 Health Survey physical component score (r = −0.30) and mental component score (r = −0.37) for the first trimester of pregnancy. [Conclusion] The Turkish version of the Pregnancy Physical Activity Questionnaire was found to be reliable and valid for assessing a pregnant woman's physical activity.

Key words: Pregnancy Physical Activity Questionnaire, Reliability, Turkish

(This article was submitted Jun. 15, 2015, and was accepted Jul. 16, 2015)

INTRODUCTION

Physical activity during pregnancy is important for the health of the mother and child and may reduce the risk of adverse maternal, fetal, and neonatal outcomes. At least 30 minutes of moderate activity or 8,000 steps/day, equivalent to approximately 7.5 metabolic equivalent (MET) hr/wk, is recommended by the American College of Obstetricians and Gynecologists for beneficial results[1].

Physical activity includes occupational, sports, and exercise activities as well as household and caregiving duties. The measurement of physical activity is typically divided into patient-reported outcomes, such as questionnaires or exercise diaries, and objective assessments using equipment such as an accelerometer or a pedometer[2–4]. The questionnaire is a feasible tool for assessing physical activity in large populations for various applications, including epidemiologic research or public health surveillance[5, 6]. It is easy to administer, relatively inexpensive, and noninvasive, and it allows for accurate estimation of the intensity and for identification of the type of physical activity (e.g., exercise, household, transportation). A major disadvantage of questionnaires and diaries is the potential for recall bias. However, studies have demonstrated reasonable reliability and validity for such measures and significant correlations of these measures with pregnancy outcomes. In contrast, objective measures, such as accelerometers or pedometers, do not suffer from self-reporting error, but these devices have limitations. Some of these tools cannot accurately measure physical activities involving upper body movement, pushing or carrying a load, stationary exercise such as bicycling and weight lifting, and water activities. Also, they are unable to determine the intensity of the physical activity level (e.g., mild or moderate physical activity)[7, 8].

In general, the majority of patient-reported outcomes reflect the language and social culture of the community in

*Corresponding author. Atalay Ekin (E-mail: atalayekin@hotmail.com)

which they were established. Therefore, it is necessary to translate and culturally adapt the scoring systems for use in other communities. Several questionnaires have been developed to measure physical activity among pregnant women[2]. In addition, the self-administered Kaiser Physical Activity Survey was modified from its original format for use among pregnant women[9]. The third Pregnancy Infection and Nutrition Study (PIN3) Physical Activity Questionnaire[10] was developed to assess moderate to vigorous physical activity among pregnant women in the past week and the perceived intensity of each physical activity type. The Pregnancy Physical Activity Questionnaire (PPAQ) is being used in an ongoing cohort study (STORK) of pregnant women[8]. This questionnaire includes questions on trimester-specific physical activity, and it measures four areas: transportation, occupation, household activities, child-care activities, and sport/exercise. For all three questionnaires, absolute intensity can be assigned to each activity using metabolic equivalent values[11]. Additionally, some physical activity questionnaires, such as the International Physical Activity Questionnaire (IPAQ), are suited for both pregnant and non-pregnant women[12].

The development of the PPAQ in 2004 resulted in a widely available tool for assessing physical activity in pregnant women[13]. The PPAQ includes activities that are important discriminators of physical activity among pregnant women, and it is able to measure the duration, frequency, and intensity of the total activity patterns in pregnant women. The PPAQ is short in length, self-administered, and easily understood by respondents in a variety of settings, making it useful for epidemiologic research. It has been translated into English, Japanese, French, Vietnamese, and Portuguese and has been psychometrically tested[13–17]. Cross-cultural adaptation into Turkish may contribute to a better understanding of the measurement properties of the PPAQ. It is important to generate a Turkish version of the PPAQ (PPAQ-Turkish) because there is no outcome measure for the assessment of physical activity in Turkish-speaking pregnant women.

The main hypothesis of this study was that it would be possible to translate and culturally adapt the PPAQ to Turkish so that it would be easily understood by Turkish-speaking individuals. Additionally, we hypothesized that the PPAQ-Turkish would show adequate internal consistency and test-retest reliability, acceptable construct validity compared with the Short Form 36 Health Survey (SF-36) physical component scale, and a weak correlation with the mental component scale of the SF-36. The purpose of this study was to translate and culturally adapt the PPAQ for use with a Turkish population and to determine its validity and reliability.

SUBJECTS AND METHODS

The PPAQ was translated into Turkish and culturally adapted in accordance with the stages recommended by Beaton et al[18]. Three Turkish individuals with a good command of English undertook the literal and conceptual translation of the PPAQ. Two of the translators were doctors with knowledge of this subject, and one was a tourism professional with no knowledge of this subject. The native language of all translators was Turkish, and all were fluent in English. The translations were completed independently. The translations were compared and reviewed by an individual who spoke both languages and who highlighted any conceptual errors or inconsistencies to produce the first Turkish translation. Subsequently, two native English speakers with a good command of Turkish separately translated the finalized Turkish version back into English. Both translators were unaware of the purpose of the study and had no access to the original document. The back-translated version of the PPAQ was compared with the initial English version by a committee of four translators as well as the original author. This committee approved the Turkish version and titled it the PPAQ-Turkish. Once the measure was approved, a pilot test was conducted with 20 pregnant women who fulfilled the eligibility criteria of the study to determine the comprehensibility of the PPAQ-Turkish. When the pregnant women had completed the questionnaire, they were interviewed by physical therapists. The participants were questioned about any difficulties in understanding the questions immediately after completing the form. The questions that were difficult to understand were noted, and the participants were asked for their recommendations for revising the questions.

This longitudinal study testing the reliability and criterion validity of the PPAQ-Turkish was conducted in the Faculty of Medicine, Dokuz Eylul University, and the Department of Obstetrics and Gynecology, Tepecik Education and Research Hospital. We selected two older and larger hospitals in Izmir because the validity and reliability could be assured by recruiting a wide range of participants of varying ages, cultures, income levels, working statuses, pregnancy trimesters, education levels, and residential areas. A total of 434 subjects representing the first, second, and third trimesters of pregnancy were recruited during a pregnancy check-up at the hospitals. Of these women, 328 fulfilled the criteria outlined below. The eligibility criteria were as follows: 1) over 16 but under 40 years of age and 2) able to read and write in Turkish. The exclusion criteria included 1) contraindication or inability to exercise or undertake physical activity; 2) medical or obstetric contraindication to exercise and physical activity, including hemodynamically significant heart disease or restrictive lung disease; 3) incompetent cervix (cerclage); 4) multiple gestations; 5) severe anemia; 6) chronic bronchitis; 7) type 1 diabetes; 8) orthopedic limitations; 9) poorly controlled seizure disorder; 10) poorly controlled hyperthyroidism; 11) heavy smoker; and 12) hypertension or chronic renal disease.

The PPAQ is a semiquantitative questionnaire that asks respondents to report about the time spent participating in 32 activities, including household/caregiving activities (13 activities), occupational activities (five activities), sports/exercise (eight activities), commuting (three activities), and inactivity (three activities). For each activity, respondents are asked to select a response category that is closest to the amount of time spent per day or week during the respondent's current trimester. The durations ranged from 0 to 6 or more hours per day and from 0 to 3 or more hours per week. An open-ended section at the end of the PPAQ allows each respondent to add activities not already listed. Self-administration of the PPAQ takes approximately 10–15 min.

The types of sport and exercise activities assessed in the PPAQ include walking, jogging, prenatal exercise classes, swimming, and dancing. To calculate weekly energy expenditure using the PPAQ, the duration of time spent in these exercise activities is multiplied by specific intensities (i.e., MET values), and scores are expressed as MET-hours per week (MET-hr/wk). The self-reported time spent on each activity was multiplied by activity intensity (in METs) to arrive at a measure of average daily energy expenditure (MET-hr/day). Activity intensity was based on field-based measurements of pregnant women and the 2000 version of the compendium-based MET values[19]. Additionally, each activity was classified by intensity, sedentary (<1.5 METs), light (1.5–3.0 METs), moderate (3.0–6.0 METs), or vigorous (≥6.0 METs), and the average MET-hours per week expended at each intensity level were calculated. Activities were also classified by type (household/caregiving, occupational, and sports/exercise), and the average MET-hours per week spent in each activity type was calculated[13].

The SF-36 was used to establish a health profile consisting of scores for eight items, with each score transformed to a score between 0 and 100 to identify the patient's physical and mental state. These eight items include physical functioning (PF), role limitations due to physical function (PF), bodily pain (BP), general health perceptions (GH), vitality (VH), social function (SF), emotional function (RE), and mental health (MH). The eight domain scores were aggregated into physical and mental component summary scores (SF-36 PCS and SF-36 MCS). The SF-36 PCS and SF-36 MCS, which were derived as the weighted sum of the item scores using the US standard SF-36 scoring algorithms, were considered the primary outcomes of this study. The Turkish version of the SF-36 has been shown to be valid and reliable[20].

The IPAQ was originally developed in 1998 to assess self-reported health-related physical activity in populations across several countries, and it was further validated by Craig and associates in 2003[21]. The IPAQ assesses physical activity across a variety of different domains including leisure time, domestic, work, and transport-related physical activity. For each domain, the questionnaire assesses walking, moderate activity, and vigorous physical activity performed for at least 10 consecutive minutes each day over 7 days. An average MET score was calculated for total physical activity performed per week as a continuous variable as follows: total physical activity in MET-minutes/week = sum of total [Walking + Moderate + Vigorous] MET-minutes/week. Individual MET scores for walking, moderate activity, and vigorous activity were calculated within each domain and combined to provide a total score using the following equation: total MET-minutes/week = MET level × minutes per day × days per week, where 1 MET is equivalent to the resting energy expenditure. The short and long forms of the Turkish version of the IPAQ have been shown to be reliable and valid in assessing physical activity[22].

The study consisted of two parts, a validity study and a reliability study. Participants were asked to complete the PPAQ-Turkish and the previously validated Turkish versions of the IPAQ and SF-36[20, 22]. The three questionnaires were given to the participants in random order using computer-generated random numbers. Physical therapists distributed the PPAQ-Turkish, PAQ, and SF-36 to the participants in the waiting room after a pregnancy check-up with a doctor. After each participant completed the questionnaire, a physical therapist checked for missing responses. Participants who missed an answer on the questionnaire were asked to give a reason. Any difficulty in understanding the question or incompatibility with their problem was noted. For the reliability study, two additional exclusion criteria were used: a reported change in the condition of the pregnancy between measurements and failure to complete the second PPAQ-Turkish questionnaire within the predetermined period of 7 to 14 days after completing the first PPAQ-Turkish questionnaire. A total of 90 subjects were eligible for inclusion in the reliability study, as 238 subjects did not return the second PPAQ-Turkish. Prior to the start of the study, all participants completed informed consent forms, which had been approved by the ethics committee at Dokuz Eylul University.

Reliability refers to the consistency of measurement and includes internal consistency and test-retest reliability. The homogeneity of the questions within the questionnaire (internal consistency) was assessed using Cronbach's α coefficient. An α of 0.7 was considered fair, 0.8 was considered good, and 0.9 was considered excellent internal consistency[23]. In this study, data for the pregnant women included in the first assessment of the PPAQ-Turkish were used to assess internal consistency. The test-retest reliability represents a scale's effectiveness in producing consistent results when administered on different days when an individual's status has remained stable[24]. The outcome measure was applied and then reapplied after 7 to 14 days. The results were compared for agreement by means of an intra-class correlation coefficient (ICC) to measure the test-retest reliability of the PPAQ-Turkish. Correlation values of 0.4 or greater were considered satisfactory (specifically, $r \geq 0.81$–1.0 was deemed excellent; 0.61–0.80, very good; 0.41–0.60, good; 0.21–0.40, fair; and 0.00–0.20, poor)[25].

Validity refers to the degree to which a study accurately reflects or assesses the specific concept that the researcher is attempting to measure. In this study, we examined three aspects of validity: construct, convergent/divergent, and content validity. Evidence for construct validity of the PPAQ-Turkish was provided by determining its relationship with the Turkish version of IPAQ and the SF-36. The SF-36 PCS was used to assess convergent validity. Evidence for divergent validity was provided by determining the relationship with the SF-36 MCS. Pearson correlation coefficients and their 95% confidence intervals were calculated to assess construct and convergent/divergent validity. Correlation values of 0.4 or greater were considered satisfactory (specifically, $r \geq 0.81$–1.0 was deemed excellent; 0.61–0.8, very good; 0.41–0.6, good; 0.21–0.4, fair; and 0–0.2, poor). Ceiling and floor effects in content validity occur where extreme scores are missing. The percentages of responders who had the lowest (i.e., none) or highest (i.e., 3 or more hours per day) scores in each subdomain on the PPAQ-Turkish were documented. Descriptive statistics (mean values, standard deviations, and quartiles) were calculated to determine the distribution and ceiling/floor effects, which were considered relevant if they were experienced by more than 30% of

subjects.

All statistical analyses were performed with the Statistical Package for the Social Sciences 21 (SPSS Inc., Chicago, IL, USA). This analysis included frequency counts and percentages for nominal variables, the measures of central tendency (means, medians), and the dispersion (standard deviations, ranges) for continuous variables. The Kolmogorov-Smirnov test was used to assess the distribution. The first and second assessments of the score were found to be normally distributed. The measurement properties analyzed in this study included internal consistency, test-retest reliability, construct validity, and ceiling and floor effects.

RESULTS

Based on the translators' comments, the questions were clear, concise, and easy to translate because the words of in the English language corresponded one-to-one with Turkish words; however, some words had Turkish equivalents that are not commonly used; e.g., items 18 and 19 in the original PPAQ refer to the use of a lawnmower. These items were not deleted or altered because there are times when a lawnmower is used, such as in local regions or in luxury housing. Some of the language was altered to improve the experiential equivalence. The only significant changes involved conversion from English measurements (gallons, pounds) to the metric equivalents (liters, kg) for item 33.

The instructions were found to have a problem in terms of experiential equivalence. The original instructions referred to "during this trimester", which was difficult for participants to understand. Pregnant women did not accurately recognize the divisions between trimesters. Therefore, this was changed to "during the last month". These considerations yielded a preliminary Turkish version of the PPAQ. The response method presented a problem in terms of technical equivalence. The original instructions stated, "Fill in the circles completely". However, this instruction was difficult, and it was modified to "Please check the box of the corresponding answer". We changed the structure of the original PPAQ to help the Turkish women understand the questions fully. We added the words "how much time do you usually spend . . ." to all of the questions and "It takes . . ." to all of the answers. Example: How much time do you usually spend preparing meals?

-None
- It takes less than 1/2 hour per day
- It takes 1/2 to almost 1 hour per day
- It takes 1 to almost 2 hours per day

No data were missing for any of the items in the pretest, and the response time was very short. These results suggest that the Turkish version of the PPAQ was easily understood and that responding was easy. The subjects required approximately 5–10 minutes to complete the PPAQ-Turkish. A total of 328 subjects completed all questionnaires in the first assessment, and 90 completed the second assessment for the test-retest reliability. Table 1 illustrates the demographic and clinical characteristics of the patients. The mean values (METs × hours/day) from the first and second measurements during the first, second, and third trimesters of pregnancy for total activity, activity intensity, and type of activity are shown in Table 2. The internal consistency of the total activity of the PPAQ-Turkish during the first and third trimesters of pregnancy was excellent, with Cronbach's α values of 0.93 and 0.95, respectively. During the second trimester of pregnancy, the internal consistency of the total activity of the PPAQ-Turkish was fair, with a Cronbach's α of 0.72. The mean interval between the two assessments was 11.1 ± 2.1 days. The test-retest reliability of the PPAQ-Turkish is presented in Table 2. The results were between 0.57 and 0.97 for the two subdomains for the first administration of the PPAQ-Turkish questionnaire.

Table 1. Demographics of study groups

Variable	Reliability study	Validity study
All participants (n)	90	328
Age (years)	28.5±5.5	28.3±4.6
Range	(18–40)	(18–40)
BMI (kg/m^2)	25.8±3.9	26.3±3.6
Parity		
0	36 (40%)	134 (40.9%)
≥1	54 (60%)	194 (59.1%)
Trimester		
First trimester (%)	20 (22.2%)	50 (15.2%)
Second trimester (%)	28 (31.1%)	141 (43%)
Third trimester(%)	42 (46.7%)	137 (41.8%)
Age (years) (mean±SD)		
First trimester	31.9±5.28	29.4±4.6
Second trimester	26.3±4.7	27.8±4.5
Third trimester	28.3±5.5	28.4±4.6

BMI: body mass index; SD: standard deviation
Values for the continuous variables are presented as the mean±SD. Values for the categorical variables are presented as the number/total number of cases (%).

Means, standard deviations, standard errors, and confidence intervals for each of the scores are displayed in Table 3. The correlation coefficient between the total activity of the PPAQ-Turkish and the IPAQ estimates of the total MET min^{-1} day^{-1} were reported as fair to poor during the first, second, and third trimesters of pregnancy ($r = 0.17$, $r = 0.17$, $r = 0.21$, respectively; $p < 0.05$). The correlation coefficients between the activity intensity of the PPAQ-Turkish and the IPAQ estimates for the light level of activity (MET-min^{-1} day^{-1}) were poor to good during the first, second, and third trimesters of pregnancy ($r = 0.32$, $r = 0.26$, $r = 0.29$, respectively; $p < 0.05$). Poor correlations were found between the PPAQ-Turkish and the subscores of the SF-36 (PF, $r = -0.22$; RP, $r = -0.53$; BP, $r = -0.21$; GH, $r = -0.20$; VT, $r = 0.30$; SF, $r = -0.38$; RE, $r = -0.58$; MH, $r = -0.28$; PCS, $r = -0.30$; MCS, $r = -0.37$ for the first trimester of pregnancy) (PF, $r = -0.14$; RP, $r = -0.42$; BP, $r = -0.19$; GH, $r = 0.14$; VT, $r = -0.21$; SF, $r = -0.23$; RE, $r = -0.25$; MH, $r = -0.14$; PCS, $r = -0.28$; MCS, $r = -0.22$ for the second trimester of pregnancy) (PF, $r = -0.35$; RP, $r = 0.05$; BP, $r = 0.29$; GH, $r = 0.10$; VT, $r = 0.06$; SF, $r = 0.04$; RE, $r = 0.03$; MH, $r = 0.02$; PCS, $r = 0.10$; MCS, $r = 0.06$ for the third trimester of pregnancy) (Table 3). Ceiling and floor effects and the number of items answered were identical during the

Table 2. Test-retest reliability of the components of the pregnancy physical activity questionnaire

	First trimester PPAQ-T n = 20				Second trimester PPAQ-T n = 28				Third trimester PPAQ-T n = 42			
	Reliability				Reliability				Reliability			
	T1	T2	ICC	Cronbah's alpha	T1	T2	ICC	Cronbah's alpha	T1	T2	ICC	Cronbah's alpha
Summary activity scores												
Total activity (light and above)	69.9±28.8	75.8±33.5	0.88	0.93	45.3±14.9	45.1±12.6	0.57	0.72	40.8±19.5	41.7±20.5	0.91	0.95
Intensity												
Sedentary (<1.5 METs)	10.9±6.5	12.3±6.8	0.90	0.95	8.6±5.5	9.5±5.2	0.95	0.97	5.4±4.1	5.7±4.3	0.93	0.96
Light (1.5–3.0 METs)	24.6±14.3	24.6±13.6	0.96	0.98	21.9±8.6	20.7±9.1	0.61	0.75	18.9±8.4	17.5±7.2	0.87	0.93
Moderate (3.0–6.0 METs)	27.9±20	27.5±23.9	0.91	0.95	9.7±7.5	9.3±6.9	0.97	0.98	11.6±7.8	12.1±8.2	0.89	0.94
Vigorous (≥6.0 METs)	6.2±6.5	8.3±7.3	0.74	0.85	5.1±6.5	5.2±6.4	0.65	0.79	4.9±6.4	4.5±6.7	0.93	0.96
Type												
Household/ caregiving	24.6±15.6	22.3±13.6	0.92	0.96	22.5±11.7	21.9±13.1	0.91	0.95	20.4±9.1	19.3±6.9	0.85	0.92
Occupational	24.3±26.4	27.4±31.1	0.88	0.94	6.6±12.6	5.2±8.9	0.92	0.96	1.1±2.7	1.0±2.6	0.90	0.94
Sports/ exercise	9.5±6.5	11.9±6.7	0.59	0.74	7.5±6.7	7.3±6.6	0.68	0.81	9.1±11.1	8.9±12.2	0.97	0.98

ICC: internal consistency; MET: metabolic equivalent; PPAQ-T: Pregnancy Physical Activity Questionnaire-Turkish; SD: standard deviation; T1: first test; T2: second test (retest)
Values for the continuous variables are presented as the mean±SD. Values for the categorical variables are presented as the number/ total number of cases (%).

test and retest assessments. None of the patients achieved the minimum or maximum overall score of the PPAQ-Turkish, implying the absence of floor and ceiling effects.

DISCUSSION

The most important finding of the present study was that the PPAQ-Turkish has good measurement properties, high reliability, and appropriate construct validity, indicating that it can be used to evaluate physical activity levels in pregnant Turkish women. The original version of the PPAQ was successfully translated and adapted to the Turkish language. The internal consistency analysis using Cronbach's α showed all questionnaires to be within the recommended range of values (0.70–0.95)[23]. Interestingly, the original version of the PPAQ questionnaire was not assessed in terms of internal consistency. In addition, validation studies in the literature for other PPAQ questionnaires in other languages that employed Cronbach's α did not evaluate internal consistency. The Cronbach's α revealed high internal consistency in our study.

The test-retest indicated adequate to excellent and very good reliability for the subscales and the PPAQ-Turkish questionnaire as a whole, respectively[25]. In the literature, the test-retest reliability of the overall PPAQ questionnaire has been excellent or very good (ranging from 0.61 to 0.94)[13–15], with the exception of the results reported by Ota et al.[16] (range, 0.88 to 0.94). Furthermore, with regard to subscales of the original version of the PPAQ, Chasan-Taber et al. found reliability of 0.83 for sports/exercise[13]. In our study, the reliability of the sports/exercise subscale was 0.61, similar to the results of Matsuzaki et al[15]. This may be because the first and second PPAQ were conducted at different times. The time interval between repeated measurements is an important factor when determining test-retest reliability. Reliability tends to be higher when a short time interval is used (7 days or less)[13], whereas short test-retest intervals carry the risk of patients' becoming familiar with the questions and simply answering based on memory of the first assessment. However, the risk is higher that the condition of the pregnant women would change over a longer period of time. Therefore, an interval of 7–14 days was chosen for the retest assessment to decrease the possibility of participants' remembering the questions. In addition, we believe that the conditions of the pregnant women would not be expected to change over this time period.

In recent studies, PPAQ validity has been investigated by determining the relationship of PPAQ results with an objective measure of physical activity[13–17]. Objective assessments compared with the PPAQ have included accelerometers and pedometers. However, although objective measures are not subject to self-report errors, they do have several limitations. Some of these tools cannot accurately measure physical activities involving upper body movement, pushing or carrying a load, and stationary exercise such as bicycling, weight lifting, and water activities[2]. Additionally, Lindseth and Vari[7]

Table 3. The mean values, standard deviations, median values, minimum and maximum values, standard errors, and confidence intervals for the outcomes

PPAQ	First trimester (n = 50)				Second trimester (n = 141)				Third trimester (n = 137)			
	Mean±SD	Median (min-max)	SE	95% CI	Mean±SD	Median (min-max)	SE	95% CI	Mean±SD	Median (min-max)	SE	95% CI
Summary activity scores												
Total activity (light and above)	48.4±24.2	15.4–119.7	3.42	41.56–55.3	47.3±24.9	2.3–162.2	2.1	43.1–51.5	42.1±20.5	7.4–113.2	1.75	38.5–45.4
Intensity												
Sedentary (<1.5 METs)	8.4±5.5	0.25–21.3	0.78	6.9–10	7.6±5.2	0.25–24.1	0.44	6.7–8.3	6.6±5.1	6–25.5	0.43	5.8–7.5
Light (1.5–3.0 METs)	19.9±11.4	6.6±46.7	1.61	16.9–23.3	20.9±9.4	0–54	0.79	19.3–22.5	19.1±8.5	2.4–47.9	0.72	17.6–20.6
Moderate (3.0–6.0 METs)	14.8±14	0–73.3	1.98	11.1–19.1	13.6±13.5	0.79	1.14	11.5–15.8	12.1±10.1	0.54	0.85	10.4–13.7
Vigorous (≥6.0 METs)	4.7±5.8	0–22.3	0.83	3.3–6.4	4.9±6.3	0.37	0.53	3.9–6.1	4.3±5.6	0–27.2	0.47	3.4–5.3
Type												
Household/caregiving	20.2±13.5	4.7–55.3	1.91	16.4–23.9	21.3±11.9	0.60	1.01	19.3±23.4	19.8±11.1	1.8–55.3	0.95	18.1–21.8
Occupational	9.8±16.3	6–84.6	2.31	5.2–14.7	7.7±13.5	0–76.1	1.14	5.6±9.9	5.4±10.1	0–61.5	0.86	3.8–7.2
Sports/exercise	8.3±7.1	0–26.1	1.00	6.62–10.2	8.8±9.7	0–56.7	0.82	7.4–10.5	8.1±8.9	0–41.2	0.76	6.6–9.7
IPAQ												
Total MET-min/week	1,058.8±1,084.1	33–4,752	153.3	781.4–1,374.6	1,192.6±2,048.3	33–1,6632	173.74	901.6–1,597.3	1,223.3±2,388.1	33–22,194	204.7	900.7–1,661.2
Light/walk MET-min/week	211.6±197.3	60–900	27.91	210.5–320.3	195.8±172.4	30–840	14.80	168.8–224.1	170.2±176.5	6–1,200	15.14	151.3–208.7
Moderate/walk MET-min/week	786±184.5	600–9,160	25.6	211.6–318.2	625.1±139.5	600–6,283	14.72	172.8–243.1	-	-	-	-
SF-36												
Physical Functioning	46.5±3.5	38.5–48.8	1.0	44.3–48.3	41.8±6.1	32–50.9	1.24	39.6–44.3	33.4±8.3	19.4–52.9	1.35	30.8–36.2
Role Physical	38.8±9.7	28–56.2	2.80	33.8–44.2	37.7±9.9	28–56–2	2.02	33.9–41.8	37.8±9.4	28–56.2	1.51	35–41.1
Bodily Pain	43.1±10.2	28.9–55.9	2.94	37.8–48.6	43.8±9.5	24.2–55.9	1.93	40.4–47.5	43.7±11.6	19.9–62.7	1.89	40.2–47.4
General Health	40.9±5.1	35.9–52.3	1.47	38.3–43.7	46.3±6.2	33.6–57.9	1.27	43.9–48.7	46.5±6.8	33.6–57.9	1.09	44.3–48.5
Vitality	42.4±6.4	34.9–56.2	1.83	39–45.9	43.2±7.7	32.5–53.8	1.57	40.3–46.3	48.9±8.6	32.5–65.6	1.39	46.8–51.9
Social Functioning	42.7±7.8	30–57.1	2.24	38.6–47.2	40±9.4	24.6–57.1	1.92	36.2–43.8	39.4±10	19–57.1	1.62	36.4–42.4
Role Emotional	35.11±14.5	23.7–55.3	4.19	27.2–43	33.4±12.4	23.7–55.3	2.52	28.5–38.6	35.3±12.4	23.7–55.3	2.01	31.7–39.2
Mental Health	39.3±8.4	23.7–52.7	2.43	34.6–44	40.9±6.7	27.2–52.7	1.36	38.1–43.4	42.8±8.9	27.7–64.1	1.45	40.4–45.7
Physical Component Score	42.6±3.9	36.7–50.3	1.14	40.6–44.8	42±5.9	30.8–52.9	1.20	39.7–44.4	38.3±6.5	27–58.1	1.06	36.5–40.4
Mental Component Score	38.9±9.6	29.2–54.3	2.78	34.1–44.2	38.9±8.2	29.5–53.8	1.67	35.6–42.3	43.9±9.4	37.9–66.9	1.52	41.2–47

CI: confidence interval; IPAQ: International Physical Activity Questionnaire; MET: metabolic equivalent; PPAQ: Pregnancy Physical Activity Questionnaire; SD: standard deviation; SE: standard error; SF-36: Short Form 36 Health Survey

Values for the continuous variables are presented as the mean±SD and median (min–max). Values for the categorical variables are presented as the number/total number of cases (%).

found that results from a pedometer were significantly correlated with results of a self-reported questionnaire. Evenson et al.[2] reported that five studies in their review included assessment of validity using another self-reported measure of physical activity to compare against the questionnaire, including a clinical review, a 7-day diary, and the PPAQ. Their results ranged from fair to almost perfect. For this reason, evidence for validity was obtained by determining the relationship between the PPAQ-Turkish and the Turkish version of the IPAQ as well as the SF-36 in this study. In the original PPAQ validation study, the observed correlations between the PPAQ and three published cutoff points used to classify actigraph data ranged from 0.08 to 0.43 for total activity, 0.25 to 0.34 for vigorous activity, 0.20 to 0.49 for moderate activity, and −0.08 to 0.22 for high-intensity activity[13]. Ota et al.[16] reported that the Pearson correlation coefficient for measurements between the PPAQ and pedometer was 0.29. The correlation coefficient between the PPAQ-Turkish and the Turkish version of the IPAQ was similar to those found in other studies (ranging from 0.17 to 0.32)[2, 13–16]. The correlation between the SF-36 and scores of specific instruments are typically weak. This confirms that the SF-36 measures additional aspects of physical health and provides more comprehensive, but less specific, information about a patient's overall health compared with condition-specific questionnaires. As expected, the PPAQ was more strongly related to concurrent measures of physical function than to concurrent measures of mental function. However, in this study, the PPAQ-Turkish and the PCS and MCS subscales of the SF-36 showed fair correlations, but these correlations were poor for the third trimester of the pregnancy.

In the assessment of ceiling and floor effects, it was found that none of the patients scored the minimum or maximum score in the PPAQ-Turkish. This suggests that the PPAQ-Turkish is an appropriate tool for pregnant woman. This study has some limitations. First, the validity of the results was affected by the Turkish language when measuring physical activity. Second, although occupational and household/caregiving activities tend to be consistent across seasons, there may be seasonal variation in sports/exercise activities. However, we did not take this into consideration.

In conclusion, based on the results of this study, it can be concluded that there is enough evidence of acceptable reliability and validity to use the PPAQ-Turkish questionnaire for pregnant women in Turkish-speaking societies.

REFERENCES

1) Moore SC, Patel AV, Matthews CE, et al.: Leisure time physical activity of moderate to vigorous intensity and mortality: a large pooled cohort analysis. PLoS Med, 2012, 9: e1001335. [Medline] [CrossRef]
2) Evenson KR, Chasan-Taber L, Symons Downs D, et al.: Review of self-reported physical activity assessments for pregnancy: summary of the evidence for validity and reliability. Paediatr Perinat Epidemiol, 2012, 26: 479–494. [Medline] [CrossRef]
3) Lee JH, Park JH, Kim YJ: Sensitivity of the accelerometer as a measurement tool for upper extremity movement by stroke patients: a comparison with the action research arm test. J Phys Ther Sci, 2015, 27: 1053–1054. [Medline] [CrossRef]
4) Abe Y, Sakamoto M, Nakazawa R, et al.: Relationship between joint motion and acceleration during single-leg standing in healthy male adults. J Phys Ther Sci, 2015, 27: 1251–1256. [Medline] [CrossRef]
5) Makabe S, Makimoto K, Kikkawa T, et al.: Reliability and validity of the Japanese version of the short questionnaire to assess health-enhancing physical activity (SQUASH) scale in older adults. J Phys Ther Sci, 2015, 27: 517–522. [Medline] [CrossRef]
6) Cho YK, Kim DY, Jung SY, et al.: Synergistic effect of a rehabilitation program and treadmill exercise on pain and dysfunction in patients with chronic low back pain. J Phys Ther Sci, 2015, 27: 1187–1190. [Medline] [CrossRef]
7) Lindseth G, Vari P: Measuring physical activity during pregnancy. West J Nurs Res, 2005, 27: 722–734. [Medline] [CrossRef]
8) Haakstad LA, Gundersen I, Bø K: Self-reporting compared to motion monitor in the measurement of physical activity during pregnancy. Acta Obstet Gynecol Scand, 2010, 89: 749–756. [Medline] [CrossRef]
9) Schmidt MD, Freedson PS, Pekow P, et al.: Validation of the Kaiser Physical Activity Survey in pregnant women. Med Sci Sports Exerc, 2006, 38: 42–50. [Medline] [CrossRef]
10) Evenson KR, Wen F: Measuring physical activity among pregnant women using a structured one-week recall questionnaire: evidence for validity and reliability. Int J Behav Nutr Phys Act, 2010, 7: 21. [Medline] [CrossRef]
11) Ainsworth BE, Haskell WL, Herrmann SD, et al.: 2011 Compendium of Physical Activities: a second update of codes and MET values. Med Sci Sports Exerc, 2011, 43: 1575–1581. [Medline] [CrossRef]
12) Harrison CL, Thompson RG, Teede HJ, et al.: Measuring physical activity during pregnancy. Int J Behav Nutr Phys Act, 2011, 8: 19. [Medline] [CrossRef]
13) Chasan-Taber L, Schmidt MD, Roberts DE, et al.: Development and validation of a pregnancy physical activity questionnaire. Med Sci Sports Exerc, 2004, 36: 1750–1760. [Medline] [CrossRef]
14) Chandonnet N, Saey D, Alméras N, et al.: French Pregnancy Physical Activity Questionnaire compared with an accelerometer cut point to classify physical activity among pregnant obese women. PLoS ONE, 2012, 7: e38818. [Medline] [CrossRef]
15) Matsuzaki M, Haruna M, Nakayama K, et al.: Adapting the Pregnancy Physical Activity Questionnaire for Japanese pregnant women. J Obstet Gynecol Neonatal Nurs, 2014, 43: 107–116. [Medline] [CrossRef]
16) Ota E, Haruna M, Yanai H, et al.: Reliability and validity of the Vietnamese version of the Pregnancy Physical Activity Questionnaire (PPAQ). Southeast Asian J Trop Med Public Health, 2008, 39: 562–570. [Medline]
17) Silva FT, Costa FD: Transcultural adaptation of the pregnancy physical activity questionnaire-PPAQ to Portuguese: a tool for evaluation of physical activity in Brazilian pregnant. FIEP Bull, 2009, 79: 318–322.
18) Beaton DE, Bombardier C, Guillemin F, et al.: Guidelines for the process of cross-cultural adaptation of self-report measures. Spine, 2000, 25: 3186–3191. [Medline] [CrossRef]
19) Roberts DE, Fragala MS, Pober D, et al.: Energy cost of physical activities during pregnancy. Med Sci Sports Exerc, 2002, 34: S124. [CrossRef]
20) Kocyigit H, Aydemir O, Fisek G, et al.: Validity and reliability of Turkish version of Short form 36: a study of a patients with romatoid disorder. J Drug Ther, 1999, 12: 102–106.
21) Craig CL, Marshall AL, Sjöström M, et al. IPAQ Consensus Group and the IPAQ Reliability and Validity Study Group: International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc, 2003, 35: 1381–1395. [Medline] [CrossRef]
22) Saglam M, Arikan H, Savci S, et al.: International physical activity questionnaire: reliability and validity of the Turkish version. Percept Mot Skills, 2010, 111: 278–284. [Medline] [CrossRef]
23) Terwee CB, Bot SD, de Boer MR, et al.: Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol, 2007, 60: 34–42. [Medline] [CrossRef]
24) Marx RG, Menezes A, Horovitz L, et al.: A comparison of two time intervals for test-retest reliability of health status instruments. J Clin Epidemiol, 2003, 56: 730–735. [Medline] [CrossRef]
25) Landis JR, Koch GG: The measurement of observer agreement for categorical data. Biometrics, 1977, 33: 159–174. [Medline] [CrossRef]