Psychometric Testing of GPAQ among the Thai Population

นิพนธ์ต้นฉบับ

วนิดา วิสุทธิพานิช*

ภาควิชาการพยาบาลเด็ก คณะพยาบาลศาสตร์ มหาวิทยาลัยศรีนครินทรวิโรฒ อ.องครักษ์
จ.นครนายก 26120

* ติดต่อผู้นิพนธ์: vanida@swu.ac.th

Original Article

Vanida Visuthipanich*

Department of Pediatric Nursing, Faculty of Nursing, Srinakharinwirot University,
Ongkharak, Nakhonnayok, 26120, Thailand

* **Corresponding author:** vanida@swu.ac.th

บทคัดย่อ

วัตถุประสงค์: การศึกษาวิจัยเชิงพรรณนานี้มีวัตถุประสงค์เพื่อทดสอบความตรง
และความเชื่อมั่นแบบทดสอบซ้ำของแบบสอบถามกิจกรรมการเคลื่อนไหวออกแรง
และการออกกำลังกาย GPAQ ฉบับแปลในคนไทย **วิธีการศึกษา:** ทดสอบ
คุณสมบัติทางจิตวิทยาของแบบสอบถามในคนไทยจำนวน 160 คนอายุ 15 – 65
ปี ร่วมกับการประเมินเชิงภววิสัยที่ใช้เครื่องวัดกิจกรรมการเคลื่อนไหวในการ
ทดสอบความตรงและความเชื่อมั่นแบบทดสอบซ้ำของ GPAQ ที่ผ่านการแปล
ย้อนกลับแล้ว **ผลการศึกษา:** พบว่า GPAQ มีค่าความสอดคล้องและค่าดัชนี
ความตรงของเครื่องมือที่แปลเท่ากับ 0.88 และ 0.99 ตามลำดับ พบว่า GPAQ มี
ความตรงตามสภาพระหว่างเมื่อเทียบกับ IPAQ ฉบับยาวในระดับยอมรับได้ (ρ =
0.75) ความตรงตามเกณฑ์ที่ค่อนข้างต่ำ (ρ = 0.33) และความเชื่อมั่น
แบบทดสอบซ้ำที่ยอมรับได้ (ρ = 0.77) **สรุป:** เครื่องมือแบบสอบถามกิจกรรม
การเคลื่อนไหวออกแรงและการออกกำลังกาย GPAQ มีความตรงและความ
เชื่อมั่นแบบทดสอบซ้ำที่ยอมรับได้ แต่ไม่สูงมากนัก ควรพัฒนาต่อไป

คำสำคัญ: แบบสอบถามกิจกรรมการเคลื่อนไหวออกแรงและการออกกำลังกาย,
ความตรง, ความเชื่อมั่น, การแปลกลับ

Abstract

Objective: This descriptive cross-sectional study aimed to assess the
validity and test-retest reliability of the back-translated version of the Global
Physical Activity Questionnaires (GPAQ). **Methods:** GPAQ was completed
by 160 Thai persons aged 15 to 65 years old coupled with objective
measurement of physical activity (PA) using motion sensor to determine the
validity and test-retest reliability of the back-translated questionnaire.
Results: The inter-rater agreement and item-level for content validity index
(I-CVIs) of the GPAQ were 0.88 and 0.99, respectively. Concurrent validity
between the GPAQ and the long form International Physical Activity
Questionnaire (LF IPAQ) was acceptable (ρ = 0.75). Criterion validity was
relatively low (ρ = 0.33). Test-retest reliability was acceptable (ρ = 0.77).
Conclusion: Validity and test-retest reliability of the GPAQ were in an
acceptable level, if not too high. Further improvement is needed.

Keywords: physical activity questionnaire, validity, reliability, back
translation

Introduction

The evidences from previous studies suggested that sufficient engagement in physical activity could result in reducing many of the risk factors for non-communicable diseases including reducing blood pressure, lowering body mass index, and improving blood cholesterol among healthy adults.[1] There has been a recommendation for adults aged 18 – 64 that throughout the week, they should perform at least 150 minutes of moderate-intensity aerobic physical activity or at least 75 minutes of vigorous-intensity aerobic physical activity (PA), or an equivalent combination of moderate- and vigorous-intensity activity.[2] However, there were inconsistent findings during the past decade that most of Thai population (70 to 80%) performed a sufficient level of PA of which an increased incident of chronic diseases and mortality rate could be the result.[3] The inconsistency of findings may be in part due to the validity and reliability of the Thai version of the PA measuring instrument.

A valid and reliable measuring instrument has been needed PA assessment. This instrument is crucial for the development of PA survey research and of a strategic plan to promote health and reduce chronic diseases and mortality rates among the Thai population. In most large-scaled population-based studies, PA surveillance has been commonly undertaken using a self-reported questionnaire due to its simple, inexpensive, and convenient process, and its somewhat acceptable reliability and validity.[4,5] There has been a need for an objective tool, such as a motion sensor, to help improve the validity and reliability of this PA self-reported questionnaire.[6] The most common motion sensor used is the pedometer and accelerometer. However, there has been a use of an accelerometer as a motion sensor since it has been accepted as the gold standard for objective measurement. Accelerometer (Actigraph®) is a uniaxial, small, light, and unobtrusive device.

The Global Physical Activity Questionnaires (GPAQ) has been used as a self-reported questionnaire in Thailand with no systematic psychometric tests in Thai population.[7,8] The GPAQ was developed in 2003 - 2005,[9] of which the GPAQ

version 2 was developed specifically for developing countries. Using pedometer and accelerometer as the standard, a poor to fair criterion validity of GPAQ was reported (r = 0.06 to 0.35 for all domains, P = 0.01). Concurrent validity was moderate to strong based on the inter-method comparisons with the short-form International Physical Activity Questionnaire (IPAQ) (r = 0.45 to 0.65, P = 0.01). The findings of good to very good 3- to 7-day test-retest reliability of the GPAQ were reported (kappa = 0.67 to 0.73; Spearman rho = 0.67 to 0.81).[9] The GPAQ version 2 consists of 16 questions with three domains including 1) activity at work (paid and unpaid), 2) travel to and from place (walking or cycling), and 3) discretional activities (leisure, recreation, etc.), and the time spent in sedentary activities.[9] The IPAQ has been known to be inferior to the GPAQ as Bull et al addressed limitations of both forms of IPAQ that the lack of domain-specific estimates both from the short-form (SF) and long-form (LF) versions of the IPAQ could potentially prohibit their use within the WHO's STEPwise approach to NCD risk factor surveillance or STEPS.[9] We therefore postulated that GPAQ could be applicable in more settings and programs. In our present study, we proposed to test the psychometric properties of the GPAQ with a systematic back translation for the first time.

The long form International Physical Activity Questionnaire (LF IPAQ) was initially developed in 1998 - 1999 and tested for psychometric property in 12 countries with convenience samples that represented generalized age, education, income, and activity levels.[11] The IPAQ has acceptable measurement properties for PA monitoring among population aged 18 - 65 years old across countries. Based on the objective measures by accelerometers, the criterion validity of the LF IPAQ was low with a correlation coefficient of only 0.33 (95% CI: 0.26 - 0.39) Spearman's correlation coefficient of 0.67 (95% CI: 0.64 - 0.70) between IPAQ both forms (inter-method) indicated a reasonable agreement. The test-retest reliability was found to be good with a Spearman's correlation coefficient of 0.81 (95% CI: 0.79 - 0.82).[11] The LF IPAQ comprises 27 items covering the PA domains of occupation, self-powered transport, household/yard/garden work, recreation, walking, cycling, and sedentary activity. The SF IPAQ consists of 7 items covering all activities of vigorous- and moderate-intensity activity, walking (vigorous, moderate, and slow pace), and sedentary activity. In Thailand, a Thai version of LF IPAQ has been used but it was not translated by a systematic back translation process. In our present study, we back translated the IPAQ and used it to test concurrent validity of the GPAQ.

The objective of this study was to conduct a set of psychometric tests on the GPAQ among Thai population aged 15 to 65 years old. The tests included concurrent validity, criterion validity, and reliability using accelerometer which was a motion sensor as the gold standard for objective measurement.

Methods

A descriptive cross-sectional research design was used to investigate psychometric properties of the GPAQ which was developed and tested in western countries. This study consisted of two phases. The first phase included the translation and back-translation process of both the GPAQ and LF IPAQ questionnaires into Thai language. The translation of the LF IPAQ was also carried out since its Thai version with a systematic back-translation had never been done before. The second phase was psychometric testing of the translated questionnaire. The inter-method testing between the GPAQ and the LF IPAQ was conducted to test the concurrent validity of the GPAQ. The subjective measures provided by the GPAQ and LF IPAQ were tested for their accuracy by comparing with a more objective physiologic measure obtained from the accelerometer. The seven-day test-retest reliability was also conducted. An overview diagram of the GPAQ testing is depicted in Figure 1 and addressed as follows:

Phase I: Translation and back-translation of the GPAQ and IPAQ questionnaire to Thai language

Translation procedures

The Thai translation of the GPAQ and LF IPAQ was carried out by an expert in the Thai culture who did not participate in any part of the research. The questionnaire was back translated into English by a second Thai translator who was proficient in English language and had lived in a western country for 5 years. To maintain the original meaning including the cultural equivalency of the instruments,[10] the later drafts of the Thai translation of the GPAQ and LF IPAQ were reviewed and edited by the study investigator.
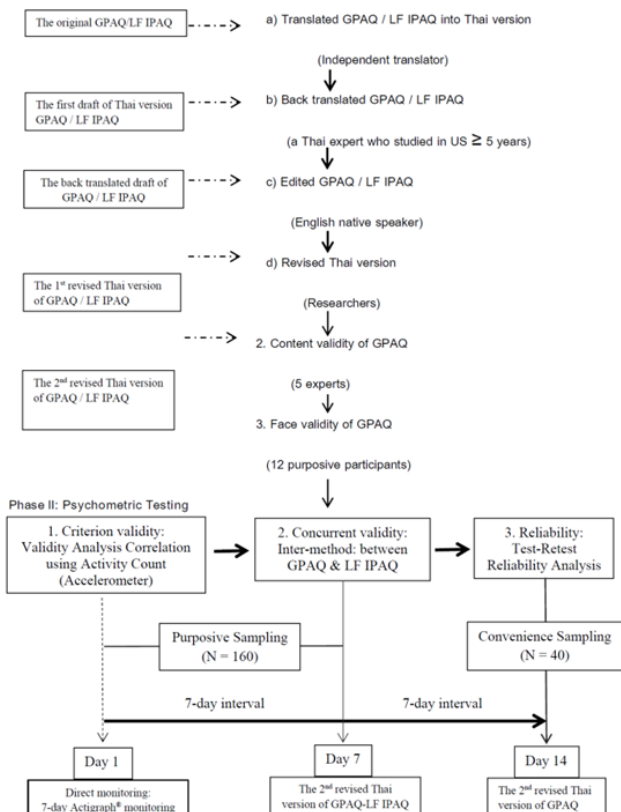
Figure 1 Overview diagram of GPAQ psychometric testing.

GPAQ = Global Physical Activity Questionnaire;
LF IPAQ = Long form Physical Activity Question

Tests for content validity and face validity

Instructions of how to complete the survey were mailed to five professors based on their expertise with the topic of this study who had been working as faculty members at the other comparable universities. For content validity, the experts were asked to identify the content relevance of each item and evaluate the clarity of the item's meaning in the Thai cultural context. Face validity was conducted on a purposive sample of 12 participants, three for each of the four age groups, specifically 15 to 24 years old (young adult), 25 to 44 (adult), 45 to 59 (mid-adult) and 60 to 65 years old (late adult) . All participants were asked to identify the content relevance of each item and evaluate the clarity of each item's meaning under the Thai culture.

Phase II: Psychometric Testing

To evaluate the psychometric properties of the Thai version of the GPAQ including concurrent validity, criterion validity and test-retest reliability, we recruited 200 eligible participants who met the inclusion criteria. They were directly monitored using the accelerometer (Actigraph®) and completed the GPAQ and the LF IPAQ surveys. Each participant was asked to wear the accelerometer (Actigraph®) attached to a Velcro strap belt around the hip at the right antero-lateral side near iliac crest all day for seven-consecutive days (five weekdays and two weekend days). The Actigraph® was removed during sleeping, bathing, and swimming times. As a gold standard for objective measure of PA, data from the accelerometer were used to test the **criterion validity** of the GPAQ. Each participant also completed the long-form IPAQ survey for **concurrent validity** testing of the GPAQ. Forty convenience participants, (10 participants per age group) from 200 participants were further asked to complete the second GPAQ survey for a **test-retest reliability** testing.

Setting, Sample, and Participants Selection

The study population consisted of Thai individuals aged 15 to 65 years old who resided in the metropolitan area of Bangkok. The sample sizes calculated and sampling method were based on the purpose of study. Firstly, 12 purposive participants were recruited to determine face validity of the questionnaire. The sample size in the phase of psychometric testing was estimated as follows. As each item of the questionnaire needed at least 10 participants to achieve an adequate statistical power, 160 participants were needed for the 16 items of the GPAQ. To compensate for a 25% attrition rate, 40 more participants were needed and the final sample size of 200 participants was indicated.

For a representative purpose of the sample, all participants were recruited from three zones of Bangkok using the multistage random sampling. The sample was divided into four groups of participants according to their age as previously described. All participants met the inclusion criteria including being ambulatory, aged 15 to 65 years old, willing to participate, and able to communicate in Thai language. In addition, ten participants from each age group were selected from a convenience sample of 200 participants to determine reliability as previously described.

Ethical considerations

This study was approved by the Ethical Review Committee for Research in Human Subjects, Ministry of Public Health, Thailand. All participants completed the written informed consent form prior to completing the questionnaire.

Data Collection Tool

Data collection instruments included personal data sheet, the Thai version of GPAQ and LF IPAQ, and self-monitoring accelerometer (Actigraph®). The personal data sheet consisted of demographic data questionnaire (age, gender, education level, marital status, and health status) and the PA Readiness Questionnaire (PAR-Q and You) which was a set of seven self-administered questions to use as a screening device to determine medical conditions and cardiac risk factors before PA conduction.[12]

The Thai version of GPAQ consisted of three dimensions. The first dimension namely activities at work had six items covering vigorous and moderate intensity PA. The second dimension which was travel to and from place had three items covering walking and bicycling. The last dimension, i.e., recreational activity, had six items covering vigorous and moderate intensity PA. The GPAQ also contained an item of sedentary behavior. The participants were asked to complete the actual questions if they conducted the physical activities. If they did not, they were asked to skip the questions. For the IPAQ, the participants were asked to complete in the same manner. The final actual measures of PA, subjective and objective, were the total energy expenditure. Various energy expenditure values were assigned on each of the intensity of the PA in each activity in the GPAQ and IPAQ to calculate the energy expenditure. The calculation method of the total energy expenditure for the data from the accelerator, GPAQ and IPAQ are described in the data analysis section.

Data collection procedure

Thais with age between 15 and 65 years old were recruited from the list of names and contact information from each recruiting site provided by the center administrator. Of the 245 Thais who met the inclusion criteria, 200 individuals (81.63%) agreed to participate in the study. Forty-five individuals (18.37%) declined to participate in this study due to inconvenience to conduct the direct monitoring. Potential participants were informed of the study via letters and direct contact. Eligible participants were scheduled for an appointment to complete direct monitoring and PA surveys. All participants were asked to complete the PA surveys after completing direct monitoring. The test-retest reliability was conducted in the same manner a week after the first survey to determine reliability of the instrument. The sequence of the presence of GPAQ and LF IPAQ questionnaires was also of concern because the responders could have remembered the answers from the first survey. Therefore two sequences of the questionnaires were provided. Specifically, for those who completed the GPAQ first at the first survey, they completed the LF IPAQ first at the second survey, and vice versa. Participants took about 30 to 45 minutes for each PA questionnaire to complete. Therefore, at each survey, participants were allowed to rest for about 30 minutes with refreshments served between the GPAQ and LF IPAQ administration because they could have been too exhausted to complete the survey accurately.

Data analysis

The responses from the experts related to the content relevance of each item were used to determine the inter-rater agreement index and content validity index of individual items (I-CVIs).[13,14] Inappropriate and ambiguous items were reviewed and refined according to the suggestions from the experts. The revision of the items based on the feedback from the 12 participants (face validity sample) was conducted. Also, a simple descriptive statistical analysis was used to analyze demographic data and PA level including frequency with percentage and mean with standard deviation. Analysis of difference between the groups such as age groups and intensity of the activities, was employed to determine the differences in PA level.

The measures of IPAQ, GPAQ and the accelerator were the total energy expenditure. The exact number of the frequency and duration of each PA activity was recorded. Scoring and data entry were conducted. Incomplete information was excluded. The total energy expenditure (EE) per week was ultimately presented as the unit of total metabolic equivalent task (MET)*minutes per week.

The energy expenditure values from the GPAQ and the LF IPAQ were computed in the same manner as originally designed by the equation of frequency × duration × intensity, and summed across all related activities for each individual.

Each specified activity was adjusted to the MET value depending on the intensity of the activity on the PA listed in the 2011 Compendium of PA.[15] In other words, the intensity in the equation above was filled with the MET value for each intensity of a given activity. Data from the GPAQ and LF IPAQ were summed as guided by the creators of the two scales into each intended PA domain. MET*min/week values of GPAQ and IPAQ were further categorized into two levels as moderate- and vigorous-activities according to the summation of EE in all activities of each PA questionnaire. Moreover, moderate- and vigorous- MET value of the GPAQ (a mean MET of 4 and 8, respectively) and IPAQ (3 - 6 and > 6, respectively) for PA calculation were used.

Since we found that the data of MET, from measures, were not normally distributed, nonparametric Spearman's rho coefficients were calculated to determine the correlation of the total METS-min/wk) of moderate and vigorous PA measured by the direct monitoring and the subjective measures by the GPAQ and LF IPAQ. All data sets were used to analyze criterion validity (PA data and direct monitoring data), concurrent validity (inter-method comparing between the GPAQ and LF IPAQ), and reliability (seven-day test-retest of the GPAQ data)

Results

Phase I: Translation and back-translation of Questionnaire to Thai language and culture

Content Validity

It was found that the inter-rater agreement index and I-CVIs were 0.88 and 0.99, respectively for the GPAQ. There were minor changes of wording in the GPAQ equivalency and face validity, such as "per week" changed to "in a typical week" and "when you go to another place" changed to "continuously to get to and from places."

Phase II: Psychometric testing

Descriptive demographic data

Among 200 eligible participants, 40 participants (25% attrition rate) were excluded from the study due to incomplete seven-consecutive days direct monitoring. Demographic data of the 160 participants were analyzed. Twice as many females (63.75%) as males participated in the study. Half of the participants were single (53.8%). Most of them graduated with a Bachelor's degree. Participants presented similar proportions of occupation including those staying at home (31.88%), students (33.13%), traders (40.00%), and business employees (36.25%). Over half of the participants (66%) had normal body mass index, one-third were overweight (30%), and a small percentage were obese (4%) according to the Department of Health's standard (18 kg/m^2 ≤ BMI < 23 kg/m^2).[16] Although all obese participants were female, two-thirds of male participants presented waist circumferences higher than standard (90 cm. for male and 80 cm. for female).[17] One-fourth of the participants (28.13%) experienced loss of balance due to dizziness and had chronic health problem. The common health problems reported by participants were hyperlipidemia (24%), hypertension (22%), bone and joint problems (17%), diabetes mellitus (12%), and cardiovascular disease (11%).

Measurements of PA

Both the GPAQ and LF IPAQ presented an overestimated PA level when compared with that of the Actigraph® (Table 1). For the total METS-min/week, the GPAQ and the IPAQ resulted a 126.10% and 153.59% of the overall METS-min/week when compared with that of the Actigraph®. An opposite direction was found in the vigorous intensity activity where the GPAQ overestimated (136.31%), the IPAQ on the other hand underestimated the EE (95.09%) when compared with the Actigraph®. For the moderate intensity activity, both the GPAQ and IPAQ overestimated the EE about three times of that of the Actigraph®.

Table 1 Mean value of EE as METS-min/week from the GPAQ, LF IPAQ and Actigraph® measures (N = 200).

Measurement	GPAQ	LF IPAQ	Actigraph®
Overall total METS-min/week	3,197.85	3,894.93	2,535.97
	(126.10%)	(153.59%)	(100%)
Total METS-min/week of	1,171.75	817.41	859.62
Vigorous intensity activity	(136.31%)	(95.09%)	(100%)
Total METSs-min/week of	1,339.51	1,297.62	399.71
Moderate intensity activity	(335.12%)	(324.64%)	(100%)

Note: GPAQ = Global Physical Activity Questionnaire; MET-min/wk = Metabolic Equivalents Task

In terms of correlations between EE in each domain of the GPAQ and the LF IPAQ with the EE of the vigorous- and moderate-intensity PA activities, the results are shown in Table 2. The highest and significant correlation values (Spearman's rho; ρ) of the GPAQ in activity at work were 0.81, 0.81, 0.64, 0.83, and 0.73 among participants with hyperlipidemia, hypertension, bone & joint problems, diabetes mellitus, and cardiovascular disease, respectively.

Table 2 Correlation (Spearman's rho) of GPAQ with its own PA dimensions in participants with various chronic illnesses.

Measures	GPAQ	
	Activity at Work	Recreation
Hyperlipidemia (N=24 / 15%)		
Total METS-min/week	0.55**	0.79**
Total METS-min/week of Vigorous activity	0.81**	0.33
Total METS-min/week of Moderate activity	0.23	0.72**
Hypertension (HT) (N=22 / 13.75%)		
Total METS-min/week	0.63**	0.79**
Total METS-min/week of Vigorous activity	0.81**	0.46*
Total METS-min/week of Moderate activity	0.36	0.69**
Bone & Joint Problems (N=17 / 10.63%)		
Total METS-min/week	0.53*	0.52*
Total METS-min/week of Vigorous activity	0.64*	0.52*
Total METS-min/week of Moderate activity	0.50*	0.59*
Diabetes Mellitus (DM) (N=12 / 7.5%)		
Total METS-min/week	0.97**	0.66*
Total METS-min/week of Vigorous activity	0.55	0.79**
Total METS-min/week of Moderate activity	0.83**	0.72**
Cardiovascular Disease (CAD) (N=11 / 6.88%)		
Total METS-min/week	0.92**	0.32
Total METS-min/week of Vigorous activity	0.24	0.52

* P = 0.05 ** P = 0.01
GPAQ = Global Physical activity questionnaire
PA = Physical activity

Psychometric Testing

The results from the psychometric testing of the GPAQ indicated an acceptable criterion validity with ρ of 0.33 (P < 0.01) (Table 3). Concurrent validity when GPAW was compared with LF IPAQ (inter-method) was acceptable with ρ of 0.75 (P = 0.01). The GPAQ also had a good-to-very good seven-day test-retest reliability with ρ of 0.77 (P < 0.01).

Table 3 Spearman's coefficients (ρ) of criterion and concurrent validity, and seven-day test-retest reliability of the GPAQ.

Psychometric evaluation	The present study	Hoos et al, 2012	Au et al, 2010	Bull et al, 2009
Criterion validity	**0.33****	NS	0.39**	0.06** -
(Total METS-min/wk of GPAQ		0.003		0.35**
vs that of direct monitor)		(VLPA 0.4**)		
Concurrent validity	**0.75****	-	0.32**	0.45** -
(Total METS-min/wk of GPAQ				0.65**
vs. that of LF IPAQ)				
Seven-day test-retest reliability	**0.76****		0.39**	0.67** -
(Total METS-min/wk GPAQ vs.				0.81**
GPAQ)				

** P < 0.01   *P < 0.05   NS = Not significance
GPAQ = Global Physical activity questionnaire,
VLPA = Vigorous leisure-time physical activity
LF IPAQ = Long Form International Physical Activity Questionnaire

When levels of PA intensity level was considered, seven-day test-retest reliability of the GPAQ were in acceptable to very good levels with Spearman's rho of 0.77, 0.81, and 0.70 for the total METS-min/week, the total METS-min/week of vigorous intensity activity, and the total METS-min/week of moderate intensity activity, respectively (Table 4). When each domain of the GPAQ were considered, seven-day test-retest reliability of the GPAQ with vigorous intensity for activity at work and exercise with ρ of 0.96 and 0.76 respectively (P < 0.01) (Table 5). Finally, once age groups were considered, the criterion validity of the GPAQ in each age group was not significant (Table 6). On the other hand, concurrent validity and test-retest reliability in all age groups were statistically significant.

Table 4 Seven-day test-retest Spearman's reliability coefficients (ρ) of the GPAQ by the PA intensity level (N = 40).

Measurement	GPAQ
Total METS-min/week	0.77**
Total METS-min/week of Vigorous intensity activity	0.81**
Total METS-min/week of Moderate intensity activity	0.70**

** P < 0.01
GPAQ = Global Physical activity questionnaire,
PA = Physical activity

Table 5 Seven-day test-retest reliability of GPAQ and LF IPAQ by their domains (Spearman's rho) (N = 40).

	Activity at Work		Exercise	
	Vigorous intensity	Moderate intensity	Vigorous intensity	Moderate intensity
GPAQ	0.96*	0.85*	0.76*	0.71*

* P < 0.01

Table 6 Psychometric testing of GPAQ and LF IPAQ among age groups (Spearman's rho).

Tests	Spearman's rho by age group (yrs)			
	15 – 24	25 – 44	45 – 59	60 – 65
Criterion validity	0.12	0.29	-0.05	0.03
(compared with direct monitor)				
Concurrent validity	0.54**	0.74**	0.75**	0.74**
(compared with LF IPAQ)				
Test-retest reliability	0.75**	0.78**	0.82**	0.61**

** P = 0.01

Discussions and Conclusion

In this study, we found that self-administered GPAQ resulted in an overestimated of the energy expenditure as measured by the direct monitoring by Actigraph®. It was

overt especially for the moderate intensity activity of which the estimation was about three time higher than the direct monitoring. This finding was consistent with various studies.[7,8,17,18] This overestimation could be attributable to cultural bias and limited ability to classify participants into low and high levels of PA.

In terms of content validity, the results of the inter-rater agreement and I-CVIs of the GPAQ from this study were over 80%. The resultant inter-rater agreement and I-CVIs score from this study was acceptable according to Davis (1992) that it should seek an 80% agreement (0.80) for new instrument and 70% agreement (0.70) for used instrument.[13]

Regarding criterion validity of the GPAQ, although found to be a rather low correlation (0.33), it was comparable to a few studies in other countries. For example, the Vietnamese version GPAQ indicated acceptable correlation (r = 0.39, P < 0.01), [19] while the criterion validity of the GPAQ among population in nine countries were poor to fair with correlation coefficients ranging from 0.06 to 0.35 (P < 0.01).[9] Similar findings were found among Latino population in vigorous leisure-time activity (r = 0.40, P < 0.01) although there was no statistical significance between the total MET-min/week against accelerometer.[20]

Several studies that measured the validity of both the GPAQ and the LF IPAQ found little to large magnitude of the correlations with the direct monitoring. In patients with schizophrenia, subjective PA instrument had small correlations with accelerometer (r = 0.37, P < 0.001).[21] The criterion validity coefficient of the GPAQ in our study was similar to the one from previous PA survey studies in Asian countries. For example, several PA psychometric testing studies in Asian countries found little to large correlation coefficients with P value less than 0.01 when GPAQ was compared with other questionnaires, for example Shanghai PA questionnaire (ρ = 0.30),[22] Thai PA questionnaire (r = 0.20 to 0.32),[23] self-reported PA questionnaire (r = 0.31),[17] and the Modified Thai Adolescent's PA questionnaire (r = 0.59).[24]

In terms of the concurrent validity of the GPAQ, our study indicated a high agreement between GPAQ and LF IPAQ with a coefficient ρ of 0.75. This correlation coefficient was higher than those in previous studies. For example, the findings of the inter-method comparison between GPAQ and IPAQ showed moderate to strong correlation among population in the nine countries (r = 0.45 to 0.65, P < 0.01)

and small magnitude of the GPAQ Vietnamese version (0.32, P < 0.01).[9,19]

For test-retest reliability, it is appropriate to determining the reliability of a measure when the concept being tested is thought to be relatively stable over time. Although, the time interval of the recall period varies across studies, the standard assessment periods used commonly by the researchers.[25] This study selected a seven-day period for time interval to repeat the test since the time interval for assessing specific time recall physical activity should match the instrument's time frame. The results from this study showed that GPAQ indicated satisfactory reliability when compared to the original studies where they found r of 0.67 to 0.81.[9] Other studies also found high reliability coefficients, such as Shanghai PA questionnaire (r = 0.65),[22] Thai PA questionnaire (ICC of r = 0.85),[23] the modified Thai adolescents PA questionnaire (r = 0.76, ρ = 0.01),[24] and self-reported PA questionnaire (0.93; P = 0.01).[17]

Our study had some limitations. First, participants' recall of PA level could result in a certain level of bias. Second, there was a limitation of objective device to measure sedentary activity. Third, the participants' concerns regarding the fear of equipment damage or loss, itching from the belt, forgetting to wear, and a decision not to wear the Actigraph® when doing a household chore or gardening. Such concerns limited Actigraph® wearing and could lead to inaccurate results.

In the future, we recommend that randomized controlled trial on the effect of PA and physical fitness-related health among Thai population using the GPAQ should be conducted. Second, exploration of diverse participants throughout the country for a better generalization across geographical areas both rural and urban should be conducted. Finally, studies to explore other criterion related validity, such as predictive validity (physical fitness) should be conducted.

In conclusion, the study of the psychometric properties of the GPAQ among Thai population aged 15 to 65 years old found a modest criterion validity, acceptable to very good concurrent validity, and very good seven-day test-retest reliability. Although, the GPAQ had a fair criterion validity, this survey study suggested that it was relatively practical for PA surveillance in Thai population. The GPAQ was found to be the most appropriate PA questionnaire, especially in the

dimension of garden and household work among Thai population aged 45 - 59 years old.

Acknowledgement

Authors' Note

This study was presented at the 3rd NUS-NUH International Nursing Conference & 20th Joint Singapore – Malaysia Nursing Conference, Singapore, November 18-20, 2015.

Consideration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the author and/or publication of this article.

References

1. Christine Hancock. Collaborating for health review: the benefits of physical activity for health and well-being 2011. (Accessed on May 14, 2016, at http://www.c3health.org/wp-content/uploads/2009/09/C3-review-of-physical-activity-and-health-v-1-20110603.pdf&rct=j&frm=1&q=&esrc= s&sa=U&ved=0ahUKEwi8_qWWvrnOAhUDK48KHVlVD-MQFggpMAM&usg=AFQjCNGskP9zhO5QNVjcxi6oZVnjQCTiFw)

2. WHO. Global strategy on diet, physical activity and health. (Accessed on Jan. 15, 2016, at http://www.who.int/dietphysicalactivity/factsheet_adults/en/)

3. Wibulpolprasert S. Thailand Health Profile 2008 – 2010. Nonthaburi. Veterans Organization of Thailand Press, 2012.

4. Ainsworth BE, Haskell WL, Herrmann SD, et al. Compendium of physical activities: a second update of codes and MET values. *Med Sci Sports Exerc* 2011;43(8):1575-1581.

5. Hagstromer M, Ainsworth BE, Oja P, Sjostrom M. Comparison of a subjective and an objective measure of physical activity in a population sample. *J Phys Act Health* 2010;7(4):541-550.

6. Hills AP, Mokhtar N, Byrne NM. Assessment of physical activity and energy expenditure an overview of objective measure. *Frontiers in Nutrition* 2014;1(5):1-16. (doi:10.3389/fnut2014.00005)

7. Aekplakorn W, Porapukam Y, Thanipanichasakul S, Pakcharoen H, Satiannopakao W, Thaikla K. Health status surveys among Thais using health assessment: 4th version, 2008 – 2009. Bangkok. The Graphic System Limited, 2010.

8. Porapukam Y, Boonyarattapun P. Health status surveys among Thais: 3rd Health assessment, 2003 – 2004. Bangkok. Express Transportation Organization Printing, 2006.

9. Bull FC, Maslin TS, Armstrong T. Global physical activity questionnaire (GPAQ): nine country reliability and validity study. J Phys Act Health 2009 Nov;6(6):790-804.

10. WHO. Process of translation and adaptation of instruments. 2015. (Accessed on Jan. 25, 2015, at http://www.who.int/substance_abuse/research_tools/translation/en/)

11. Craig CL, Marshall AL, Sjostrom M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc* 2003;35(8): 1381-1395.

12. Canadian Society for Exercise Physiology. PAR-Q and YOU: A questionnaire for people aged 15 and up. 2014. (Accessed on Jan. 15, 2015, at http://www.google.co.th/url?url=http://lmsc.ca/styled/files/2012-13-par-q002c-registr.pdf&rct=j&frm=1&q=&esrc=s&sa=U&ved=0ahUKEwjfqdq1jrvOAhWBu48KHVAGBpgQFggjMAM&usg=AFQjCNGt6aV7wv-Xu5Eta6Jw_69c_lpEnw)

13. Davis LL. Instruments review: Getting the most from a panel of experts. *Appl Nurs Res* 1992;5:194-197. (doi.10.1016/S0897-1897(05)80008-4)

14. Polit DF, Beck CT. The content validity index: are you sure you know what's being reported? Critique and recommendations. *Res Nurs Health* 2006;29(5):489-497.

15. Department of Health, Ministry of Public Health, Nonthaburi, Thailand. Normal standard of BMI among Thai population 2015. (Accessed on Jan. 5, 2016, at http://nutrition.anamai.moph.go.th/web/bmi.html)

16. Visuthipanich V, Sirapo-ngam Y, Malathum P, Kijboonchoo K, Vorapongsathorn T, Winters-Stone K. Physical activity questionnaire development and testing among elderly community-dwelling Thais. *Thai J Nurs Res* 2009;13(4):249-267.

17. Hagstromer M, Ainsworth BE, Oja P, Sjostrom M. Comparison of a subjective and an objective measure of physical activity in a population sample. *J Phys Act Health* 2010;7(4):541-550.

18. Au TB, Blizzard L, Schmidt M, Pham LH, Magnussen C, Dwyer T. Reliability and validity of the global physical activity questionnaire in Vietnam. *J Phys Act Health* 2010;7(3):410-418.

19. Hoos T, Espinoza N, Marshall S, Arredondo EM. Validity of the Global Physical Activity Questionnaire (GPAQ) in adult Latinas. *J Phys Act Health* 2012 Jul;9(5):698-705.

20. Faulkner G, Cohn T, Remington G. Validation of a physical activity assessment tool for individuals with schizophrenia. *Schizophr Res* 2006;82(2-3):225-231.

21. Peters TM, Shu XO, Moore SC, et al. Validity of a physical activity questionnaire in Shanghai. *Med Sci Sports Exerc* 2010;42(12):2222-2230.

22. Pongurgsorn, C. A questionnaire for assessment of physical activity in Thailand. Doctoral degree dissertation, University of Illinois at Urbana-Champaign, 2001.

23. Wattanasit P, Prateepchaikul L, Petpichetchian W, Meininger JC, Kijboonchoo K. Validity and reliability of the modified Thai adolescent's physical activity questionnaire. *Pac Rim Int J Nurs Res* 2010;14(1):79-92.

24. Melnyk, Morrison-Beedy. Intervention research designs, conducting, analyzing, and funding. New York. Sprinker Publishing Company, 2012.