

The validation and cross-cultural adaptation of the Arabic version of the pregnancy physical activity questionnaire

Afnan S. Younis, MPH, SBCM, Nada A. Alyousefi, MBBS, SB-FM, Dina M. Al-Habib, MBBS, SBCM, Amal T. Al-Omran, MBBS.

ABSTRACT

الأهداف: ترجمة استبيان النشاط البدني أثناء الحمل (PPAQ) إلى اللغة العربية، وملائمتها ثقافياً واختبار موثوقيتها وصلاحتها بين النساء الحوامل السعوديات.

المنهجية: تمت ترجمة PPAQ، الذي يتألف من 36 عنصراً، إلى اللغة العربية وفقاً لإرشادات منظمة الصحة العالمية الخاصة بترجمة المقاييس (الترجمة الأولية، ولجنة الخبراء والترجمة العكسية والاختبار المسبق والمقابلات المعرفية والنسخة النهائية)، متبوعاً بالتحقق من صحة الترجمة من قبل الخبراء. وكان من بين المشاركين 118 امرأة سعودية حامل سليمة صحياً. تم تقييم صلاحية الاستبيان باستخدام مؤشرات صلاحية المحتوى ونموذج راش، بينما تم تقييم موثوقية الاستبيان بحساب الاتساق الداخلي لمحتوى الاستبيان ومعامل ألفا كرونباخ.

النتائج: كان متوسط عمر المشاركات 30.15±5.59 عام، وكان مؤشر كتلة الجسم في المستوى الطبيعي لـ 38.15% منهم. كان متوسط إجمالي إنفاق الطاقة في الأنشطة البدنية 356.1 مكافئاً أيضاً للعمل لساعة / أسبوع (المدى الربيعي = 162.3-648.3). كانت مؤشرات صلاحية المحتوى للاستبيان تتراوح بين 0.8 – 1.0 بينما كانت موثوقية الاستبيان ممتازة تبعاً لنموذج راش (>0.89).

الخلاصة: النسخة العربية من PPAQ هي أداة موثوقة وصالحة ويمكن استخدامها في الدول العربية.

Objectives: To translate the pregnancy physical activity questionnaire (PPAQ) into Arabic language, cross-culturally adapt and test its reliability and validity among Saudi pregnant women.

Methods: Pregnancy physical activity questionnaire, which consisted of 36 items, was translated to Arabic following the World Health Organization's guidelines for tool translation (forward translation, expert panel and back translation, pretesting and cognitive interviewing, and final version), followed by validation by experts. This is a cross-sectional study and data were collected from 118 healthy pregnant Saudi women from May to June 2019. Validity included content validity indices (CVI) and construct validity by Rasch analysis. Reliability was assessed by test-retest reliability and Cronbach's alpha coefficient.

Results: The mean age of the participants was 30.15±5.59 years; 38.2% of them had normal pre-gestational body mass index (n=45). The median of total energy expenditure in physical activity was 356.1 METs.h/week (IQR=162.3-648.3). Item content validity index was good ranging between 0.8-1. Rasch analysis showed good construct validity and excellent reliability for all types of physical activity (>0.89).

Conclusion: This Arabic PPAQ is a reliable and valid tool that can be used in Arab countries.

Keywords: Saudi women, Arabic, pregnancy physical activity questionnaire, validity, reliability, Rasch analysis

*Saudi Med J 2021; Vol. 42 (5): 499-508
doi: 10.15537/smj.2021.42.5.20200826*

From the Department of Family and Community Medicine, College of Medicine (Younis, Alyousefi, Al-Omran), King Saud University and from the Deputyship for Public Health, Ministry of Health (Al-Habib), Riyadh, Kingdom of Saudi Arabia.

Received 29th December 2020. Accepted 22nd March 2021.

*Address correspondence and reprint request to: Dr. Afnan S. Younis, Department of Family and Community Medicine, College of Medicine, King Saud University, Riyadh, Kingdom of Saudi Arabia.
E-mail: afnan.younis@gmail.com
ORCID ID: <http://orcid.org/0000-0001-6219-5809>*

Physical activity, defined as any bodily movement produced by skeletal muscles and resulting in energy expenditure,¹ is one of the most important factors in preventing and controlling non-communicable diseases such as diabetes, hypertension and coronary heart disease.² It is important for the health of human body in all population subgroups including adults, children, elderly as well as pregnant women.

Physical activity during pregnancy has been shown to be not only safe,³ but also beneficial for the health of mother and baby. For instance, it has been associated with better fetal birth weight,³ decreased

risk of gestational diabetes^{4,6} and preeclampsia.^{7,8} It is recommended that pregnant women spend at least 30 minutes of moderate-intensity physical activity a day for almost every day in a week in the absence of any medical or obstetric contraindication to physical activity.^{9,10}

Despite the benefits of physical activity during pregnancy, the level of physical activity among pregnant women is generally low at a global level. A review by Povidine et al concluded that 60% of pregnant women were physically inactive.¹¹ Another review by Gatson et al that included 25 peer-reviewed papers showed similar results of less than 30% of pregnant women were sufficiently active.¹²

The recent Saudi Arabia Ministry of Health initiatives¹³ acknowledged these challenges and stressed the importance of a healthy lifestyle, including PA. A report published in 2018 recommended that more intensive efforts toward promoting PA and reducing sedentary behaviors among the Saudi population are needed.¹⁴ Moreover, recent recommendations¹⁵ have highlighted the issues of healthy pregnancies including PA for pregnant women to help adapt to physical and mental changes and be prepared for delivery and to contribute to preventing back pain and constipation.¹⁶

As direct measurement of physical activity is difficult and challenging, It is crucial to develop and improve tools that are easy to use yet accurate in assessing physical activity for pregnant women. Tools should have the ability to assess different types and intensities of physical activity. Many tools have been developed to assess physical activity among adults,¹⁵ but Pregnancy Physical Activity Questionnaire (PPAQ) was developed specifically for pregnant women and included various activities that can be performed by pregnant women.¹⁷

Pregnancy Physical Activity Questionnaire has been translated to different languages including Japanese,¹⁸ Chinese,¹⁹ French,²⁰ Polish,²¹ and Turkish,²² and have been widely used in epidemiologic studies evaluating physical activity for pregnant women. A systematic review conducted in 2018 to assess PPAQ showed sufficient reliability for total and vigorous physical activity.²³

The aim of this study is to translate the English PPAQ to Arabic, culturally adapt it and to assess its validity and reliability among pregnant Saudi women.

Methods. This was a cross-sectional study that included 118 pregnant Saudi women who were followed up at King Khalid University Hospital (KKUH), Riyadh, Saudi Arabia between May 13, 2019 and June 10, 2019. The exclusion criteria were multiple pregnancies, medical or obstetric conditions that may have prevented PA such as vaginal bleeding or paralysis, and receipt of advice for bedrest for any reason. This study was approved by the Institutional Review Board at King Saud University, Riyadh, Saudi Arabia (Research Project No.: E-15-1555).

All participants provided informed consent before their enrolment in the study.

Data collection tool. The PPAQ is a self-administered questionnaire that was originally developed in English by Chasan-Taber et al.¹⁷ The questionnaire contains 36 questions. The first 3 questions ask about the date of completion of the questionnaire, date of last menstrual period, and expected date of delivery, followed by 33 questions covering different types of PAs performed during household or caregiving activities, occupational activities, sports or exercise, and transportation, as well as including inactivity assessment. Two additional questions are open for participants to add other types of PA that are not included in the questionnaire. The participants were asked to select the approximate number of hours spent per day/week (depending on the question) in each activity during the last 3 months. Durations ranged between 0 and ≥ 6 hours per day and 0 and ≥ 3 hours per week.¹⁷ For each question, the number of hours was multiplied by the reported intensity of that activity to obtain the average weekly energy expenditure in metabolic equivalent hours per week units (MET.h/wk). The activities included in the questionnaire were further classified according to their intensity in metabolic equivalents of task (METs) as sedentary (<1.5 METs), light (1.5-3.0 METs), moderate (3.0-6.0 METs), or vigorous (>6.0 METs), based on calculations as reported by Chasan-Taber et al.¹⁷

Translation. Questionnaire translation followed the World Health Organization's guidelines for tool translation, which consisted of 4 steps: forward translation, expert panel and back translation, pretesting and cognitive interviewing, and the final version.²⁴

1) Forward translation. The questionnaire was translated by a health professional whose mother tongue was Arabic and was fluent in English. Translation was conceptual rather than literal, aiming to use common words used by the target population to avoid professional jargon, long sentences, and any words that might be considered offensive by participants.

Disclosure. Authors have no conflict of interests, and the work was not supported or funded by any drug company.

2) Expert panel and back-translation. After forward translation, a panel of 5 experts (2 family physicians, a preventive medicine specialist, an obstetrician, and a medical researcher with experience in questionnaire translation) reviewed the Arabic version of the questionnaire and compared it to the original English version. All 5 physicians were fluent in both English and Arabic. The panel suggested changing the terminology “عندما لا تكونين في العمل” (“when you are not at work”) in questions 11 and 13 to “عندما تكونين في المنزل” (“when you are at home”) as negative sentences might have been confusing for participants and changing the terminology “during this trimester” to “during the last 3 months” as the word “trimester” is not commonly used in the Arabic language.

Later, back-translation of the translated Arabic questionnaire to English was performed by a certified English linguist who was fluent in Arabic and had good knowledge about local spoken terminology and expressions. The resulting version of the questionnaire was compared with the original version for consistency. Both versions of the questionnaire were found to be consistent.

3) Pre-testing and cognitive interviewing. The translated questionnaire was piloted in 30 participants who were pregnant Saudi women being followed up at KKHU. Face-to-face interviews were conducted with participants to ask about the clarity of questions, assuring the correct understanding of the participants. These participants were not included in further analyses for the validity and reliability of the questionnaire.

4) Final version. The Arabic version of the PPAQ was subjected to validity and reliability testing.

Validity. Face and content validity were assessed by consulting a panel of 10 experts that included 3 obstetricians, 3 family physicians who run antenatal clinics, 2 general physicians, one preventive medicine specialist and one epidemiologist. The panel did not include any of the authors nor the experts who translated the questionnaire.

Experts reviewed the contents and wording of the questions. They also checked the appearance and format of the questionnaire to ensure it was clear and appropriate for the participants. They were asked to evaluate each question on a 4-point likert scale as follows (1-not relevant, 2-item needs some revision, 3-relevant but needs minor revision, 4-very relevant). Later, points 1 and 2 were combined and labeled “not relevant” and points 3 and 4 were also combined and labeled “relevant”.

Item content validity index (I-CVI) was calculated for each question. It ranges between 0 and 1 and was

calculated as the proportion of experts giving the question a rate of 3 or 4 (agreeing on that question).

Scale content validity index (S-CVI) was calculated by 2 methods. The first was scale content validity index based on average (S-CVI_{ave}), which represented the sum of I-CVI divided by the number of items. The second method was scale content validity index based on universal agreement (S-CVI_{UA}), which was calculated as the sum of universal agreement of items divided by the number of items. Universal agreement (UA) score for an item was given as one if it was rated as relevant (points 3 or 4) by all experts; otherwise, the UA score was given as 0.²⁵

Items with an I-CVI more than 0.78 were considered acceptable and those with lower I-CVI should be eliminated or revised. However, S-CVI was considered excellent if it was 0.9 or more.²⁶

Construct validity. The construct validity of this questionnaire was tested by 3 main features of Rasch analysis; item fit to Rasch model, unidimensionality of data, and item difficulty hierarchy. Rasch analysis was performed by Andrich's Rating Scale Method (RSM) as the items of this questionnaire had multiple response categories and fixed intervals between categories.²⁷ Items in each domain of the questionnaire were analyzed separately.

Item fit to Rasch model was tested by infit and outfit mean squares. Fit mean square is the ratio of the observed response to the expected response which should ideally be 1.0. As Rasch model is probabilistic it is expected to find some variations in results. The adequate range for fit mean squares is between 0.5-1.5. However, measures between 1.5-2 can be acceptable.^{28,29}

Unidimensionality means items within one domain measure the same construct. It was tested by evaluating eigenvalues from principal component analysis (PCA) on item residuals. Eigenvalues show how much of the construct is explained by residuals from Rasch model. The smaller the eigenvalues for residuals the stronger the unidimensionality of items in that domain. For each domain, unidimensionality was decided if the ratio-of-the-first-to-second-eigenvalue was less than 3.³⁰

Item difficulty hierarchy. Difficulty level for each item were calculated and expressed as logits, the natural logarithm of the odds of a person being able to perform a certain task. The greater the item logit was, the more difficult the task that item was asking about.³¹ If the items of the tool have theoretical hierarchy, difficulty measurements yielded from Rasch analysis can be compared with theoretical hierarchy in order to test the internal construct validity of the tool.³¹ Items were listed ascendingly according to the activity MET value

considering that activities with lower MET values are easier than the ones with higher MET values. These values were then compared with difficulty level from Rasch analysis to test construct validity. Internal consistency was assessed by measuring item separation index and item reliability. Item separation index was calculated for each domain and showed the ability of items in that domain to distinguish groups of people

according to the differences in their ability levels. Item separation index of 1.5 was considered acceptable, an index of 2 was considered good and an index of 3 was considered excellent.^{31,32} Reliability measure was also calculated for each domain separately and was interpreted in the same way as Cronbach's alpha.³¹

Reliability. In addition to Rasch model, internal consistency was assessed by calculating Cronbach's alpha coefficient for each subscale. A score of less than 0.50 indicated poor internal consistency, scores ranging from 0.51 to 0.69 were considered suspicious, scores ranging from 0.70 to 0.80 were considered acceptable, scores ranging from 0.81 to 0.90 were considered good, and scores greater than 0.90 indicated excellent internal consistency.³³

Test-retest reliability was checked by calculating the intraclass correlation coefficient (ICC) for each subscale. For this method, 59 participants were asked to complete the questionnaire again one or 2 weeks

Table 1 - Demographic characteristics of participants (N=118).

Variables	Total (%)
Age (years) Mean ± SD	30.15 ± 5.59
Pregestational body mass index	
Underweight	2 (2.2)
Normal weight	45 (38.2)
Overweight	28 (30.1)
Obese	18(19.4)
Trimester	
First	6 (5.1)
Second	64 (54.2)
Third	47 (39.8)
Previous abortion	
No	74 (62.7)
Yes	44 (37.3)
Gravity	
Primigravida	33 (28.0)
Multigravida	85 (72.0)
Children	
0 children	24 (20.9)
1-2 children	60 (52.2)
3-4 children	23 (20.0)
>5 children	8 (7.0)
Level of education	
Any school education	22 (20.6)
Bachelor	77 (72.3)
Higher education	8 (7.5)
Occupation	
Housewife	69 (58.5)
Employed	38 (32.1)
Student	2 (1.7)
Others	9 (7.6)
Family income (SAR)	
<5000	26 (33.9)
5000 to <10000	40 (11)
10000 to <15000	13 (7.6)
>15000	9 (25.4)
Don't know/don't want to answer	30 (25.4)
Accommodation	
Apartment	81 (75.7)
Villa	26 (24.3)
Home ownership	
Rented	55 (52.4)
Owned	30 (28.6)
Work housing	1 (1.0)
Living with husband's family	19 (18.1)

SAR: Saudi Riyals

Table 2 - Energy expenditure in METs.hours/week classified by intensity and type of activities.

Activities	25 th quartile	Median	75 th quartile
Total activity	162.3	356.1	648.3
By intensity			
Total (light and greater than light)	58.1	163.8	344.5
Sedentary activity	36.5	135.4	355.1
Light intensity	47.9	125.5	264.9
Moderate intensity	3.5	18.7	83.1
Vigorous intensity	0.0	0.0	0.4
By type			
Household/caregiving	46.7	123.5	227.8
Occupational	0.0	0.0	101
Sports	0.2	0.49	1.8
Transportation	2.2	7.3	22.4

METs.hours/week, metabolic equivalent hours per week

Table 3 - Internal consistency of PPAQ-A using Cronbach's alpha (N=118).

Activity type	Cronbach's alpha
Household activity	0.68
Transportation	0.56
Leisure/sports Activity	0.83
Occupational Activity	0.83

PPAQ-A: pregnancy physical activity questionnaire-Arabic version

Table 3 - Content validity, relevance rating from ten experts.

Item no.	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	I-CVI	UA
1	1	1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	1	1	1	1	1
3	1	1	1	1	1	1	1	1	1	1	1	1
4	1	1	1	1	1	1	1	1	1	1	1	1
5	1	1	1	1	1	1	1	1	1	1	1	1
6	1	1	1	1	1	1	1	1	1	1	1	1
7	1	1	1	1	1	1	1	1	1	1	1	1
8	1	1	1	1	1	1	1	1	1	1	1	1
9	1	1	1	1	1	1	1	1	1	1	1	1
10	1	1	1	1	1	1	1	1	1	1	1	1
11	1	1	1	1	1	1	1	1	1	1	1	1
12	1	1	1	1	1	1	1	1	1	1	1	1
13	1	1	1	1	0	1	1	1	1	1	0.9	0
14	1	1	1	1	1	1	1	0	0	1	0.8	0
15	1	1	1	1	1	1	1	1	1	1	1	1
16	1	1	1	1	1	1	1	1	0	1	0.9	0
17	1	1	1	1	1	1	1	1	1	1	1	1
18	1	1	1	1	1	1	1	1	1	1	1	1
19	1	1	1	1	1	1	1	1	0	1	0.9	0
20	1	1	1	1	1	1	1	1	1	1	1	1
21	1	1	1	1	0	1	1	1	1	1	0.9	0
22	1	1	1	1	1	1	1	1	0	1	0.9	0
23	1	1	1	1	0	1	1	1	1	1	0.9	0
24	1	1	1	1	1	1	1	1	1	1	1	1
25	1	1	1	1	1	1	1	1	1	1	1	1
26	1	1	1	1	1	1	1	1	1	1	1	1
27	1	1	1	1	1	1	1	1	1	1	1	1
28	1	1	1	1	1	1	1	1	1	1	1	1
29	1	1	1	1	1	1	1	1	1	1	1	1
30	1	1	1	0	1	1	1	1	1	1	0.9	0
31	1	1	1	0	1	1	1	1	1	1	0.9	0
32	1	1	1	1	1	1	1	1	1	1	1	1
33	1	1	1	1	1	1	1	1	1	1	1	1
34	1	1	1	1	1	1	1	1	1	1	1	1
35	1	0	1	1	1	1	1	1	1	1	0.9	0
36	1	1	1	1	1	1	1	1	1	1	1	1
Proportion relevance	1	0.97	1	0.94	0.92	1	1	0.97	0.89	1		
S-CVI _{ave}											0.94	
S-CVI _{UA}												0.72
Average proportion of items judged as relevant across the ten experts												0.97

E: expert, I-CVI: item content validity index, UA: universal agreement, S-CVI: scale content validity index, S-CVI_{ave}: scale content validity index based on average, S-CVI_{UA}: scale content validity index based on universal agreement

Table 4 - Rasch analysis, item difficulty and fit statistics.

Domain/item number*	Activity MET value	Difficulty (SE)	Infit mean square (SE)	Outfit mean square (SE)
<i>Household activity</i>				
12 Watching TV	1.0	-0.56 (0.08)	1.30 (2.43)	1.31 (2.46)
13 Sitting and reading or talking on the phone	1.1	-0.88 (0.08)	1.07 (0.62)	1.22 (1.58)
11 Sitting and writing or using computer	1.8	-0.62 (0.06)	1.46 (3.74)	1.77 (5.01)
5 Childcare while sitting	2.0	-0.41 (0.06)	0.76 (-2.35)	0.75 (-1.95)
15 Light cleaning	2.3	-0.48 (0.08)	0.89 (-0.85)	0.84 (-1.19)
16 Shopping	2.3	-0.14 (0.08)	1.14 (1.05)	1.12 (0.85)
4 Preparing meals	2.5	-0.77 (0.08)	0.91 (-0.75)	0.91 (-0.69)
7 Playing with children while sitting or standing	2.7	-0.06 (0.08)	0.80 (-1.55)	0.74 (-1.95)
17 Heavier cleaning in the house	2.8	-0.29 (0.07)	1.04 (0.34)	1.07 (0.51)
18 Mowing lawn on riding mower or cleaning the front or backyards	2.8	0.63 (0.12)	1.00 (0.06)	0.80 (-0.65)
6 Childcare while standing	3.0	-0.21 (0.07)	0.71 (-2.58)	0.66 (-2.71)
9 Carrying children	3.0	0.23 (0.09)	0.91 (-0.36)	0.74 (-1.09)
14 Playing with pets	3.2	1.04 (0.21)	1.06 (0.29)	1.44 (0.99)
8 Playing with children while walking or running	4.0	0.33 (0.10)	0.98 (-0.08)	0.87 (-0.64)
10 Taking care of an older adult	4.0	0.81 (0.18)	1.05 (0.27)	1.07 (0.34)
19 Mowing lawn or gardening	4.4	1.38 (0.24)	1.01 (0.13)	0.99 (0.06)
<i>Transportation</i>				
22 Driving or riding a car or a bus	1.5	-0.55 (0.13)	1.07 (0.49)	0.94 (-0.37)
20 Walking slowly to go places	2.5	0.05 (0.14)	0.86 (-0.81)	0.87 (-0.82)
21 Walking quickly to go places	4.0	0.50 (0.15)	1.04 (0.31)	1.00 (0.05)
<i>Leisure/sports activity</i>				
23 Walking slowly	3.2	-1.03 (0.16)	0.88 (-0.80)	0.88 (-0.82)
27 Prenatal exercise classes	3.5	0.43 (0.32)	0.88 (-0.21)	2.09 (1.22)
29 Dancing	4.5	0.64 (0.27)	1.65 (1.58)	1.37 (0.76)
24 Walking quickly	4.6	-0.57 (0.18)	0.82 (-1.22)	0.75 (-1.63)
28 Swimming	6.0	0.90 (0.36)	0.75 (-0.30)	4.09 (1.76)
25 Walking quickly up hills	6.5	-1.03 (0.20)	1.09 (0.70)	1.08 (0.55)
26 Jogging	7.0	0.66 (0.28)	0.80 (-0.53)	1.57 (0.97)
<i>Occupational activity</i>				
32 Sitting	1.6	-1.71 (0.16)	1.15 (0.76)	1.01 (0.11)
34 Standing or walking slowly not carrying things	2.2	-0.51 (0.16)	0.90 (-0.42)	0.89 (-0.43)
33 Standing or walking slowly while carrying things	3.0	0.32 (0.18)	1.01 (0.13)	0.91 (-0.21)
36 Walking quickly not carrying things	3.3	0.58 (0.18)	0.87 (-0.44)	0.95 (-0.07)
35 Walking quickly while carrying things	4.0	1.32 (0.25)	1.24 (0.92)	0.99 (-0.17)

SE: standard error; *Items listed ascendingly according to the MET value for its activity

after completing the first one. Their responses on both occasions were then checked for consistency. The ICC was calculated for log-transformed data, as PA scores were not normally distributed. Values less than 0.5 were an indication of poor reliability, values between 0.5 and 0.75 indicated moderate reliability, values between 0.75 and 0.9 indicated good reliability, and values greater than 0.90 indicated excellent reliability.³⁴ Rasch analysis

was performed by the Software (jMetrik, version 4.1.1)³⁵ all other analyses were performed by SPSS version 25.³⁶

Results. A total of 118 pregnant women were enrolled in this study. The mean age was 30.15 years (SD=5.59). Most of the participants (n=64) were in their second trimester, followed by the third trimester (n=47). About two-thirds (n=74) did not have previous

Table 5 - Item separation, item reliability and ratio of first-to-second eigen values for residuals from Rasch analysis.

Domain	Item separation	Item reliability	Ratio of first-to-second eigenvalues
Household activity	5.45	0.97	1.23
Transportation	2.83	0.89	1.48
Leisure/sports activity	2.81	0.89	1.30
Occupational activity	5.40	0.97	1.37

Table 6 - Test-retest reliability of PPAQ-A using intraclass correlation coefficients (ICC) with 95% confidence intervals (CI).

Variable	ICC (N=59)	95% CI for ICC
<i>By type of activity</i>		
Household activity	0.85	0.74-0.91
Transportation	0.59	0.30-0.76
Leisure/sports activity	0.70	0.45-0.84
Occupational activity (n=27)	0.15	-0.78-0.60
<i>By intensity</i>		
Vigorous intensity activity	0.61	-0.21-0.87
Moderate intensity activity	0.80	0.61-0.89
Light intensity activity	0.78	0.62-0.87
Sedentary activity	0.65	0.40-0.79
Total activity	0.78	0.55-0.88

PPAQ-A: pregnancy physical activity questionnaire-Arabic version

abortions. Most of the participants (n=69) were housewives, and more than one-third had normal pre-gestational body mass index (n=45) (Table 1).

The median of total energy expenditure in PA was 356.1 METs.h/week (interquartile range [IQR]=162.3-648.3). Most of this was expended in sedentary activities with a median (IQR) of 135.4 (36.5-355.1) METs.h/week. Regarding the types of PA, most of the participants' energy expenditure was in household/caregiving activities (median=123.5), while the least was in occupational activities (median=0.0) (Table 2).

Instrument content validity. Item content validity index ranged between 0.8 and 1 which indicates excellent content validity. However the S-CVI was excellent by the average method (0.94) but was not adequate by the universal agreement method (0.72). Average proportion of items judged as relevant across the 10 experts was 0.97 (Table 3).

Construct validity. Item fit to Rasch model. The infit mean square for all items was between the range of 0.5 - 1.5 except for item number 29 where it was 1.65. Item number 29 did not show perfect fit but it did not

distort the integrity of the model. Outfit mean squares which include outliers were excellent for all items except for items number 11 and 26 that were acceptable and for items number 27 and 28 which were questionable (Table 4).

Unidimensionality. The ratio-of-the-first-to-second-eigenvalue for all domains were less than 3 representing that little information was explained by residuals and thus unidimensionality of items (Table 5).

Item difficulty hierarchy. Item difficulty ranged between -1.71 and 1.38 indicating that the tool covered tasks with a variety of difficulty levels; contrarily, no items were extremely difficult or extremely easy.

When item difficulty measures were compared with activity MET values, the order of item difficulty was consistent MET values (theoretical hierarchy) for transportation and occupational activities indicating high construct validity for these 2 domains. For household activities, the order of difficulty in comparison to MET values was slightly distorted yet easy items were listed before difficult items. However, for leisure/sports activities, the order of MET values and difficulty were consistent except for items 24 and 25 which have high MET value but low scores in difficulty level (Table 4).

Internal consistency. The highest item separation index was for household activities (5.45) and the lowest was for leisure activities (2.81) indicating that at least 3 groups of participants can be distinguished by items in different domains of the questionnaire. In addition, item reliability was 0.97 for household and work activities while it was 0.89 for transportation and leisure activities representing excellent internal consistency for all domains (Table 5).

In addition, Cronbach's alpha coefficient was applied to each subscale. Cronbach's alpha coefficient was good for occupational and leisure/sports activities (0.83), while it was lower for household and transportation activities (0.68 and 0.56, respectively), indicating suspicious internal consistency (Table 6).

Based on the ICC results, the test-retest reliability was good for the total PA (0.78). Regarding the type of activity, the reliability was moderate to good (0.59-0.85), with the highest value recorded for household activity. The exception to this was that for occupational activity, for which the ICC was 0.15; however, this was statistically insignificant. Moreover, the reliability was moderate to good for activity intensity, ranging from 0.61 to 0.80. The highest reliability was for moderate-intensity and lowest was for vigorous activity, although the latter was statistically insignificant (Table 6).

Discussion. Regular PA during pregnancy can help psychological well-being, physical fitness, and weight management.¹⁰ In addition, PA can reduce the risk of gestational diabetes in obese women.¹⁰ Hence, it is essential to establish an instrument for assessing pregnant women's PA to help prevent complications and maintain a healthy pregnancy.

This study involved the translation and cross-cultural adaptation process of the PPAQ from English to Arabic for women living in Saudi Arabia. The availability of this tool in many languages adds value to this study in terms of comparing the results.

The median of the total PA (light and more vigorous in intensity) was 163.8 METs.h/w. This finding is higher than those determined in other studies. However, most of the energy expenditure in PA corresponded to light-intensity activities and household/caregiving activities. This is in line with findings of studies in different populations.^{17-22,37}

In this study, the least amount of energy was consumed in occupational activities. This may be explained by the fact that most of the participants in this study were unemployed (58.5%). However, the type of activity that consumed the second-least amount of energy in this study was sports/exercise. This agrees with the findings of previous studies showing that pregnant women's energy expenditure is the least in occupational activities and sports/exercise.^{17-22,37}

The study finding of low sports exercise activity and vigorous-intensity activity may be explained by the lack of information and common misbeliefs about exercise and sports during pregnancy in Saudi Arabia. These results are consistent with those of other studies conducted in Japan and Taiwan.^{18,38}

The content validity for the Arabic version of the tool was good. Although the S-CVI\UA was not adequate for this questionnaire, both S-CVI and I-CVI were excellent.

Construct validity for this questionnaire was good as all constructs were unidimensional with good fit to Rasch model. In addition, item difficulty order was consistent with theoretical hierarchy in almost all domains except for leisure/sports activities where 2 items (number 24 and number 25) distorted the order. These items were about walking quickly and walking quickly up hills respectively. Although these 2 activities consume higher energy in comparison to other activities, walking in general is a common activity for pregnant women and is performed regularly throughout pregnancy. This might explain the low difficulty level yielded for these 2 items.

Cronbach's alpha coefficients for different types of PA ranged from 0.56 to 0.83. However, Rasch analysis revealed good to excellent item separation index (2.81-5.45) and excellent reliability (0.89-0.97) indicating that this tool has excellent internal consistency.

The current study showed good test-retest reliability for total PA (ICC=0.78). This was comparable with other translations and adaptations of the PPAQ, as the ICC ranged from 0.75 to 0.917.¹⁹⁻²²

During the preparation of this manuscript, an Arabic version of the PPAQ was published.³⁹ It showed validity in a sample of 179 pregnant Lebanese women with different educational backgrounds, socioeconomic statuses, and gestational ages.³⁹ The advantages of the current study are that our tool is a translation of the original English PPAQ without major changes. The researchers retained the same questions and number of items with the same content and type of activities without deleting any items. During the stage of validating the tool, all questions were relevant to the population, and most of the participants answered them. Thus, we believe that this tool might be more applicable to pregnant women in the Arabian Gulf region. In addition, the reliability of this tool was tested at 2-week intervals and showed good repeatability and internal consistency.

There were some limitations in this study, as it was a self-report measure of PA that could have been limited by recall bias. Adding a pedometer in future research, for example, can add a more objective assessment method. Another limitation was the low number of participants in the post-test, which could have weakened the ICC values.

In conclusion, the results of our reliability and validity testing are in line with those of previous studies indicating successful translation and adaptation of the PPAQ to the Saudi/Arabian Gulf culture and acceptable reliability of the Arabic PPAQ.

Acknowledgment. *This research was supported by the College of Medicine Research Center, Deanship of Scientific Research, King Saud University, Riyadh, Saudi Arabia. The authors would like to thank Dr. Ali Al-Hazmi, Dr. Hayfaa Wahabi and Dr. Samia Esmail for revising the initial questionnaire after translation.*

References

1. Physical activity and cardiovascular health. *NIH Consensus Statement* 1995; 13: 1-33.
2. World Health Organization. Global recommendations on physical activity for health. Geneva (SW): World Health Organization; 2010. Available from: <https://www.who.int/publications/i/item/9789241599979>

3. Both M, Overvest M, Wildhagen M, Golding J, Wildschut H. The association of daily physical activity and birth outcome: A population-based cohort study. *Eur J Epidemiol* 2010; 25: 421-429.
4. Dempsey JC, Butler CL, Williams MA. No need for a pregnant pause: physical activity may reduce the occurrence of gestational diabetes mellitus and preeclampsia. *Exerc Sport Sci Rev* 2005; 33: 141-149.
5. Tobias D, Zhang C, van Dam R, Bowers K, Hu F. Physical activity before and during pregnancy and risk of gestational diabetes mellitus. *Diabetes Care* 2011; 34: 223-229.
6. Deierlein A, Siega-Riz A, Evenson K. Physical activity during pregnancy and risk of hyperglycemia. *J Womens Health (Larchmt)* 2012; 21: 769-775.
7. Aune D, Saugstad O, Henriksen T, Tonstad S. Physical Activity and the risk of preeclampsia. *Epidemiology* 2014; 25: 331-343.
8. Wolf H, Owe K, Juhl M, Hegaard H. Leisure time physical activity and the risk of pre-eclampsia: a systematic review. *Matern Child Health J* 2014; 18: 899-910.
9. ACOG Committee Obstetric Practice. ACOG Committee opinion. Number 267, January 2002: exercise during pregnancy and the postpartum period. *Obstet Gynecol* 2002; 99: 171-173.
10. American College of Obstetricians and Gynecologists (ACOG) Physical activity and exercise during pregnancy and the postpartum period: ACOG Committee Opinion, Number 804. *Obstet Gynecol* 2020; 135: 991-993.
11. Poudevigne M, O'Connor P. A review of physical activity patterns in pregnant women and their relationship to psychological health. *Sports Med* 2006; 36: 19-38.
12. Gaston A, Cramp A. Exercise during pregnancy: A review of patterns and determinants. *J Sci Med Sport* 2011; 14: 299-305.
13. Saudi Vision 2030. National Transformation Program (2030). [cited 2018; accessed 2020 Oct 27]. Available from: <https://vision2030.gov.sa/en/programs/NTP>
14. Al-Hazzaa H, AlMarzooqi M. Descriptive analysis of physical activity initiatives for health promotion in Saudi Arabia. *Front Public Health* 2018; 6: 329.
15. Saudi Arabian Ministry of Health. MOH Initiatives (2030). [cited 2018; accessed 2020 Oct 27]. Available from: <https://www.moh.gov.sa/en/Ministry/nehs/Pages/vision2030.aspx> 20.
16. Helmerhorst H, Brage S, Warren J, Besson H, Ekelund U. A systematic review of reliability and objective criterion-related validity of physical activity questionnaires. *Int J Behav Nutr Phys Act* 2012; 9: 103.
17. Chasan-Taber L, Schmidt MD, Roberts DE, Hosmer D, Markenson G, Freedson PS. Development and Validation of a Pregnancy Physical Activity Questionnaire. *Med Sci Sports Exerc* 2004; 36: 1750-1760.
18. Matsuzaki M, Haruna M, Nakayama K, Shiraiishi M, Ota E, Murayama R, et al. Adapting the pregnancy physical activity questionnaire for Japanese pregnant women. *J Obstet Gynecol Neonatal Nurs* 2014; 43: 107-116.
19. Xiang M, Konishi M, Hu H, Takahashi M, Fan W, Nishimaki M, et al. Reliability and validity of a Chinese-translated version of a pregnancy physical activity questionnaire. *Matern Child Health J* 2016; 20: 1940-1947.
20. Chandonnet N, Saey D, Alm eras N, Marc I. French Pregnancy Physical Activity Questionnaire compared with an accelerometer cut point to classify physical activity among pregnant obese women. *PLoS One* 2012; 7: e38818.
21. Krzepota J, Sadowska D. Measuring physical activity during pregnancy - Cultural adaptation of the Pregnancy Physical Activity Questionnaire (PPAQ) and assessment of its reliability in Polish conditions. *Ann Agric Environ Med* 2017; 23: 100-106.
22.  ıracak Y, Yılmaz G, Demir Y, Dalkılınc M, Yaman S. Pregnancy physical activity questionnaire (PPAQ): reliability and validity of Turkish version. *J Phys Ther Sci* 2015; 27: 3703-3709.
23. Helmerhorst H, Brage S, Warren J, Besson H, Ekelund U. A systematic review of reliability and objective criterion-related validity of physical activity questionnaires. *Int J Behav Nutr Phys Act* 2012; 9: 103.
24. World Health Organization. Process of translation and adaptation of instruments. [cited 2010]. Available at: https://www.who.int/substance_abuse/research_tools/translation/en/
25. Yusoff M. ABC of Content Validation and Content Validity Index Calculation. *Education in Medicine Journal* 2019; 11: 49-54.
26. Polit D, Beck C, Owen S. Focus on research methods: Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Res Nurs Health* 2007; 30: 459-467.
27. Tennant A, Conaghan P. The Rasch measurement model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis Rheum* 2007; 57: 1358-1362.
28. Linacre, J. What do infit and outfit, mean-square and standardized mean? *Rasch Measurement Transactions* 2002; 16: 878.
29. Aryadoust V, Ng LY, Sayama H. A comprehensive review of Rasch measurement in language assessment: Recommendations and guidelines for research. *Language Testing* 2021; 38: 6-40.
30. Slocum-Gori S, Zumbo B. Assessing the Unidimensionality of Psychological Scales: Using Multiple Criteria from Factor Analysis. *Social Indicators Research* 2011; 102: 443-461.
31. Duncan P, Bode R, Lai S, Perera S. Rasch analysis of a new stroke-specific outcome scale: The stroke impact scale. *Archives of Physical Medicine and Rehabilitation* 2003; 84: 950-963.
32. Linacre J. A user's guide to Winstepsvr Ministep Rasch-model computer programs. Program Manual 4.4.7. [cited 2019; accessed 2021 Feb 27]. Available at: <https://www.winsteps.com/winman/copyright.htm>
33. Cronbach L. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951; 16: 297-334.
34. Koo T, Li, M. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine* 2016; 15: 155-163.
35. Psychomeasurement Systems. jMetric (Version 4.1.1) Charlottesville, Virginia. [cited 2017]. Available at: <https://itemanalysis.com/jmetric-download/>

36. IBM Corp. IBM SPSS statistics for windows (Version 25.0) [computer software]. New York: IBM Corp.; 2017.
37. de Oliveria Melo A, Silva J, Tavares J, Barros V, Leite D, Amorim M. Effect of a Physical Exercise Program During Pregnancy on Uteroplacental and Fetal Blood Flow and Fetal Growth: A Randomized Controlled Trial. *Obstet Gynecol* 2012; 12: 302-310.
38. Ko YL, Chen CP, Lin PC. Physical activities during pregnancy and type of delivery in nulliparae. *European Eur J Sport Sci* 2016; 16: 374-380.
39. Papazian T, El Osta N, Hout H, Chammas D El, El Helou N, Younes H, et al. Pregnancy physical activity questionnaire (PPAQ): Translation and cross cultural adaption of an Arabic version. *PLoS One* 2020; 15: e0230420.