Ann Ig. 2025 Nov-Dec; 37(6): 807-815. doi: 10.7416/ai.2025.2696

Validation of the Italian version of the Pregnancy Physical Activity Questionnaire

Alessandro Bianconi[1], Rossella Messina[2], Matteo Ricci[1], Vincenza Leccese[1], Dila Parma[1], Virginia Bertini[1], Paola Rucci[2], Sofia Marini[3], Laura Dallolio[2], Alice Masini[4]

Keywords: PPAQ; pregnancy; physical activity; questionnaire; Italian validation; PPAQ-IT
Parole chiave: PPAQ; gravidanza; attività fisica; questionario; validazione italiana; PPAQ-IT

Abstract

Background. *A growing body of evidence suggests that physical activity during pregnancy may have significant health benefits. Self-reported questionnaires may represent an important instrument for gathering information and quantifying the amount of physical activity that women engage during pregnancy. In this study we analyzed the reliability and validity of the Italian version of the Pregnancy Physical Activity Questionnaire.*

Methods. *The reliability and construct validity of the Italian Version of the Pregnancy Physical Activity Questionnaire was analyzed in a sample of pregnant women. Intraclass correlation coefficients (ICCs) were calculated to investigate the test-retest reliability of the Italian Version of the Pregnancy Physical Activity Questionnaire scores at baseline and after 1 week. Intraclass correlation coefficients ≥0.70 were considered to denote an adequate reliability. Validity was assessed using Spearman's rank order correlation coefficients between the Italian Version of the Pregnancy Physical Activity Questionnaire scores and the equivalent measures obtained by accelerometers. Coefficients (ρ) ≥0.50 were considered adequate for construct validity.*

Results. *The test-retest analysis revealed an adequate reliability for sedentary (Intraclass correlation coefficient = 0.785), moderate (Intraclass correlation coefficient = 0.751), and vigorous physical activity (Intraclass correlation coefficient = 0.931), while lower reliability was found for total (Intraclass correlation coefficients = 0.689) and light physical activity (Intraclass correlation coefficients = 0.632). All Spearman's rank correlation coefficients between the Italian Version of the Pregnancy Physical Activity Questionnaire and the accelerometers measures were <0.50 and non-statistically significant.*

Conclusion. *The results of this study showed that the Italian Version of the Pregnancy Physical Activity Questionnaire has almost adequate reliability, while construct validity, based on correlation between Italian Version of the Pregnancy Physical Activity Questionnaire scores and accelerometer data, is very low. Still, the Pregnancy Physical Activity Questionnaire may represent a useful tool to capture daily life activities of pregnant women.*

[1] *Department of Medical and Surgical Sciences, ALMA MATER STUDIORUM-University of Bologna, Bologna, Italy*
[2] *Department of Biomedical and Neuromotor Sciences, ALMA MATER STUDIORUM-University of Bologna, Bologna, Italy*
[3] *Department for Life Quality Studies, ALMA MATER STUDIORUM-University of Bologna, Rimini, Italy*
[4] *Department of Translational Medicine, University of Eastern Piedmont, Novara, Italy*

Annali di Igiene : Medicina Preventiva e di Comunità (Ann Ig)
ISSN (print) 1120-9135 ISSN (online) 3035-5559

https://www.annali-igiene.it

Introduction

A growing body of scientific literature investigated the associations between the levels of physical activity (PA) during pregnancy and health outcomes, such as weight gain, gestational diabetes, and postpartum depression (1). Physical activity during pregnancy is considered safe and it is generally encouraged in pregnant women with no medical contraindications (2). In particular, the World Health Organization (WHO) recommends that women during pregnancy and in the peripartum period without contraindications should perform at least 150 minutes of moderate-intensity aerobic PA during the week (3).

Many experimental studies reported the effects of PA-based intervention on health and gestational outcomes in pregnant women (1). However, meta-analyses underlined a wide variability in the results (4), and guidelines reported low-quality evidence for some outcomes, highlighting the need for further research (3). To conduct high quality studies in different contexts, it is detrimental to use validated measurements to quantify the amount of PA in this target population. Data on levels of PA is generally collected using wearable digital devices, i.e., accelerometers, or estimated from self-reported responses to validated questionnaires (5). While questionnaires may represent an inexpensive alternative to costly objective measurements obtained from wearable devices in restrictively funded research contexts, evidence concerning the properties of self-report PA questionnaires in pregnancy is limited. A systematic review conducted to summarize the measurement properties of this type of questionnaire, recommended the use of the Pregnancy Physical Activity Questionnaire (PPAQ) to assess PA in pregnancy, although the study revealed insufficient construct validity (6).

The PPAQ was developed in 2004 to collect frequency, duration, and intensity of PA during any trimester of pregnancy (7). The questions are structured to gather information about the time spent in various forms of PA, and at different intensities. The original PPAQ has been translated, culturally adapted, and validated in different languages (6,8,9). The various translated versions showed mixed levels of reliability and construct validity (6). To our knowledge, there are no validated Italian versions of the PPAQ in the current literature. The aim of the present study is to provide evidence on the cross-cultural validity and reliability of the Italian version of the PPAQ (PPAQ-IT) in a sample of pregnant women.

Materials and Methods

1. Study design and participants

The present study was conducted as part of an experimental study that evaluated the effects of a PA-based intervention on a wide range of self-reported health and gestational outcomes in pregnant women (10). Pregnant women were recruited between January and September 2022 during routine visits at the Obstetrics Unit of the University Hospital of Bologna (IRCCS Azienda Ospedaliero-Universitaria di Bologna, Italy). The study was approved by the Ethical Review Committee of the Emilia–Romagna Region (Prot. n. 984/2020/Sper/AOUBo of 19/11/2020). Written informed consent was individually obtained prior to participation.

Inclusion criteria were: (a) age $\geq$ 18 years; (b) singleton pregnancy; (c) pregnancy without medical complications; (d) the ability to read and speak Italian; (e) willingness to sign the informed consent for participation.

1.1 Data collection

Participants' informations, including age, weight, height, education level, week of gestation, marital status and occupational status, were gathered at baseline. Data about PA levels were collected using the PPAQ-IT at baseline after one week, and using accelerometer during the week following the baseline assessment.

Specifically, participants were asked to wear accelerometers for a 7-days period to obtain weekly PA levels.

1.2 Accelerometers data

The time spent in active and sedentary behaviors was monitored using ActiLife6 wGT3X-BT accelerometers (Actigraph LLC, Pensacola, FL, USA). The Actigraph accelerometer models GT3X (ActiGraph LLC, Pensacola, FL, USA) are used to monitor objectively the daily PA and sedentary status. The accelerometers were placed on an adjustable belt on the right hip, and were taken off only while bathing, showering, and swimming (11,12). Participants recorded their activities in a diary when they were not wearing the accelerometer. After 1 week of monitoring, participants came back to the hospital and returned the accelerometers. Participants that did not wear the accelerometer for at least 10 hours for 4 days during the monitored week were excluded from the study (11).

The actigraph data were summarized into minutes per week spent doing PA. Freedson Adult counts per

minutes (CPM) cut points were used to categorize data as sedentary, light, moderate, vigorous, and very vigorous activity. In particular, the cut-points used to classify PA levels by intensity were 0-100 CPM for sedentary activity, 100–1951 CPM for light activity, 1952–5724 CPM moderate activity, 5725–9498 CPM for vigorous activity, and >9498 for very vigorous activity (13). Total activity referred for PA from light to above. Accelerometer data were managed with ActiLife 6.13.3 software (ActiGraph, LCC, Pensacola, FL, USA).

1.3 PPAQ

The PPAQ collects the answers to 33 multiple choice questions, each describing a different activity. The 33 activities are aggregated in 4 domains by type of activity: 16 household/caregiving activities, 5 occupational activities, 9 sports/exercise activities, and 3 transportation activities. Participants are required to select, among 6 frequency categories, the one that best approximates the time they spent doing the reported activity during the current trimester of pregnancy. Timeframes span between "none" to "3 or more hours per day" and from "none" to "3 or more hours per week".

The PPAQ responses are used to estimate the weekly energy expenditure for each activity. Time spent in each activity is estimated using the response option selected and expressed in hours/week, and then multiplied by the corresponding Metabolic Equivalent of Task (MET). Results for each activity are expressed in MET-hours/week. PA are also classified in 4 intensity categories: Sedentary, Light, Moderate and Vigorous, using the following cut-offs: ≤1.5 METs Sedentary, >1.5 - <3.0 METs light intensity, 3.0 - <6.0 METs moderate intensity, and ≥6.0 METs Vigorous intensity. Total activity indicates light to vigorous activity. Self-administration of the PPAQ-IT takes approximately 15 minutes.

2.1 Translation and cultural adaptation

The PPAQ was translated to Italian and then translated back to English by a native English speaker. In order to improve the intelligibility of the questionnaire, items were reviewed by a team of experts including 1 psychologist and 2 researchers in sport science. The draft version was administered to a sample of 10 women enrolled during childbirth preparation classes in 2020. Cognitive interviewing methodology was used to assess the readability, comprehensibility, and interpretability of each item (14). Once filled the questionnaire, participants were asked to provide comments on the terminology used in each item. No face and content validity issues emerged. No major revisions were made between the draft version and the final version of the PPAQ-IT.

2.2 Statistical analysis

The sample size was defined to detect a correlation ≥0.40 between the MET-hours/week of each PPAQ intensity domain and the weekly time spent in the corresponding level of PA intensity measured by the accelerometers, with 80% power and a type-I error of 0.05. The minimum sample size was estimated at 46 participants using G*Power Software (15). Continuous variables were summarized as mean and standard deviation or as median and interquartile range, and categorical variables as absolute and relative frequencies.

Intraclass correlation coefficients (ICCs) were calculated to investigate the test-retest reliability between the PPAQ-IT results at baseline (t0) and the results of the second assessment after 1 week (t1). ICCs were calculated for the categories "Total physical activity", the four groups of PA aggregated by intensity ("Sedentary", "Light", "Moderate", and "Vigorous"), and the three groups of PA aggregated by type ("Household/caregiving", "Occupational", and "Sport/exercise"). ICC values $\geq$ 0.70 were considered to denote an adequate reliability of the questionnaire, according to the Qualitative Attributes and Measurement Properties of Physical Activity Questionnaires Checklist (QAPAQ) (16). Data were log-transformed to normalize the frequency distribution of the variables.

To assess the validity of the PPAQ-IT, Spearman's rank order correlation coefficients were computed between the PPAQ-IT results for sedentary, light, moderate and vigorous PA and the equivalent measures obtained by accelerometers. The PPAQ does not discriminate between vigorous and very vigorous PA. To harmonize data between the PPAQ-IT and the actigraph measures, accelerometers-derived time spent doing Vigorous and Very Vigorous PA were aggregated in a single category labeled as "Vigorous + Very Vigorous PA". Thus, for vigorous PA, Spearman's correlation coefficient was calculated between the PPAQ MET-hours/week for vigorous activities and the accelerometers-derived variable "Vigorous + Very Vigorous PA". According to the QAPAQ Checklist (16), correlation coefficients of ≥0.50 were considered adequate for construct validity when comparing questionnaire-derived measures of TPA, MPA, and VPA to accelerometer-derived ones. For the purpose

of this study, the threshold of 0.50 was applied also for sedentary activity and LPA.

In October 2023, an updated version of the PPAQ was published (17). Compared to the first version used for this study, only minor changes on how the questions are formulated were made. However, important updates on the conversion tables used to estimate each activity intensity should be highlighted. The new estimates are based on the 2011 Compendium of Physical Activities MET values (18). Although the PPAQ-IT is based on the translation and cultural adaptation of the first English version, all validation analyses were also replicated using the updated conversion values.

Statistical significance was set at $p \leq 0.05$. All statistical analyses were performed using R-Studio Software (R version 4.3.2, RStudio, PBC, Boston, MA, USA) The code used for the analyses is available on Open Science Framework at the following link: https://osf.io/h96fy/files/osfstorage/664b566e2f167d118c0e7af3.

Results

Of the 54 enrolled women that wore the accelerometer, 51 provided complete data. Participants had a mean age of 34.0 (4.6). The median week of gestation at the time of the study was 33.0 (19.5, 37.0). The complete descriptive analyses have been reported in Table 1.

Test-retest reliability was adequate for Sedentary (0.785, 95% CI 0.653 - 0.871), Moderate PA (0.751, 95% CI 0.601 - 0.850), and Vigorous PA (0.931, 95% CI 0.883 - 0.960), and it was slight below the cut-off of 0.70 for total PA (0.689, 95% CI 0.508 - 0.811) and Light PA (0.632, 95% CI 0.422 - 0.775) (Table 2). Regarding the PPAQ domains by type of activities, adequate ICCs were obtained for Household/caregiving PA (0.790, 95% CI 0.630 - 0.880) and Sport/Exercise (0.764, 95% CI 0.621 - 0.858), but not for Occupational PA (0.686, 95% CI 0.510 - 0.807).

The same analysis repeated using non-log-transformed data revealed a general reduction in the ICCs values, with the ICC for moderate PA falling below the threshold for sufficient reliability (ICC > 0.70), indicating potential overestimation of reliability when using log-transformation (Supplementary Material, Table S1).

As reported in Table 3, replication of the analysis using the revised MET values from the updated version of the PPAQ revealed similar ICCs, except for Moderate PA that was below the threshold of 0.70. The

Table 1 - Summary of the sample characteristics, and results of the PPAQ and accelerometers measurements.

Characteristics	n=51*
Age	34.0 (4.6)
BMI	24.9 (3.9)
Gestation Week	33.0 (19.5, 37.0)
Age of the First Pregnancy	32.0 (30.0, 36.0)
Education Level	
Middle school degree or lower	1 (2%)
High school degree	10 (20%)
University degree	40(79%)
Marital Status	
Married/Cohabiting	44 (86%)
Single	7 (14%)
Employment Status	
Paid Work	33 (65%)
Student	1 (2%)
On Maternity Leave	17 (33%)
PPAQ scores (t0) **	
Total PA	78.2 (47.7, 114.6)
By intensity	
Sedentary	35.2 (15.7, 77.2)
Light PA	58.8 (33.7, 73.4)
Moderate PA	16.6 (8.2, 45.4)
Vigorous PA	0.0 (0.0, 1.6)
By type	
Household/caregiving	32.8 (24.7, 51.8)
Occupational	45.2 (0.0, 67.4)
Sport/exercise	3.6 (0.9, 7.5)
PPAQ scores (t1) **	
Total PA	64.1 (51.1, 90.6)
By intensity	
Sedentary	32.0 (17.2, 73.6)
Light PA	48.2 (32.6, 68.9)
Moderate PA	18.0 (9.2, 30.1)
Vigorous PA	0.0 (0.0, 1.6)
By type	
Household/caregiving	27.4 (19.7, 39.3)
Occupational	16.8 (0.0, 68.7)
Sport/exercise	3.6 (0.8, 6.3)
Actigraph ***	
Total PA	1341.7 (1032.2, 1580.7)
By intensity	
Sedentary	7248.3 (6515.8, 7592.0)
Light PA	1,059.2 (853.9, 1,377.8)
Moderate PA	193.5 (114.7, 298.3)
Vigorous PA	0.8 (0.3, 1.5)
Very Vigorous PA	0.0 (0.0, 0.0)
Vigorous + Very Vigorous PA	0.8 (0.3, 1.7)

* Mean (standard deviation), median (interquartile range), or N (%).
** MET-hours/week. *** minutes/week.
PPAQ = pregnancy physical activity questionnaire. PA = physical activity. MET = Metabolic Equivalent of Task.

Table 2 - Test-retest reliability between physical activity categories of the PPAQ-IT (log-transformed values).

	ICC	95% CI	p-value
Total Physical Activity	0.689	0.508 - 0.811	< 0.0001
By intensity			
Sedentary	0.785	0.653 - 0.871	< 0.0001
Light	0.632	0.422 - 0.775	< 0.0001
Moderate	0.751	0.601 - 0.850	< 0.0001
Vigorous	0.931	0.883 - 0.960	< 0.0001
By type			
Household/caregiving	0.790	0.630 - 0.880	< 0.0001
Occupational	0.686	0.510 - 0.807	< 0.0001
Sport/exercise	0.764	0.621 - 0.858	< 0.0001

ICC = Intraclass correlation coefficient. CI = confidence interval.

Table 3 - Test-retest reliability between physical activity categories of the PPAQ-IT calculated using the updated Metabolic Equivalent of Task values from the novel version of the PPAQ (log-transformed values).

	ICC	95% CI	p-value
Total Physical Activity	0.668	0.485 - 0.796	< 0.0001
By intensity			
Sedentary	0.739	0.585 - 0.842	< 0.0001
Light	0.662	0.477 - 0.792	< 0.0001
Moderate	0.627	0.418 - 0.773	< 0.0001
Vigorous	0.968	0.944 - 0.981	< 0.0001
By type			
Household/caregiving	0.810	0.689 - 0.887	< 0.0001
Occupational	0.692	0.518 - 0.811	< 0.0001
Sport/exercise	0.765	0.622 - 0.858	< 0.0001

ICC = Intraclass correlation coefficient. CI = confidence interval.

Table 4 - Spearman correlation coefficients between intensity categories of physical activity calculated with PPAQ-IT and with accelerometers.

	Rho	p-value
Total Physical Activity	0.005	0.973
By intensity		
Sedentary	-0.087	0.542
Light	-0.072	0.612
Moderate	0.189	0.183
Vigorous	0.236	0.096

rho = Spearman correlation coefficient.

Table 5 - Spearman correlation coefficients between intensity categories of physical activity calculated using the updated Metabolic Equivalent of Task values from the novel version of the PPAQ and using accelerometers.

	Rho	p-value
Total Physical Activity	0.079	0.582
By intensity		
Sedentary	-0.134	0.350
Light	0.056	0.695
Moderate	0.102	0.478
Vigorous	0.236	0.096

rho = Spearman correlation coefficient.

use of non-log-transformed data to compute the latter analysis did not substantially change the interpretation of the results, although a potential overestimation of the ICCs obtained using log-transformed values occurred also in this case (Supplementary Material, Table S2).

Spearman's rank correlation coefficients were very small, indicating no relationship between the PPAQ-IT and the accelerometers measures. In particular, correlation coefficients were $\rho = 0.005$ (p=0.973) for Total PA, $\rho = -0.087$ (p=0.542) for Sedentary, $\rho = -0.072$ (p=0.612) for Light PA, $\rho = 0.189$ (p=0.183) for Moderate PA, and $\rho = 0.236$ (p=0.096) for Vigorous PA. Results concerning construct validity are summarized in Table 4.

As reported in Table 5, the same analysis was replicated using the revised MET values from the updated version of the PPAQ. The newly calculated correlation coefficients for each domain of PA were comparable with the ones reported in Table 5.

Discussion and conclusions

The present study aimed to assess the reliability and construct validity of the PPAQ-IT in pregnant women. The findings indicate a mixed performance of the questionnaire across different domains and intensities of PA. While the PPAQ-IT demonstrated sufficient reliability for certain domains of PA, such as moderate/vigorous PA and sedentary behavior, it exhibited a lower reliability for total and light PA levels. Additionally, the low correlations with objective accelerometer-derived measures raises questions about the validity of self-reported data collected using the PPAQ-IT.

Since the use of log-transformed values in ICC analyses for test-retest reliability assessment can introduce biases in the estimation of reliability, we compared ICCs derived from log-transformed values with those obtained from non-log-transformed values to control discrepancies that may arise. Log transformation is often employed to normalize skewed data distributions and stabilize variance; however, it can distort the interpretation of ICCs (19). In our study, comparing ICCs derived from log-transformed values with those from non-log-transformed values revealed notable differences. Consequently, caution is warranted when interpreting reliability estimates derived from log-transformed data.

Comparing the findings of the main analyses, which utilized the metabolic equivalent (MET) values from the original version of the PPAQ, with the results of the same analyses using MET values from the updated version of the PPAQ, revealed consistent results overall, with the only discrepancy regarding the reliability in measuring moderate PA. Construct validity results remain consistent between the two analyses, with all findings highlighting non-sufficient, non-statistically significant correlation between PPAQ-derived and accelerometers-derived measures.

A floor effect was observed in certain items or domains of the PPAQ-IT, indicating a limitation in capturing the full spectrum of PAs commonly performed by pregnant women (16). In our study, the floor effect was particularly prominent in items such as childcare-related activities and certain occupational tasks. For example, among primiparous women who had yet to experience childcare responsibilities, responses to childcare-related items tended to cluster around the lowest scores, indicating minimal engagement in such activities. Another example may be those participants in maternity leave, who, thus, tended to report minimal occupational activities. This skewing towards lower scores suggests that the questionnaire may not adequately capture the range and diversity of activities performed by all pregnant women. The same applies to infrequent reporting of vigorous PA, for which most of the participants declare to perform it rarely or never. This discrepancy may reflect cultural norms or safety concerns that discourage pregnant women to engage in intense physical efforts (20).

In line with our study, previous research investigating the construct validity of the PPAQ and its translated versions has frequently reported Spearman's rank correlation coefficients below the predetermined threshold of 0.50 when comparing the questionnaire-derived measures with accelerometer-derived ones (6). Different interpretations of the results may have been reported in each single validation study, depending on which thresholds are used for assessing sufficient construct validity. Sattler et al. analyzed different translated version of the PPAQ, and it emerged that coefficients for total PA ranged between $\rho = 0.02$, for the Japanese version of the PPAQ (21), to $\rho = 0.50$, for the French version (22), being the only version reporting a correlation coefficient ≥ 0.50 comparing the PPAQ-IT measured total PA with accelerometer-derived data (6). This heterogeneity may depend on the generally small sample sizes of the studies, and the various study designs that occurs in the literature about the PPAQ, with different CPM cut-points for categorize the accelerometer data by intensity of PA (7,22), and even MET values cut-points for classify PPAQ activity categories that are different with the ones reported in the questionnaire itself (22). These results exhibit a limited comparability between the validity findings of the various PPAQ versions. In particular, Freedson et al. (13) CPM cut-offs were used in this study. According to other studies (7,21), the PA levels derived using Freedson et al. (13) cut-points tend to show lower correlation with the PPAQ-derived values, when compared with other cut-points, such as Swartz et al. (23) and Hendelman et al. (24).

The test-retest reliability coefficients in previous studies on other PPAQ versions was generally higher than in the present study. In some cases, all of ICCs in test-retest analyses over a 1-week period exceeded the threshold of 0.70, indicating adequate reliability over time (6). These higher ICC values suggest greater consistency in participants' responses when completing the PPAQ on separate occasions.

While accelerometers represent a widely accepted objective measure for assessing PA, they could not be considered a "gold standard" for assessing PA levels, since they may present certain limitations that should be considered. One of the primary limitations is the risk for measurement error associated with placement and positioning. Depending on where the accelerometer is worn, certain types of movements may not be adequately captured, or may be inaccurately assessed (7). Additionally, compliance with wearing accelerometers during the day can vary among participants, leading to incomplete data collection (25). These limitations may partly explain the inconsistent results obtained for construct validity.

The PPAQ and similar questionnaires used to assess PA in pregnancy offer several strengths. Firstly, they provide an inexpensive instrument for collecting data on PA behaviors of pregnant women, when compared

to more objective measurements obtained by wearable devices (5). This allows researchers to pursue research attempts in contexts with funding limitations. Additionally, these questionnaires typically capture a wide range of specific activities relevant to pregnancy, providing a comprehensive overview of patients' daily behaviors to their healthcare professionals (5). Moreover, the availability of questionnaires in different languages can facilitate the applicability and universality of the results from various population studies, permitting to compare local PA levels with those of other countries in the world (6-9). Despite their usefulness, the PPAQ and other questionnaires may also have certain limitations. Notably, they are based on self-reported data, which introduces the potential for recall and social desirability biases (5,16). Furthermore, as suggested by the floor effect highlighted in this study, questionnaires may not evaluate all relevant activities (26). Cultural differences and variations in lifestyle behaviors across populations may impact the applicability of questionnaires in different settings (5). These limitations highlight the need for caution when interpreting findings derived from the use of the PPAQ-IT, and underscore the importance of employing complementary objective measures when possible.

This study presents some limitations that should be discussed. First, the already discussed use of accelerometers as a reference standard, which may not capture certain activities accurately, may have hindered the correct assessment of construct validity. In particular, accelerometers placed on the hip may not have fully detected upper body movements and cannot have measured aquatic sports or swimming pool activities. Second, the accelerometers and the PPAQ estimate the levels of PA at different intensities in two different ways. Both the actigraph CPM cut-offs and the PPAQ MET values cut-offs used to discriminate time spent in sedentary behavior, light, moderate, and vigorous activity may be at risk of misclassification bias, leading to potential erroneous estimates of the correlation tests results. Third, while the minimum number of participants required by the a-priori power analyses has been reached, the sample size may still be considered small, increasing the risk of false discovery findings. Further validation analyses with larger sample size may provide more precise estimates of the validity of the PPAQ-IT.

The results of this study showed that the Italian translated and culturally adapted version of the PPAQ is comparable in terms of reliability with other versions of the questionnaire. As reported also in other PPAQ validation studies, correlation tests using accelerometer data to assess construct validity led to inconclusive results. While the PPAQ-IT may represent an alternative instrument to wearable devices for measuring PA levels in pregnant women in research and clinical practice contexts with funding restriction, its results should be interpreted with extreme caution, due to the mixed reliability and the low construct validity. Further research is needed to refine and contextualize survey instruments such as the PPAQ-IT to better capture the diverse activity patterns of pregnant women across different cultural backgrounds.

Acknowledgements

Statements of ethical approval: The study was approved by the Ethical Review Committee of the Emilia–Romagna Region (Prot. n. 984/2020/Sper/AOUBo of 19/11/2020).
Funding: This study received no funding.
Competing interests: All authors declare no conflicts of interest.

Riassunto

Validazione della versione italiana del Pregnancy Physical Activity Questionnaire

Introduzione. Un crescente numero di evidenze suggerisce che l'attività fisica durante la gravidanza possa avere benefici significativi per la salute. I questionari rappresentano uno strumento importante per raccogliere informazioni e quantificare la quantità di attività fisica svolta dalle donne durante la gravidanza. In questo studio abbiamo analizzato l'affidabilità e la validità della versione italiana del Pregnancy Physical Activity Questionnaire.

Metodi. L'affidabilità al test-retest e la validità costruttiva della versione italiana del Pregnancy Physical Activity Questionnaire sono state analizzate in un campione di donne in gravidanza. Sono stati calcolati i coefficienti di correlazione intraclasse per indagare l'affidabilità al test-retest dei punteggi della versione italiana del Pregnancy Physical Activity Questionnaire al baseline e dopo 1 settimana e coefficienti maggiori di 0,70 sono stati considerati indicativi di un'adeguata affidabilità. La validità è stata valutata mediante i coefficienti di correlazione di Spearman tra i punteggi della versione italiana del Pregnancy Physical Activity Questionnaire e le misurazioni equivalenti ottenute dagli accelerometri. Coefficienti (ρ) maggiori di 0,50 sono stati considerati indicativi di validità costruttiva adeguata.

Risultati. L'analisi test-retest ha evidenziato un'adeguata affidabilità per le misurazioni riguardanti l'attività sedentaria (coefficiente di correlazione intraclasse = 0,785), moderata (coefficiente di correlazione intraclasse = 0,751) e vigorosa (coefficiente di correlazione intraclasse = 0,931), mentre è stata riscontrata una minore affidabilità per l'attività totale (coefficiente di correlazione intraclasse = 0,689) e leggera (coefficiente di correlazione intraclasse = 0,632). Tutti i coefficienti di correlazione di Spearman tra le misurazioni della versione italiana del Pregnancy Physical Activity Questionnaire e quelle degli accelerometri erano inferiori a 0,50 e non statisticamente significativi.

Conclusione. I risultati di questo studio suggeriscono che la versione italiana del Pregnancy Physical Activity Questionnaire possiede un'affidabilità parzialmente adeguata, mentre la validità costruttiva, basata sulla correlazione tra i punteggi della versione italiana del Pregnancy Physical Activity Questionnaire e i dati degli accelerometri, risulta essere molto bassa. Ciononostante, il Pregnancy Physical Activity Questionnaire potrebbe comunque rappresentare uno strumento non invasivo utile per rilevare le attività quotidiane delle donne in gravidanza nella pratica clinica.

References

1. DiPietro L, Evenson KR, Bloodgood B, Sprow K, Troiano RP, Piercy KL, et al. Benefits of Physical Activity during Pregnancy and Postpartum: An Umbrella Review. Med Sci Sports Exerc. 2019 Jun;**51**(6):1292-302. doi: 10.1249/MSS.0000000000001941. PMID: 31095086; PMCID: PMC6527310.
2. Physical Activity and Exercise During Pregnancy and the Postpartum Period: ACOG Committee Opinion, Number 804. Obstet Gynecol. 2020 Apr;**135**(4):e178-88. doi: 10.1097/AOG.0000000000003772. PMID: 32217980.
3. WHO guidelines on physical activity and sedentary behaviour. Geneva: World Health Organization; 2020. Licence: CC BY-NC-SA 3.0 IGO.
4. Ruchat SM, Mottola MF, Skow RJ, Nagpal TS, Meah VL, James M, et al. Effectiveness of exercise interventions in the prevention of excessive gestational weight gain and postpartum weight retention: a systematic review and meta-analysis. Br J Sports Med. 2018 Nov 1;**52**(21):1347-56. doi: 10.1136/bjsports-2018-099399. PMID: 30337461.
5. Strath SJ, Kaminsky LA, Ainsworth BE, Ekelund U, Freedson PS, Gary RA, et al. Guide to the Assessment of Physical Activity: Clinical and Research Applications. Circulation. 2013 Nov 12;**128**(20):2259-79. doi: 10.1161/01.cir.0000435708.67487.da. Epub 2013 Oct 14. PMID: 24126387.
6. Sattler MC, Jaunig J, Watson ED, van Poppel MNM, Mokkink LB, Terwee CB, et al. Physical Activity Questionnaires for Pregnancy: A Systematic Review of Measurement Properties. Sports Med. 2018 Oct 1;**48**(10):2317-46. doi: 10.1007/s40279-018-0961-x. PMID: 30094797; PMCID: PMC6132497.
7. 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 Oct;**36**(10):1750–60. doi: 10.1249/01.mss.0000142303.49306.0d. PMID: 15595297.
8. Rovcanin M, Jankovic S, Mikovic Z, Sipetic Grujicic S, Ersk IRB, et al. The Translation and Cross-Cultural Adaptation of the Pregnancy Physical Activity Questionnaire: Validity and Reliability of a Serbian Version (PPAQ-SRB). Healthc Basel Switz. 2022 Aug 7;**10**(8):1482. doi: 10.3390/healthcare10081482. PMID: 36011143; PMCID: PMC9408768.
9. Santos PC, Maciel LYS, Abreu S, Mesquita AR, Mesquita CC, Lopes S, et al. Cultural adaptation and validation of the "Pregnancy Physical Activity Questionnaire" for the Portuguese population. PLoS One. 2023 Jan 10;**18**(1):e0279124. doi: 10.1371/journal.pone.0279124. PMID: 36626393; PMCID: PMC9831324.
10. Marini S, Parma D, Masini A, Bertini V, Leccese V, Caravita I, et al. Co-Design and Evaluation of the Feasibility and the Efficacy of a Multiple-Targeted Adapted Physical Activity Intervention to Promote Quality of Life, Well-Being and Physical Activity Levels in Pregnant Women: The "WELL-DONE!" Study Protocol. Sustainability. 2021 Jan;**13**(21):12285. https://doi.org/10.3390/su132112285.
11. Migueles JH, Cadenas-Sanchez C, Ekelund U, Delisle Nyström C, Mora-Gonzalez J, Löf M, et al. Accelerometer Data Collection and Processing Criteria to Assess Physical Activity and Other Outcomes: A Systematic Review and Practical Considerations. Sports Med. 2017 Sep;**47**(9):1821-45. doi: 10.1007/s40279-017-0716-0. PMID: 28303543; PMCID: PMC6231536.
12. Rich C, Geraci M, Griffiths L, Sera F, Dezateux C, Cortina-Borja M. Quality Control Methods in Accelerometer Data Processing: Identifying Extreme Counts. PLoS One. 2014 Jan 13;**9**(1):e85134. doi: 10.1371/journal.pone.0085134. PMID: 24454804; PMCID: PMC3890298.
13. Freedson PS, Melanson E, Sirard J. Calibration of the Computer Science and Applications, Inc. accelerometer. Med Sci Sports Exerc. 1998 May;**30**(5):777-81. doi: 10.1097/00005768-199805000-00021. PMID: 9588623.
14. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine. 2000 Dec 15;**25**(24):3186-91. doi: 10.1097/00007632-200012150-00014. PMID: 11124735.
15. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods. 2007 May 1;**39**(2):175-91. doi: 10.3758/bf03193146. PMID: 17695343.
16. Terwee CB, Mokkink LB, van Poppel MNM, Chinapaw MJM, van Mechelen W, de Vet HCW. Qualitative Attributes and Measurement Properties of Physical Activity Questionnaires. Sports Med. 2010 Jul 1;**40**(7):525-37. doi: 10.2165/11531370-000000000-00000. PMID: 20545379.
17. Chasan-Taber L, Park S, Marcotte RT, Staudenmayer J, Strath S, Freedson P. Update and Novel Validation of a Pregnancy Physical Activity Questionnaire. Am J Epidemiol. 2023 Oct 10;**192**(10):1743-53. doi: 10.1093/aje/kwad130. PMID: 37289205; PMCID: PMC11484608.
18. Ainsworth BE, Haskell WL, Herrmann SD, Meckes N, Bassett DR Jr, Tudor-Locke C, et al. 2011 Compendium of Physical Activities: a second update of codes and MET values. Med Sci Sports Exerc. 2011 Aug;**43**(8):1575-81. doi: 10.1249/MSS.0b013e31821ece12. PMID: 21681120.
19. Hammouri HM, Sabo RT, Alsaadawi R, Kheirallah KA. Handling Skewed Data: A Comparison of Two Popular Methods. Appl Sci. 2020 Jan;**10**(18):6247. https://doi.org/10.3390/app10186247.
20. Guelfi KJ, Wang C, Dimmock JA, Jackson B, Newnham JP, Yang H. A comparison of beliefs about exercise during pre-

gnancy between Chinese and Australian pregnant women. BMC Pregnancy Childbirth. 2015 Dec 22;**15**(1):345. doi: 10.1186/s12884-015-0734-6. PMID: 26693690; PMCID: PMC4689036.
21. Matsuzaki M, Haruna M, Nakayama K, Shiraishi M, Ota E, Murayama R, et al. Adapting the Pregnancy Physical Activity Questionnaire for Japanese Pregnant Women. J Obstet Gynecol Neonatal Nurs. 2014 Jan 1;**43**(1):107-16. doi: 10.1111/1552-6909.12267. Epub 2013 Dec 19. PMID: 24428147.
22. Chandonnet N, Saey D, Alméras 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 Jun 11;**7**(6):e38818. doi: 10.1371/journal.pone.0038818. Epub 2012 Jun 11. PMID: 22701717; PMCID: PMC3372468.
23. Swartz AM, Strath SJ, Bassett DR, O'Brien WL, King GA, Ainsworth BE. Estimation of energy expenditure using CSA accelerometers at hip and wrist sites. Med Sci Sports Exerc. 2000 Sep;**32**(9 Suppl):S450-456. doi: 10.1097/00005768-200009001-00003. PMID: 10993414.
24. Hendelman D, Miller K, Baggett C, Debold E, Freedson P. Validity of accelerometry for the assessment of moderate intensity physical activity in the field. Med Sci Sports Exerc. 2000 Sep;**32**(9 Suppl):S442-S449. doi:10.1097/00005768-200009001-00002. PMID: 10993413.
25. O'Brien WJ, Shultz SP, Firestone RT, George L, Breier BH, Kruger R. Exploring the challenges in obtaining physical activity data from women using hip-worn accelerometers. Eur J Sport Sci. 2017 Aug;**17**(7):922-30. doi: 10.1080/17461391.2017.1323952. Epub 2017 May 15. PMID: 28504054.
26. Shephard RJ. Limits to the measurement of habitual physical activity by questionnaires. Br J Sports Med. 2003 Jun;**37**(3):197-206; discussion 206. doi: 10.1136/bjsm.37.3.197. PMID: 12782543; PMCID: PMC1724653.

Corresponding author: Dr. Matteo Ricci, Department of Medical and Surgical Sciences, ALMA MATER STUDIORUM-University of Bologna, Bologna, Italy
e-mail: matteo.ricci18@studio.unibo.it

SUPPLEMENTARY MATERIAL

Table S1 - Test-retest reliability between physical activity categories of the PPAQ-IT (non-log-transformed values)

	ICC	95% CI	p-value
Total Physical Activity	0.604	0.387 - 0.755	< 0.0001
By intensity			
Sedentary	0.756	0.609 - 0.853	< 0.0001
Light	0.498	0.250 - 0.683	0.0001
Moderate	0.636	0.442 - 0.774	0.0001
Vigorous	0.847	0.826 - 0.940	< 0.0001
By type			
Household/caregiving	0.803	0.658 - 0.887	< 0.0001
Occupational	0.460	0.220 - 0.650	0.0002
Sport/exercise	0.860	0.767 - 0.917	< 0.0001

ICC = Intraclass correlation coefficient. CI = confidence interval.

Table S2 - Test-retest reliability between physical activity categories of the PPAQ-IT calculated using the updated Metabolic Equivalent of Task values from the novel version of the PPAQ (non-log-transformed values)

	ICC	95% CI	p-value
Total Physical Activity	0.509	0.279 - 0.686	< 0.0001
By intensity			
Sedentary	0.663	0.478 - 0.792	< 0.0001
Light	0.599	0.391 - 0.749	< 0.0001
Moderate	0.412	0.162 - 0.614	0.0010
Vigorous	0.908	0.844 - 0.946	< 0.0001
By type			
Household/caregiving	0.843	0.741 - 0.907	< 0.0001
Occupational	0.393	0.140 - 0.600	0.0016
Sport/exercise	0.851	0.753 - 0.912	< 0.0001

ICC = Intraclass correlation coefficient. CI = confidence interval.