

Validity and reliability of the International Physical Activity Questionnaire (IPAQ) long-form in a subsample of female Mexican teachers

Catalina Medina, DSc,⁽¹⁾ Adriana Monge, MSc,⁽²⁾ Edgar Denova-Gutiérrez, DSc,⁽¹⁾ Ruy López-Ridaura, PhD,⁽³⁾ Simón Barquera, PhD,⁽¹⁾ Isabelle Romieu, PhD,⁽⁴⁾ Martín Lajous, PhD.^(2,5)

Medina C, Monge A, Denova-Gutiérrez E, López-Ridaura R, Barquera S, Romieu I, Lajous M. Validity and reliability of the International Physical Activity Questionnaire (IPAQ) long-form in a subsample of female Mexican teachers. *Salud Publica Mex.* 2022;64:57-65. <https://doi.org/10.21149/12889>

Medina C, Monge A, Denova-Gutiérrez E, López-Ridaura R, Barquera S, Romieu I, Lajous M. Validez y confiabilidad del Cuestionario Internacional de Actividad Física versión larga en una submuestra de maestras mexicanas. *Salud Publica Mex.* 2022;64:57-65. <https://doi.org/10.21149/12889>

Abstract

Objective. To evaluate the test-retest reliability and validity of the long-form International Physical Activity Questionnaire (IPAQ-long form). **Materials and methods.** A subsample of 86 teachers underwent telephone interviews and clinic visits in Mexico City. Through 22 months, physical activity (PA) levels were obtained from accelerometers (months 6 and 22) and IPAQ long-form (months 1, 3, 6, 9 and 12). Minutes per week of moderate (MPA), walking (WPA), vigorous (VPA) and moderate-to-vigorous physical activity (MVPA) were calculated for each intensity and domain. Variables were logarithmically transformed. Intra-class and Pearson correlations were used to determine the relationship between instruments. **Results.** Test-retest reliability of the average IPAQs long-form varied across intensities and domains (ranged from $r = 0.24$ to $r = 0.65$). Minutes per week of VPA and MVPA accelerometer 1 (AC1) were fairly associated to International Physical Activity Questionnaire-long form 3 (IPAQ3) ($r = 0.60$ and $r = 0.31$, respectively) one week apart. **Conclusions.**

Resumen

Objetivo. Evaluar la confiabilidad y validez del Cuestionario Internacional de Actividad Física versión larga (por sus siglas en inglés IPAQ-long form). **Material y métodos.** Un total de 86 maestras realizaron entrevistas y visitas a clínicas en la Ciudad de México. En un periodo de 22 meses, se obtuvieron los niveles de actividad física (AF) por medio de un acelerómetro (AC) (meses 6 y 22) y autorreporte de actividad física por medio del IPAQ versión larga (meses 1, 3, 6, 9 y 12). Se estimaron los minutos por semana de actividad física moderada (AFM), caminando (AFC), vigorosa (AFV) y actividad física moderada-vigorosa (AFMV) para cada intensidad y dominio. Las variables se transformaron logaritmicamente. Se utilizaron correlaciones de intraclass y de Pearson para determinar la relación de los minutos por semana de AFM, AFC, AFMV entre instrumentos. **Resultados.** La confiabilidad entre el promedio de los IPAQs varió entre intensidades y dominios (de $r = 0.24$ a $r = 0.65$). Los minutos por semana de AFV y AFMV del acelerómetro 1

- (1) Center for Nutrition and Health Research, National Institute of Public Health. Mexico City, Mexico.
- (2) Center for Research on Population Health, National Institute of Public Health. Mexico City, Mexico.
- (3) National Center for Preventive Programs and Disease Control, Ministry of Health. Mexico City, Mexico.
- (4) Section of Nutrition and Metabolism, International Agency for Research on Cancer. Lyon, France.
- (5) Department of Global Health and Population, Harvard University T.H. Chan School of Public Health. Boston, United States.

Received on: May 28, 2021 • Accepted on: September 9, 2021 • Published online: February 25, 2022

Corresponding author: Simón Barquera. Director del Centro de Investigación en Nutrición y Salud. Instituto Nacional de Salud Pública. Av. Universidad 655, col. Santa María Ahuacatitlán. 62100 Cuernavaca, Morelos, México
email: sbarquera@insp.mx

License: CC BY-NC-SA 4.0

IPAQ long-form is fairly reliable for MVPA and valid for assessing MVPA and VPA in a subsample of Mexican teachers.

Keywords: validity; women; measurement; questionnaire; accelerometer

(AC1) se asociaron con los del Cuestionario Internacional de Actividad Física 3 (IPAQ3) ($r = 0.60$ y $r = 0.31$, respectivamente) reportado una semana después. **Conclusiones.** La versión larga del IPAQ es confiable para estimar AFMV y válido para evaluar AFMV y AFV en una submuestra de maestras mexicanas.

Palabras clave: validez; mujeres; mediciones; cuestionario; acelerómetro

Non-communicable chronic diseases (NCDs), including cancer, are the leading cause of death globally.¹ Among the main risk factors include the lack of physical activity (PA).² Measurement of physical activity levels is crucial to identify current levels of activity, to assess the effectiveness of intervention and/or prevention programs and to estimate its association with NCDs.³

Assessment of physical activity is possible through a variety of objective and subjective methods. Accelerometers (AC) have the advantage of measuring physical activity levels through acceleration, however, this instrument could underestimate some activities, specifically, aquatic and upper body movements.^{4,5} Self-reported questionnaires are the most common, low-cost, and easy to apply method used to measure physical activity in population-based studies.³ However, these instruments have the disadvantage of misclassifying physical activity levels and thus compromise the ability to detect true associations.³

One of the most commonly used questionnaires is the International Physical Activity Questionnaire (IPAQ), both the short and long version.⁶ The short version has been validated in several settings including Mexico, with a test-retest reliability value of 0.55 and validity value of 0.31 when compared with accelerometer values.^{6,9} However, there is a lack of information related to the reliability and validity of the long version within the Mexican context.

Therefore, the purposes of this study were to determine the test-retest reliability and validity of the IPAQ long-form in a subsample of female Mexican teachers from the Mexican Teachers' Cohort (MTC).

Materials and methods

The MTC is a prospective study of 115 345 female teachers aged 25 years and older in Mexico. This study began in 2006 in two states and expanded to 10 more states in 2008. Teachers are members of the Teachers Incentives Program (TIP) (or *carrera magistral* in Spanish). This is

a well-established voluntary economic incentives program for public education teachers.^{10,11}

Participants. In 2013, a random sample of 161 women aged ≥ 35 years from Mexico City were invited to participate in a validation study. In total, 99 teachers agreed to participate. Written informed consent was obtained from all participants. There were no statistically significant differences in mean weight, height, body mass index (BMI), and hip circumference between teachers that agreed to participate and those that disagreed. However, a statistically higher mean waist circumference was observed for those teachers that did not want to participate (data not shown).

Procedure. Teachers underwent telephone interviews and trained staff collected anthropometric data during 22 months. In the first month, teachers answered the IPAQ long-form (IPAQ1) by phone and underwent anthropometric measurements. At month three, a second IPAQ long-form (IPAQ2) was administered by phone. At month six, teachers visited the clinic and received an Actical accelerometer (AC1) (Mini Mitter Company, Bend, OR, USA). They were asked to wear this device, using an elastic belt all the time (except when in water) for the following nine days. In addition, teachers received a daily log that contained illustrated instructions on how to wear the accelerometer, removals, frequently asked questions and support contacts. After nine days, teachers returned the accelerometer, and a third IPAQ long-form (IPAQ3) was administered. At month nine, a fourth IPAQ long-form (IPAQ4) was self-reported by phone. During month 12, teachers answered a fifth IPAQ long-form (IPAQ5) by phone. And at month 22, a subsample of 69 teachers received and wore an Actical accelerometer (AC2) during seven consecutive days. The study was reviewed and approved by the *Instituto Nacional de Salud Pública* institutional review board (No. 1130). Written informed consent was obtained from all participants.

IPAQ long-form. The IPAQ was developed by a group of experts to understand and compare physical activity patterns around the world.⁶ This is a 31-item questionnaire that estimates physical activity levels related to household and yard work activities, work-related activity, self-powered transport, walking, leisure time and sedentary activities during the last seven days in bouts of at least 10 minutes.⁶ It can be applied by either face-to-face or telephone interview in 15 minutes.⁶ The Spanish version of the questionnaire is available within the IPAQ webpage.¹²

Data cleaning of the IPAQ long-form was performed based on the IPAQ protocol according to the following criteria:¹³ 1) data collected in hours were converted into minutes, 2) results reported as a weekly frequency were converted into an average daily time, and 3) "do not know", "refused" or "missing data" for duration or frequency were removed from the analysis. Truncation was performed for all daily duration values exceeding 180 minutes.¹³

Days and minutes per day of vigorous, moderate and walking were multiply to obtain minutes per week of vigorous (VPA), moderate (MPA) and walking (WPA) physical activity for each domain. Then, moderate and walking minutes per week were summed to create (MWPA) during household and yard activities, work-related, transportation, walking and leisure time. In addition, minutes per week of MWPA and VPA were added to create total moderate to vigorous physical activity (MVPA) minutes per week.

Actical accelerometer. The Actical is a waistband-mounted omnidirectional accelerometer (dimensions: 2.8 x 2.7 x 1.0 centimeters, weight: 17 grams), which has been previously used in adults to validate and measure physical activity.¹⁴ The Actical accelerometer output is expressed as total counts per registered time (counts-min-1).

Data management and cleaning were performed using IBM SPSS software; version 24 (SPSS Inc., an IBM company Chicago Illinois, United States) and the Personal Activity and Location Measurement System (University of California, San Diego, California, United States). Compliance criteria for wearing accelerometers were defined as a minimum of four weekdays of wearing the accelerometer for at least 10 hours per day. All periods of 60 or more consecutive minutes with zero epochs were removed prior to calculating wear time for a given day.^{15,16} To determine physical activity levels for participants, we used the following cut-points for each set of epochs (minutes of physical activity data):¹⁴ sedentary (≤ 1.5 Mets, <100 accelerometer counts) or

engaged in light (1.5 to 2.9 Mets, 100 to $<1\ 500$ accelerometer counts), moderate (3.0 to 5.9 Mets, 1 500 to $<6\ 500$ accelerometer counts), or vigorous (≥ 6 Mets, $\geq 6\ 500$ accelerometer counts) intensity activity. Finally, MVPA that occurred in bouts of at least 10 consecutive minutes (with allowance of two minutes per each 10 minutes below the 1 530 epoch cut-point) was summed for each valid day and averaged to determine weekly moderate, vigorous, and MVPA.

Physical activity classification (average IPAQs and average AC). Teachers with complete IPAQs and AC data (n=58) were classified as inactive if they accumulated <150 minutes of MVPA, moderately active if they achieved 150-299 minutes of MVPA, and active if they accumulated 300 or more minutes of MVPA based on WHO physical activity recommendations.¹⁷

Anthropometry. Trained personal measured weight and height to the nearest 0.1 kg and 0.1 cm, and the BMI was calculated as kg/m^2 .¹⁸ Waist and hip circumferences were measured to the nearest 0.1 cm.

Statistical analysis. Descriptive analyses (means and standard deviation) were used to characterize the study population. Kolmogorov-Smirnov test was used to test normality in all physical activity variables. Variables that did not meet the normality criteria were logarithmically transformed. Log-transformed mean minutes per week per instrument were compared using student t-test.

Test-retest reliability for the log-transformed work-related, transportation, household, leisure-time, MWPA, VPA and MVPA per week of two accelerometers and five IPAQs long-form were tested using Pearson correlation coefficients. Linear Regression was used to estimate 95% confidence intervals (CI) for the correlation between instruments.

Pearson correlation coefficients and Intra-class correlation (two-way mixed, average measurements and absolute agreement) were used to determine the relationship between log-transformed minutes per week of MWPA, VPA, MVPA for the log-transformed minutes per week of five IPAQ long-form and accelerometers.

A linear regression was used to assess the association between MVPA minutes per week of IPAQ3 and AC1. The intercept and slopes of the regression line and their associated 95%CI was examined to observe if values were different from 0 and 1, respectively. Level of significance was set at $p<0.05$.

Results

Descriptive characteristics

Characteristics of participants at baseline are described in table I. In total, 86 women answered the IPAQs and on average 61 had two valid accelerometer values. The mean age was 48.2 ± 6.8 years. The mean BMI was 28.2 ± 4.8 kg/m².

Physical activity means and differences

With respect to intensities of physical activity, higher MVPA and MWPA were observed on IPAQ2 (504 ± 421 and 463 ± 393 min/week, respectively) than IPAQ3 (387 ± 289 and 341 ± 251 min/week, respectively), IPAQ4 (328 ± 254 and 289 ± 223 min/week, respectively), and IPAQ5 (333 ± 306 and 300 ± 295 min/week, respectively). There were no differences between IPAQ1 and IPAQ2 for all intensities. MWPA minutes per week were significantly higher based on IPAQ1 and IPAQ3 compared to IPAQ4 and IPAQ5 (451 ± 350 and 341 ± 251 vs. 289 ± 223 and 300 ± 295 , respectively). Based on domains of physical activity, teachers reported higher minutes per week of household physical activities for all IPAQs compared to other domains ($p < 0.01$) (figure 1).

Related to household physical activity, statistically significant differences were observed between IPAQ1 and IPAQ3, IPAQ4 and IPAQ5 (200 ± 198 and 162 ± 106 , 138 ± 133 , 126 ± 174 minutes per week, respectively) ($p < 0.05$). The same pattern was observed for active transportation (IPAQ1: 100 ± 114 vs. IPAQ3: 72 ± 92 , IPAQ4: 63 ± 100 , IPAQ5: 66 ± 93 minutes per week, $p < 0.05$). Leisure time physical activity was significantly different between IPAQ2 and IPAQ5, and IPAQ3 and IPAQ5 (111 ± 142 and 83 ± 130 minutes per week) and (106 ± 162 and 83 ± 130 minutes per week) respectively

Table I
CHARACTERISTICS OF PARTICIPANTS IN A STUDY ON THE RELIABILITY AND VALIDITY OF IPAQ LONG-FORM (N= 86). MEXICO, 2013-2015

	Mean	SD*
Age (years)	48.2	6.8
Weight (kg)	68.1	12.3
Height (m)	155.4	6.2
Body mass index (kg/m ²)	28.2	4.8
Waist circumference (cm)	88.2	10.9
Hip circumference (cm)	104.4	9.9

* Standard deviation

($p < 0.05$). Work-related physical activity was statistically significant different between IPAQ2 and IPAQ3 (66 ± 187 vs. 46 ± 115 minutes per week) (figure 1).

Based on accelerometer measures, there were no statistically significant differences between minutes per week in AC1 and AC2 at any intensity. Statistically significant differences were found between average accelerometers and average IPAQs long-form for all intensities ($p < 0.001$) (figure 1).

Reliability

The test-retest correlation between five IPAQs long-form is presented in table II. IPAQ1 minutes per week of MWPA, VPA, MVPA, work-related, transportation, household and leisure-time physical activity were significantly correlated to those minutes of IPAQ2, IPAQ3, IPAQ4 and IPAQ5 with r values ranging from 0.29 to 0.59. The same result was observed for the correlation between minutes per week of MWPA, VPA and MVPA, work-related, transportation, household and leisure-time physical activity IPAQ2 vs. IPAQ3, IPAQ4, IPAQ5, IPAQ3 vs. IPAQ4 and IPAQ5, and IPAQ4 vs. IPAQ5 (range: 0.26 to 0.65), except for the correlation between minutes per week of VPA IPAQ2 and IPAQ4 and minutes per week of transportation IPAQ2 and IPAQ5.

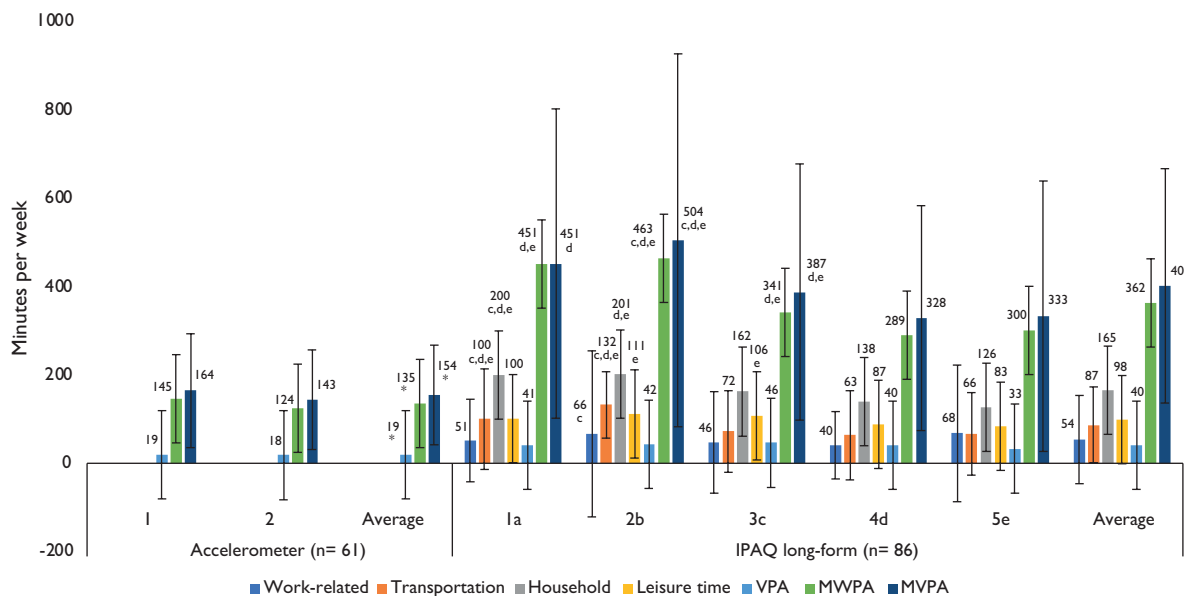
Based on accelerometer estimates, a high correlation ($r \geq 0.60$) was observed between AC1 and AC2 minutes per week of MWPA, VPA and MVPA (table II).

Validity

Compared to the log-transformed minutes per week AC1 values, VPA and MVPA were significantly correlated to average log-transformed minutes per week IPAQ long-form values (Pearson correlation: 0.31, 95%CI: 0.11, 0.47 and 0.25, 95%CI: 0.04, 0.55, respectively). Long-transformed minutes per week AC1 VPA and MVPA measures were significantly correlated ($p < 0.05$) to IPAQ3 values $r = 0.44$ and $r = 0.24$, respectively. Higher correlations were observed using Intra-class correlation for all intensities (table III).

As shown in figure 2, when IPAQ3 MVPA measures were used to predict accelerometer 1, the intercept was 1.57 ($p < 0.001$) and the slope was 0.22 ($p < 0.05$). This interception suggests that those teachers with lower MVPA measured by accelerometers over-reported MVPA levels, however, this over-report decreased as objectively measured MVPA increased. The interception between the line of identity and the Pearson correlation line was at 126 minutes per week of MVPA.

The distribution of teachers across physical activity categories was significantly different between average



MWPA: moderate and walking activities
 VPA: vigorous physical activity
 MVPA: moderate-to-vigorous physical activity
 Line: standard deviation
 Number: mean

* Statistically significant differences between different methods ($p < 0.05$)
 a,b,c,d,e: Statistically significant differences between IPAQs (a- IPAQ1, b- IPAQ2, c- IPAQ3, d- IPAQ4, e- IPAQ5) ($p < 0.05$)

FIGURE 1. ACCELEROMETERS AND IPAQs-LONG FORM MEASURED PHYSICAL ACTIVITY (MEAN AND STANDARD DEVIATION). MEXICO 2013-2015

IPAQs and average AC. Based on average AC, 56.9% were classified as inactive, 34.5% as moderately active and 8.6% as active. Whereas, 12.1% were classified as inactive, 27.6% as moderately active and 60.3% as active based on average IPAQs (data not shown).

Discussion

The main purposes of this study were to determine the test-retest reliability and validity of the IPAQ long-form in a sub sample of Mexican teachers. Our findings indicated that minutes per week of three different intensities of physical activity (VPA, MWPA and MVPA) between IPAQs long-form were correlated (range: $r = 0.26$ to 0.65). Minutes per week of four domains of physical activity (work-related, transportation, housework and leisure-time) had a correlation between 0.24 and 0.64 . Furthermore, minutes per week of MVPA and VPA AC1 were correlated to average MVPA and VPA IPAQs ($r = 0.25$ and $r = 0.31$), and MVPA and VPA estimates of AC1 were correlated with IPAQ3-long form measures answered one-week apart ($r = 0.24$ and $r = 0.44$).

Teachers reported higher average mean values for all intensities based on IPAQs long-form compared to the Actical accelerometer estimates. This over-report has been observed in other self-reports¹⁹ and in the IPAQ short-form for Mexican population.⁷ The overestimation of physical activity levels by the questionnaire could be due to social desirability and/or social approval bias. However, the underestimation of the Actical accelerometer could be related to the lack of record of water activities and strength exercises such as weight lifting.^{4,5,7}

Data from this study showed that in three out of five self-reported IPAQs teachers accumulated more than 150 minutes per week on household physical activities. In other words, teachers achieved physical activity recommendations only by performing activities at home. In addition, compared to other domains, higher number of minutes per week of household physical activity were reported. The same pattern has been observed in previous Latin American studies.²⁰ There are many possible reasons of this observation; first, there could be a misclassification of physical activity intensities during

Table II
TEST-RETEST RELIABILITY OF THE ACCELEROMETER AND THE IPAQ LONG-FORM. MEXICO 2013-15

(Min/week) r (CI95%)	Accelerometer (n= 61)					IPAQ long-form (n= 86)					
	1,2	1,2	1,3	1,4	1,5	2,3	2,4	2,5	3,4	3,5	4,5
Domains of PA											
Work-related	NA	0.40 (0.20,0.60)	0.39 (0.19,0.61)	0.46 (0.28,0.68)	0.36 (0.14,0.52)	0.45 (0.26,0.66)	0.37 (0.17,0.59)	0.51 (0.31,0.66)	0.38 (0.18,0.58)	0.34 (0.13,0.50)	0.38 (0.16,0.52)
Transportation	NA	0.35 (0.15,0.58)	0.36 (0.14,0.51)	0.34 (0.13,0.56)	0.36 (0.16,0.56)	0.48 (0.25,0.58)	0.40 (0.19,0.59)	0.20 (-0.01,0.39)	0.51 (0.40,0.79)	0.42 (0.24,0.67)	0.24 (0.02,0.43)
Household	NA	0.53 (0.29,0.60)	0.44 (0.19,0.50)	0.51 (0.25,0.53)	0.57 (0.30,0.57)	0.38 (0.16,0.54)	0.32 (0.10,0.47)	0.42 (0.25,0.60)	0.36 (0.15,0.54)	0.31 (0.09,0.50)	0.42 (0.22,0.62)
Leisure-time	NA	0.49 (0.31,0.60)	0.59 (0.42,0.78)	0.56 (0.39,0.76)	0.45 (0.25,0.64)	0.50 (0.30,0.67)	0.49 (0.29,0.67)	0.53 (0.33,0.69)	0.59 (0.42,0.77)	0.56 (0.37,0.73)	0.64 (0.47,0.80)
Intensity of PA											
MWPA	0.60 (0.39,0.80)	0.35 (0.19,0.70)	0.47 (0.38,0.91)	0.51 (0.32,0.68)	0.48 (0.29,0.67)	0.45 (0.28,0.70)	0.61 (0.33,0.60)	0.34 (0.10,0.42)	0.60 (0.30,0.55)	0.43 (0.17,0.45)	0.62 (0.45,0.80)
VPA	0.61 (0.43,0.81)	0.33 (0.13,0.54)	0.43 (0.24,0.64)	0.29 (0.09,0.52)	0.32 (0.12,0.57)	0.26 (0.05,0.48)	0.18 (-0.03,0.41)	0.28 (0.08,0.52)	0.38 (0.19,0.60)	0.45 (0.27,0.68)	0.48 (0.29,0.69)
MVPA	0.65 (0.44,0.83)	0.35 (0.18,0.69)	0.49 (0.40,0.89)	0.52 (0.31,0.66)	0.52 (0.35,0.74)	0.54 (0.38,0.77)	0.64 (0.35,0.60)	0.38 (0.15,0.49)	0.64 (0.34,0.57)	0.49 (0.24,0.54)	0.65 (0.54,0.91)

NA: Not applicable

PA: physical activity

MWPA: moderate and walking activities

VPA: vigorous physical activity

MVPA: moderate-to-vigorous physical activity

Bold type: statistically significant correlations

CI: confidence interval

household activities,²¹ second, questionnaire structure problems (ex.: since IPAQ is conformed by four different domains, it could be difficult for respondents to understand, differentiate and report minutes per week of each domain),²⁰ and third, there is good evidence that women truly perform more household activities than in any other domain.²²

Mean differences were observed in almost all minutes per week for transportation, household, MWPA and MVPA between IPAQs. Similar results have been observed in previous repeatability studies worldwide^{23,24} and in Mexico.⁷ These differences could be related to the fact that teachers accumulated different minutes of physical activity throughout the year, over or underestimation of minutes per week of physical activity and the recognition of activities performed throughout the year.^{7,23,24}

The correlation between accelerometers 17 months apart was good for MVPA ($r=0.65$) and was slightly lower for VPA than for MWPA ($r= 0.61$ vs. $r= 0.60$, respectively). Some studies have found different correlation values between accelerometers (r ranged from 0.47 to 0.96).²⁵⁻²⁸ The main differences in the correla-

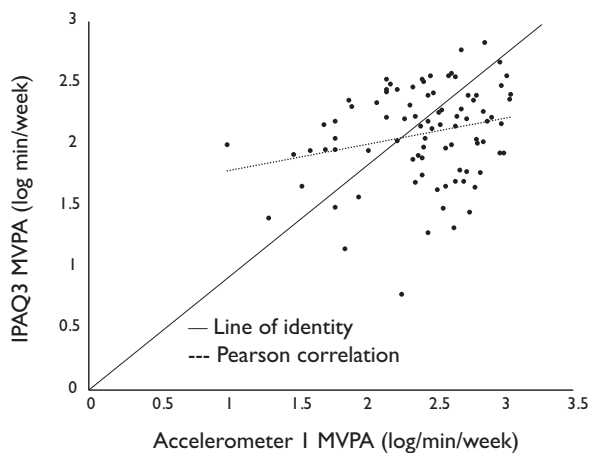
tion values could be a shorter interval range between measurements,²⁵ the use of laboratory settings and controlled activities,²⁵ seasonal physical activity variability,²⁷ differences between weekdays and weekends,²⁸ intra-individual variance,²⁸ types of accelerometer,²⁵⁻²⁹ changes in physical activity participation,³⁰ and accelerometer hip position (left or right).²⁹

The test-retest reliability for IPAQ long-form MVPA was lower compared to previous studies.^{6,31-34} Some explanations of the lower MVPA correlations could be related to the fact that most of the studies that reported higher correlation values had a shorter time interval between surveys.^{6,31-35} Lower correlation values have been observed for VPA than for MWPA.^{31,32} This could be attributed to the low frequency of VPA. In addition, in almost all IPAQ test-retests, lower repeatability correlation values were observed for transportation physical activity (ranging from 0.20 to 0.51) compared to other domains. In contrast, higher correlation values were observed for the leisure-time domain (ranging from 0.45 to 0.64). These results could be explained by the fact that transportation activities could vary from day to day and leisure-time activities are the most regular physical activities.^{31,32}

Table III
PEARSON AND INTRACLASS CORRELATIONS
BETWEEN LOG-TRANSFORMED MINUTES PER WEEK
ACCELEROMETERS AND LOG-TRANSFORMED MINUTES
PER WEEK IPAQ LONG-FORM. MEXICO 2013-15

	Variables (r (95%CI))	
	Pearson	ICC
Average IPAQ vs.AC1 (n= 82)		
MWPA	0.17 (-0.06, 0.44)	0.17 (-0.06, 0.26)
VPA	0.31 (0.11, 0.47)	0.42 (0.06, 0.63)
MVPA	0.25 (0.04, 0.55)	0.25 (-0.13, 0.51)
IPAQ3 vs.AC1 (n= 87)		
MWPA	0.16 (-0.07, 0.42)	0.21 (-0.12, 0.45)
VPA	0.44 (0.33, 0.83)	0.60 (0.39, 0.74)
MVPA	0.24 (0.04, 0.50)	0.31 (-0.04, 0.54)

ICC: intra-class correlation
 IPAQ: International Physical Activity Questionnaire – long form
 AC: accelerometer; MWPA: moderate and walking activities
 VPA: vigorous physical activity; MVPA: moderate-to-vigorous physical activity
 Bold type: statistically significant correlations (p<0.05)
 CI: confidence interval
 MVPA: moderate-to-vigorous physical activity
 MWPA: moderate and walking activities



MVPA: moderate-to-vigorous physical activity
FIGURE 2. LOG-TRANSFORMED MINUTES PER WEEK
OF MVPA BETWEEN IPAQ3 AND ACCELEROMETER
I. MEXICO 2013-2015

Minutes per week of MVPA and VPA for average IPAQ vs. AC1 and IPAQ3 vs. AC1 were significantly correlated. However, minutes per week of MWPA were not correlated. Similar results have been observed in

previous studies.³⁶ This could be explained by the fact that vigorous physical activities are easier to remember and report than moderate physical activities.³⁶ In addition, underreport of minutes of moderate physical activity.^{34,36}

The correlation of the average minutes MVPA between the IPAQ long-form vs. AC1 was $r=0.25$. Compared to other self-report instruments, slightly higher correlation values have been observed in different settings and countries for IPAQ long-form ranged from 0.30 to 0.38.^{6,19,32,35} One explanation for the low correlation observed in this type of analysis for all intensities could be the fact that subjects wore an Actical accelerometer one week, and this week does not represent the physical activity performed during almost 12 months (IPAQ assessment). The difficulty for participants to estimate the physical activity intensity in which they are engaged in,³⁵ and social desirability bias might also produce an overestimation of physical activity levels.³⁵ However, a slightly low correlation was observed between AC1 and IPAQ3 answered one week apart. This result is similar to what has been observed in previous studies using IPAQ-long form ($r=0.31$ Mexico vs. $r=0.33$ global)⁶ and similar compared to IPAQ-short form ($r=0.31$ long form vs. $r=0.31$ short form).⁷

The prevalence of physical inactivity was higher based on average AC (56.9%) than for average IPAQs (12.1%). This result is similar to previous studies in Mexican population.⁷ Caution should be taken when using self-report physical activity questionnaires, because misleading prevalence could conduct to underreported associations (physical activity vs. determinants, and vs. health outcomes).⁷

Strengths and limitations

This is one of the few studies that validated the IPAQ long-form in a Mexican sample. Although participants used the Actical accelerometer day and night during seven days, we asked them to remove their device every time they had contact with water (ex.: swimming, showers). This could result in an underestimation of the physical activity levels. It is well known that accelerometers and questionnaires measure different constructs of physical activity, accelerometers capture body movements, whereas questionnaires measures behaviors, making it difficult to compare.^{4,5} The 10-minute cut off point of the IPAQ could lead to overestimation, in other words, respondents could have rounded the minutes of physical activity less than 10 minutes. In addition, self-report questionnaires based on seven previous days are limited to last week, this might not represent necessarily the usual physical activity. We could not test the validity

of domain-specific physical activity because accelerometers could not differentiate between these domains. In addition, it is important to remind that this study was limited to female Mexican teachers from Mexico City, which increases its internal validity, but results may not represent women with other professions, residents from other cities, nor men.

Conclusion

Test-retest reliability of the IPAQ long-form was modest for MVPA and the validity against accelerometer was fair for MVPA and VPA in a sub sample of Mexican female teachers. Although these correlation values were similar to those reported previously, future studies should consider the use of objective measures to estimate physical activity levels.

Funding

This work was supported by the American Institute for Cancer Research grant number 10A035. CM received support from the CONACYT scholarship (CVU 359065).

Acknowledgments

To Antonio García for database management. In addition, Victor Sastré, director of *Promoción de la Función por Incentivos* of the *Secretaría de Educación Pública*, the State Coordinators and ISSSTE's (*Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado*) Medical directorate staff for their support in contacting the MTC's participants and assisting with logistics operations during the validation study. To the Mexican Teachers' Cohort participants and to Elsa Yunes-Díaz, Erin Comerford, Jaime Jaquez, Jose Isaias Rangel Rosas, Maricruz Morales, Martha Tamez, Nataly Mercado, and Rocío Santoyo Vistrain for their participation at field work operations.

Declaration of conflict of interests. The authors declare that they have no conflict of interests.

References

1. Lee IM, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT, et al. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *Lancet*. 2012;380(9838):219-29. [https://doi.org/10.1016/S0140-6736\(12\)61031-9](https://doi.org/10.1016/S0140-6736(12)61031-9)
2. Stern MC, Barnoya J, Elder JP, Gallegos-Carrillo K. Diet, physical activity, obesity and related cancer risk: strategies to reduce cancer burden in the Americas. *Salud Publica Mex*. 2019;61(4):448-55. <https://doi.org/10.21149/9753>

3. Shephard RJ. Limits to the measurement of habitual physical activity by questionnaires. *Br J Sports Med*. 2003;37(3):197-206. <https://doi.org/10.1136/bjism.37.3.197>
4. Troiano RP, McClain JJ, Brychta RJ, Chen KY. Evolution of accelerometer methods for physical activity research. *Br J Sports Med*. 2014;48(13):119-23. <https://doi.org/10.1136/bjsports-2014-093546>
5. Chen KY, Bassett DR. The technology of accelerometry-based activity monitors: current and future. *Med Sci Sports Exerc*. 2005;37(11 Suppl):S490-500.
6. Craig CL, Marshall AL, Sjoström M, Bauman AE, Booth M, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc*. 2003;35(8):1381-95. <https://doi.org/10.1249/01.MSS.0000078924.61453.FB>
7. Medina C, Barquera S, Janssen I. Validity and reliability of the International Physical Activity Questionnaire among adults in Mexico. *Rev Panam Salud Publica*. 2013;34(1):21-8.
8. Pereira MA, FitzGerald SJ, Gregg EW, Joswiak ML, Ryan WJ, Suminski RR, et al. A collection of Physical Activity Questionnaires for health-related research. *Med Sci Sports Exerc*. 1997;29(6 Suppl):S1-205.
9. van Poppel MN, Chinapaw MJ, Mokkink LB, van Mechelen W, Terwee CB. Physical activity questionnaires for adults: a systematic review of measurement properties. *Sports Med*. 2010;40(7):565-600. <https://doi.org/10.2165/11531930-000000000-00000>
10. Instituto Nacional de Salud Pública. EsMaestras [Internet]. Cuernavaca: ISNP; 2020 [cited July 2021]. Available from: <https://www.insp.mx/centros/salud-poblacional/esmaestras-521.html>
11. Lajous M, Ortiz-Panoso E, Monge A, Santoyo-Vistrain R, García-Anaya A, Yunes-Díaz E, et al. Cohort Profile: The Mexican Teachers' Cohort (MTC). *Int J Epidemiol*. 2017;46(2):e10. <https://doi.org/10.1093/ije/dyv123>
12. IPAQ-group. International Physical Activity Questionnaire. Downloadable questionnaires [Internet]. IPAQ Web site [cited October 20, 2020]. Available from: https://sites.google.com/site/theipaq/questionnaire_links
13. IPAQ-group. Guidelines for data processing and analysis of the International Physical Activity Questionnaire (IPAQ)-Short and Long Forms [Internet]. IPAQ Web site, 2005 [cited July 2021]. Available from: <https://sites.google.com/site/theipaq/>
14. Wong SL, Colley R, Connor-Gorber S, Tremblay M. Actical accelerometer sedentary activity thresholds for adults. *J Phys Act Health*. 2011;8(4):587-91.
15. Colley R, Connor-Gorber S, Tremblay MS. Quality control and data reduction procedures for accelerometry-derived measures of physical activity. *Health Rep*. 2010;21(1):63-6.
16. Colley RC, Garriguet D, Janssen I, Craig CL, Clarke J, Tremblay MS. Physical activity of Canadian children and youth: accelerometer results from the 2007 to 2009 Canadian Health Measures Survey. *Health Rep*. 2011;22(1):15-23.
17. World Health Organization. Guidelines on physical activity and sedentary behaviour [Internet]. Geneva:WHO; 2020 [cited January 2021]. Available from: <https://www.who.int/publications/i/item/9789240015128>
18. World Health Organization. Obesity: preventing and managing the global epidemic. WHO Technical Report Series No. 894. 2000 [Internet]. Geneva:WHO [cited July 2021]. Available from: <https://apps.who.int/iris/handle/10665/42330?locale-attribute=es&>
19. Boon RM, Hamlin MJ, Steel GD, Ross JJ. Validation of the New Zealand Physical Activity Questionnaire (NZPAQ-LF) and the International Physical Activity Questionnaire (IPAQ-LF) with accelerometry. *Br J Sports Med*. 2010;44(10):741-46. <https://doi.org/10.1136/bjism.2008.052167>
20. Hallal PC, Gomez LF, Parra DC, Lobelo F, Mosquera J, Florindo AA, et al. Lessons learned after 10 years of IPAQ use in Brazil and Colombia. *J Phys Act Health*. 2010;(Suppl2):S259-64. <https://doi.org/10.1123/jpah.7.s2.s259>
21. Sebastiao E, Gobbi S, Chodzko-Zajko W, Schwingel A, Papini CB, Nakamura PM, et al. The International Physical Activity Questionnaire-long form overestimates self-reported physical activity of Brazilian adults. *Public Health*. 2012;126(11):967-75. <https://doi.org/10.1016/j.puhe.2012.07.004>

22. Jauregui A, Salvo D, Medina C, Barquera S, Hammond D. Understanding the contribution of public- and restricted-access places to overall and domain-specific physical activity among Mexican adults: A cross-sectional study. *PLoS One*. 2020;15(2):e0228491. <https://doi.org/10.1371/journal.pone.0228491>
23. Melanson EL, Freedson PS. Physical activity assessment: a review of methods. *Crit Rev Food Sci Nutr*. 1996;36(5):385-396. <https://doi.org/10.1080/10408399609527732>
24. Hallal PC, Victora CG. Reliability and validity of the International Physical Activity Questionnaire (IPAQ). *Med Sci Sports Exerc*. 2004;36(3):556. <https://doi.org/10.1249/01.mss.0000117161.66394.07>
25. Sallis JF, Buono MJ, Roby JJ, Carlson D, Nelson JA. The Caltrac accelerometer as a physical activity monitor for school-age children. *Med Sci Sports Exerc*. 1990;22(5):698-703.
26. Mahoney M, Freedson P. Assessment of physical activity from caltrac and Baecke questionnaire technique. *Med Sci Sports Exerc*. 1990;22(S80).
27. Washburn RA, Cook TC, LaPorte RE. The objective assessment of physical activity in an occupationally active group. *J Sports Med Phys Fitness*. 1989;29(3):279-84.
28. Gretebeck R, Montoye, H. Variability of some objective measures of physical activity. *Med Sci Sports Exerc*. 1992;14:1167-72.
29. Aadland E, Ylvisaker E. Reliability of the Actigraph GT3X+ Accelerometer in Adults under Free-Living Conditions. *PLoS One*. 2015;10(8):e0134606. <https://doi.org/10.1371/journal.pone.0134606>
30. McClain JJ, Sisson SB, Tudor-Locke C. Actigraph accelerometer interinstrument reliability during free-living in adults. *Med Sci Sports Exerc*. 2007;39(9):1509-14. <https://doi.org/10.1249/mss.0b013e3180dc9954>
31. YanJun R, Meng S, QingMin L, YaYun T, YuKun D, LiMing L, et al. Validation of the Simplified Chinese-character Version of the International Physical Activity Questionnaire-Long Form in Urban Community-dwelling Adults: a cross-sectional study in Hangzhou, China. *Biomed Environ Sci*. 2017;30(4):255-63. <https://doi.org/10.3967/bes2017.035>
32. Macfarlane D, Chan A, Cerin E. Examining the validity and reliability of the Chinese version of the International Physical Activity Questionnaire, long form (IPAQ-LC). *Public Health Nutr*. 2011;14(3):443-50. <https://doi.org/10.1017/S1368980010002806>
33. Oyeyemi AL, Bello UM, Philemon ST, Aliyu HN, Majidadi RW, Oyeyemi AY. Examining the reliability and validity of a modified version of the International Physical Activity Questionnaire, long form (IPAQ-LF) in Nigeria: a cross-sectional study. *BMJ Open*. 2014;4(12):e005820. <https://doi.org/10.1136/bmjopen-2014-005820>
34. Sallis JF, Saelens BE. Assessment of physical activity by self-report: status, limitations, and future directions. *Res Q Exerc Sport*. 2000;71(2 Suppl):S1-14.
35. Vandelanotte C, De Bourdeaudhuij I, Philippaerts R, Sjöström M, Sallis J. Reliability and validity of a computerized and dutch version of the International Physical Activity Questionnaire (IPAQ). *J Phys Act Health*. 2005;2:63-75. <https://doi.org/10.1123/jpah.2.1.63>
36. Hagstromer M, Oja P, Sjostrom M. The International Physical Activity Questionnaire (IPAQ): a study of concurrent and construct validity. *Public Health Nutr*. 2006;9(6):755-62. <https://doi.org/10.1079/phn2005898>