

# Reliability and validity of two multidimensional self-reported physical activity questionnaires in people with chronic low back pain

Citation for published version (APA):

Carvalho, F. A., Morelhao, P. K., Franco, M. R., Maher, C. G., Smeets, R. J. E. M., Oliveira, C. B., Freitas Junior, I. F., & Pinto, R. Z. (2017). Reliability and validity of two multidimensional self-reported physical activity questionnaires in people with chronic low back pain. *Musculoskeletal science and practice*, 27, 65-70. <https://doi.org/10.1016/j.msksp.2016.12.014>

## Document status and date:

Published: 01/02/2017

## DOI:

[10.1016/j.msksp.2016.12.014](https://doi.org/10.1016/j.msksp.2016.12.014)

## Document Version:

Publisher's PDF, also known as Version of record

## Document license:

Taverne

## Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

## Take down policy

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.



## Original article

# Reliability and validity of two multidimensional self-reported physical activity questionnaires in people with chronic low back pain



Flávia A. Carvalho<sup>a</sup>, Priscila K. Morelhão<sup>a</sup>, Marcia R. Franco<sup>a</sup>, Chris G. Maher<sup>b</sup>,  
Rob J.E.M. Smeets<sup>c</sup>, Crystian B. Oliveira<sup>a</sup>, Ismael F. Freitas Júnior<sup>a</sup>, Rafael Z. Pinto<sup>a, b, \*</sup>

<sup>a</sup> Departamento de Fisioterapia, Faculdade de Ciências e Tecnologia, UNESP - Univ Estadual Paulista, Presidente Prudente, Brazil

<sup>b</sup> Musculoskeletal Division, The George Institute for Global Health, Sydney Medical School, University of Sydney, Sydney, Australia

<sup>c</sup> Department of Rehabilitation Medicine, Research School Caphri, Maastricht University and Libra Rehabilitation and Audiology, Eindhoven/Weert, The Netherlands

## ARTICLE INFO

## Article history:

Received 6 June 2016

Received in revised form

30 August 2016

Accepted 9 December 2016

## Keywords:

Chronic low back pain

Reliability

Validity

Self-report questionnaires

Accelerometry

## ABSTRACT

**Background:** Although there is some evidence for reliability and validity of self-report physical activity (PA) questionnaires in the general adult population, it is unclear whether we can assume similar measurement properties in people with chronic low back pain (LBP).

**Objective:** To determine the test-retest reliability of the International Physical Activity Questionnaire (IPAQ) long-version and the Baecke Physical Activity Questionnaire (BPAQ) and their criterion-related validity against data derived from accelerometers in patients with chronic LBP.

**Design:** Cross-sectional study.

**Methods:** Patients with non-specific chronic LBP were recruited. Each participant attended the clinic twice (one week interval) and completed self-report PA. Accelerometer measures >7 days included time spent in moderate-and-vigorous physical activity, steps/day, counts/minute, and vector magnitude counts/minute. Intraclass Correlation Coefficients (ICC) and Bland and Altman method were used to determine reliability and spearman rho correlation were used for criterion-related validity.

**Results:** A total of 73 patients were included in our analyses. The reliability analyses revealed that the BPAQ and its subscales have moderate to excellent reliability (ICC<sub>2,1</sub>: 0.61 to 0.81), whereas IPAQ and most IPAQ domains (except walking) showed poor reliability (ICC<sub>2,1</sub>: 0.20 to 0.40). The Bland and Altman method revealed larger discrepancies for the IPAQ. For the validity analysis, questionnaire and accelerometer measures showed at best fair correlation (rho < 0.37).

**Conclusions:** Although the BPAQ showed better reliability than the IPAQ long-version, both questionnaires did not demonstrate acceptable validity against accelerometer data. These findings suggest that questionnaire and accelerometer PA measures should not be used interchangeably in this population.

© 2016 Elsevier Ltd. All rights reserved.

## 1. Introduction

Measuring physical activity (PA) levels in chronic low back pain (LBP) is thought to have important clinical implications. Regarding clinical presentation, patients with chronic symptoms and high disability are more likely to present with low PA levels (Lin et al., 2011). From a prognosis perspective, patients considered to be physically active are more likely to have less pain and disability 1 year later (Pinto et al., 2014). Although it is still debatable whether

patients with chronic LBP decrease their PA level (Smeets et al., 2006), physical activity-based interventions are effective in managing this condition (van Middelkoop et al., 2010). More recently, in light of the evidence suggesting coexistence of chronic LBP and non-communicable diseases such as cardiovascular disease (Ha et al., 2014), diabetes (Eivazi and Abadi, 2012) and obesity (Shiri et al., 2010), clinicians are facing the challenge to concentrate efforts on interventions that not only decrease pain and disability but also improve patients' PA levels. This aligns with the Exercise is Medicine initiative, an initiative coordinated by the American College of Sports Medicine aiming to advance the implementation of evidence-based strategies to elevate the PA status in primary care settings (Sallis, 2009).

\* Corresponding author. Rua Roberto Simonsen, 305, FCT/UNESP, Campus de Presidente Prudente, Presidente Prudente 19060-900, SP, Brazil.

E-mail address: [rafaelzambelli@gmail.com](mailto:rafaelzambelli@gmail.com) (R.Z. Pinto).

Most evidence regarding PA in this area has been generated from self-reported assessment methods, such as the International Physical Activity Questionnaire (IPAQ) (Craig et al., 2003) and Baecke Habitual Physical Activity Questionnaire (BPAQ) (Baecke et al., 1982). While self-reported methods rely upon recall of physical activities, objective methods utilize technology to measure PA in real time. As technology advances, objective PA assessment, such as accelerometers, have become one of the most commonly used methods to assess PA in free-living activities, due to their small size, low participant burden and relatively low cost (Trost and O'Neil, 2014).

Although the IPAQ and BPAQ have some evidence for reliability (Baecke et al., 1982; Craig et al., 2003) and validity (Craig et al., 2003; Philippaerts et al., 2001) in the general adult population, it is unclear whether we can assume similar measurement properties in people with chronic LBP. This is important, particularly in the LBP field, in which evidence generated by self-reported and objective PA measures are often combined indiscriminately (Hendrick et al., 2011; Lin et al., 2011). Therefore, our primary aim was to determine the test-retest reliability of two multidimensional self-reported PA measures, the IPAQ (long-version) and BPAQ, and to assess the criterion-related validity of these two self-reported measures against objective PA measures derived from accelerometers in patients with chronic LBP. As a secondary aim we investigated the correlation between self-reported and objective PA measures with measures of pain and disability.

## 2. Materials and methods

### 2.1. Study design

This is a cross-sectional study conducted at an outpatient physical therapy university clinic.

### 2.2. Study population

Consecutive patients with LBP seeking physical therapy were recruited through advertisements in the community, local press and social media. This study was approved by the university ethics research committee (CAAE36332514.0.0000.5402) and all included patients gave informed consent. Patients were included if aged between 18 and 60 years old and presenting with chronic non-specific LBP, defined as pain localised below the costal margin and above the inferior gluteal folds, with or without leg pain of at least 3 months' duration. To be eligible to enter the study patients had to report moderate intensity low back pain and interference with function as measured by items 7 or 8 of the SF36 questionnaire (Ciconelli et al., 1999). Those patients presenting with known or suspected serious pathology such as nerve root compromise (at least 2 of the following signs: weakness, reflex change, or sensation loss, associated with the same spinal nerve); history of spinal surgery; or any contraindication to physical exercise were excluded.

Based on an expected ICC of 0.7, two measures per participant on two separate days and anticipating at least moderate reliability coefficient (ICC = 0.5), the sample size required was 63 participants. Nevertheless, given the high rates of noncompliance with accelerometer wearing protocol, the sample size was adjusted to allow for up to 35% dropout. Hence, the final sample size was set at 85 patients.

### 2.3. Procedures

Each participant attended the physical therapy clinic twice, approximately one week apart. At the first session information about demographic and anthropometric data; duration and

severity of LBP; disability; fear of movement; depression; and self-report PA measures were collected. Participants were also asked to wear an accelerometer during waking hours (except when showering or swimming) for the 7 days following the first session, while maintaining their typical weekly schedule. At a pre-determined time each morning participants received daily short message service (SMS) reminders to wear the accelerometer. At the second session participants returned the accelerometer and completed the self-reported PA measures for the second time.

### 2.4. Data collection

Patients completed a set of questionnaires to measure the following clinical constructs:

*Demographic and social variables* investigated were age, body mass index, highest education level, work status and duration of symptoms.

*Disability* was measured using The Quebec Back Pain Disability Scale (Kopec et al., 1995). The questionnaire consists of 20 items scored in a 0–5 scale. The total score ranges from 0 (no disability) to 100 (maximum disability).

*Average pain intensity* in the last 24 h was measured with a 0–10 numerical pain rating scale (Ross, 1997).

*Fear of movement* was measured using the 17-item version of The Tampa Scale for Kinesiophobia (Vlaeyen et al., 1995). Total score ranges from 17 to 68 with higher scores indicating greater fear of movement.

*Depression* was measured with The Beck Depression Inventory (BDI). The BDI is a 21-item, self-report rating inventory that measures characteristic attitudes and symptoms of depression (Beck et al., 1961). BDI scores range from 0 (low depression) to 63 (maximum depression).

*Self-reported PA behaviour* was measured with two questionnaires. The IPAQ long-version covers four PA domains: work-related, transportation, housework, and leisure activity (Craig et al., 2003). IPAQ asks in detail about walking, moderate-intensity and vigorous-intensity in each domain over the past seven days. Specific activity scores for each PA domain and intensity can be calculated by multiplying the number of minutes per week of the performed activities with the accompanying mean metabolic equivalent (MET) value of these activities. In this study, IPAQ measures calculated were: (i) IPAQ-Total PA (*MET-minutes/week*); (ii) PA for each domain: i.e. occupational, transportation, household, and leisure activity (*MET-minutes/week*); (iii) total time spent with walking (*minutes/week*); and (iv) total time spent in moderate-to-vigorous physical activity (MVPA) intensity (*minutes/week*) (i.e. computed by summing the *minutes/week* of reported MVPA across all domains). The BPAQ collects information on PA levels within a typical or usual week (Baecke et al., 1982). This questionnaire The BPAQ consists of 16 items and allows four indices to be calculated: work, sports, leisure time (excluding sports) and total PA (i.e. sum of all indices). The total score varies between 3 and 15 with higher score indicating a higher PA level.

*Objective PA behaviour* was evaluated by a triaxial accelerometer. The Actigraph wGT3X-BT (ActiGraph, LLC, Pensacola, FL, USA) is a non-invasive, small, lightweight device (4.6 × 3.3 × 1.5 cm, 19 g) that is worn during waking hours for 7 consecutive days on the right hip. For the purpose of this study, acceleration data were sampled at 30 Hz and analysed at 60 s epochs. Accelerometer data were analysed and computed with ActiLife 6 software (ActiGraph, LLC, Pensacola, FL, USA). For each patient, a complete data set was defined to have at least 10 h per day of monitored wear during at least 5 days (Troiano et al., 2008). Non-wear periods were defined as time intervals of at least 60 consecutive minutes of zero counts, with an activity interruption allowance of 0–100 counts min<sup>-1</sup>

lasting a maximum of 2 consecutive minutes. Outcome measures for each participant were: (i) total time spent in MVPA, (ii) total steps per day, (iii) counts per minute, and (iv) vector magnitude (VM) counts per minute. The MVPA, steps and counts were obtained from the vertical axis. MVPA was defined, according to (Troiano et al., 2008), as values greater than 2020 counts/min. Counts per minute were calculated by dividing the sum of activity counts of the vertical axis by the number of valid minutes. VM counts was obtained using data from all 3 axes and calculated by dividing the sum of activity counts of the 3 axes by the number of valid minutes.

2.5. Data analysis

Data are summarised as means with standard deviations (SDs) or medians with interquartile ranges (IQRs) depending on the data distribution. To allow comparison with previous studies, PA measures are reported as mean (SDs) and median (IQRs).

For the test-retest reliability analysis, Intraclass Correlation Coefficients (ICC<sub>2, 1</sub>) and its 95% confidence interval (CI) were calculated for all measures derived from the IPAQ and BPAQ. ICC values above 0.75 indicate excellent reliability, values between 0.75 and 0.40 indicate moderate reliability and values below 0.40 indicate poor reliability (Fleiss, 1986). The agreement between data collected at test 1 and test 2 were investigated using Bland and Altman method (Bland and Altman, 1999) and spearman rank (rho) correlations for the commonly reported self-reported PA outcomes (i.e. IPAQ-Total PA, IPAQ-MVPA and BPAQ-Total PA). Limits of agreement were defined as mean difference between test 1 minus test 2 ± (1.96 × SD of the difference).

For the validity analysis, we calculated rho correlation coefficients between self-reported PA measures and accelerometer data. We also calculated rho correlation coefficients of PA measures with pain and disability. The magnitude of association was interpreted as little or no relationship (from 0.00 to 0.25), fair relationship (from 0.25 to 0.50), moderate to good relationship (from 0.50 to 0.75), and good to excellent relationship (above 0.75) (Portney and Watkins, 2000). As the reliability of a measure can influence the magnitude of a correlation (i.e. unreliable measures may attenuate correlation coefficients), the correlation coefficients of self-reported PA measures with pain and disability were corrected using the following formula:  $r = \frac{r_{att}}{\sqrt{ICC_x ICC_y}}$ , where ICC is the reliability coefficients of measures X and Y,  $r_{att}$  is the attenuated or observable correlation and r is the error-free correlation (Fleiss, 1986). The reliability coefficients for pain (ICC<sub>(2,1)</sub> = 0.94) and disability (ICC<sub>(2,1)</sub> = 0.96) found in a similar population were used in the formula above (Costa et al., 2008; Rodrigues et al., 2009). All statistics were performed using IBM SPSS version 20.0 (IBM corporation, Somers, NY, USA).

3. Results

From October 2014 to September 2015, a total of 85 patients were recruited. Of these, 9 patients were excluded due to the lack of adherence to the accelerometer wearing protocol and 3 patients were excluded due to incomplete data in the self-reported questionnaires. Hence, data from 73 patients were included in our analyses. Characteristics of included patients are reported in Table 1.

Table 2 shows the reliability coefficients calculated for both IPAQ and BPAQ. The results for total PA revealed that BPAQ has excellent reliability, whereas IPAQ shows only poor reliability. For the PA domains, the reliability coefficients for the three BPAQ domains ranged from moderate to excellent. In the reliability analysis for the IPAQ domains, five domains showed poor reliability, and only two domains (i.e. domestic and walking) showed moderate reliability.

In the Bland and Altman analyses, IPAQ-Total PA and IPAQ-MVPA revealed significant difference between the two days (Fig. 1). The mean difference (limits of agreement) for IPAQ-Total PA and IPAQ-MVPA were 408.8 (−3027.6 to 3845.3) MET-minutes per week and 137.9 (−594.1 to 869.9) min/week, respectively. In both cases, it would appear that those patients who reported to be more physically active were the ones showing larger discrepancies in their responses (IPAQ-Total PA: rho = 0.32, p < 0.01; IPAQ-MVPA: rho = 0.31, p < 0.01). For the BPAQ-Total PA, the mean difference (limits of agreement) found was −0.05 (−1.78 to 1.81) with no observable systematic trends of correlation (rho = 0.73, p < 0.01).

For validity analysis, the correlation between self-reported and objective measures showed at best fair relationship (Table 3). IPAQ-total PA and BPAQ-total PA showed fair significant correlation with counts/min and steps/day (rho coefficients ranging from 0.25 to 0.33). The time spent in MVPA measured by the IPAQ and the accelerometer showed no significant correlation. In this case, the IPAQ clearly overestimates the total time spent in MVPA compared to data generated by the accelerometer. The correlation of subjective PA measures with clinical outcomes showed no relationship, even after assuming excellent reliability of PA measures. Fair correlation was found only between objective PA measures and disability.

4. Discussion

Our findings revealed that the BPAQ and its subscales have moderate to excellent reliability, whereas IPAQ and most IPAQ domains showed poor reliability. Bland and Altman analyses revealed large discrepancies for the IPAQ-Total PA e IPAQ-MVPA across the two test days. Not surprisingly, self-reported PA

Table 1 Demographics and clinical characteristics of 73 patients.

Characteristics	Total sample (n = 73)
Gender (n,%)	
Male	50 (68%)
Female	23 (32%)
Age (years)	37.2 (12.2)
BMI (Kg/m <sup>2</sup> )	29.9 (12.4)
Symptoms duration (months)	12 [30.0]
Education	
No education beyond primary school <sup>a</sup>	12 (17%)
No education beyond secondary school <sup>b</sup>	25 (34%)
Tertiary education incomplete	19 (26%)
Tertiary education complete	17 (23%)
Working status	
Yes	38 (52%)
No	35 (48%)
QBPD (0–100)	36.9 (20.6)
Pain (0–10)	6.9 (1.5)
BDI (0–63)	11.9 (6.9)
TSK (17–68)	44.4 (7.3)
Accelerometer data	
Counts/min	286.0 (100.0) 276.0 [114.3]
VM counts/min	608.0 (184.2) 577.1 [201.0]
MVPA min/day	21.7 (15.3) 17.1 [19.5]
Steps/day	8841.0 (2350.0) 6395.5 [2647.5]

Data are means (standard deviation), median [interquartile range] or frequency (proportion).

BMI, body mass index; BDI, Beck depression inventory; QBPD, Quebec back pain disability scale; TSK, Tampa scale for kinesiophobia; MVPA, moderate-to-vigorous physical activity.

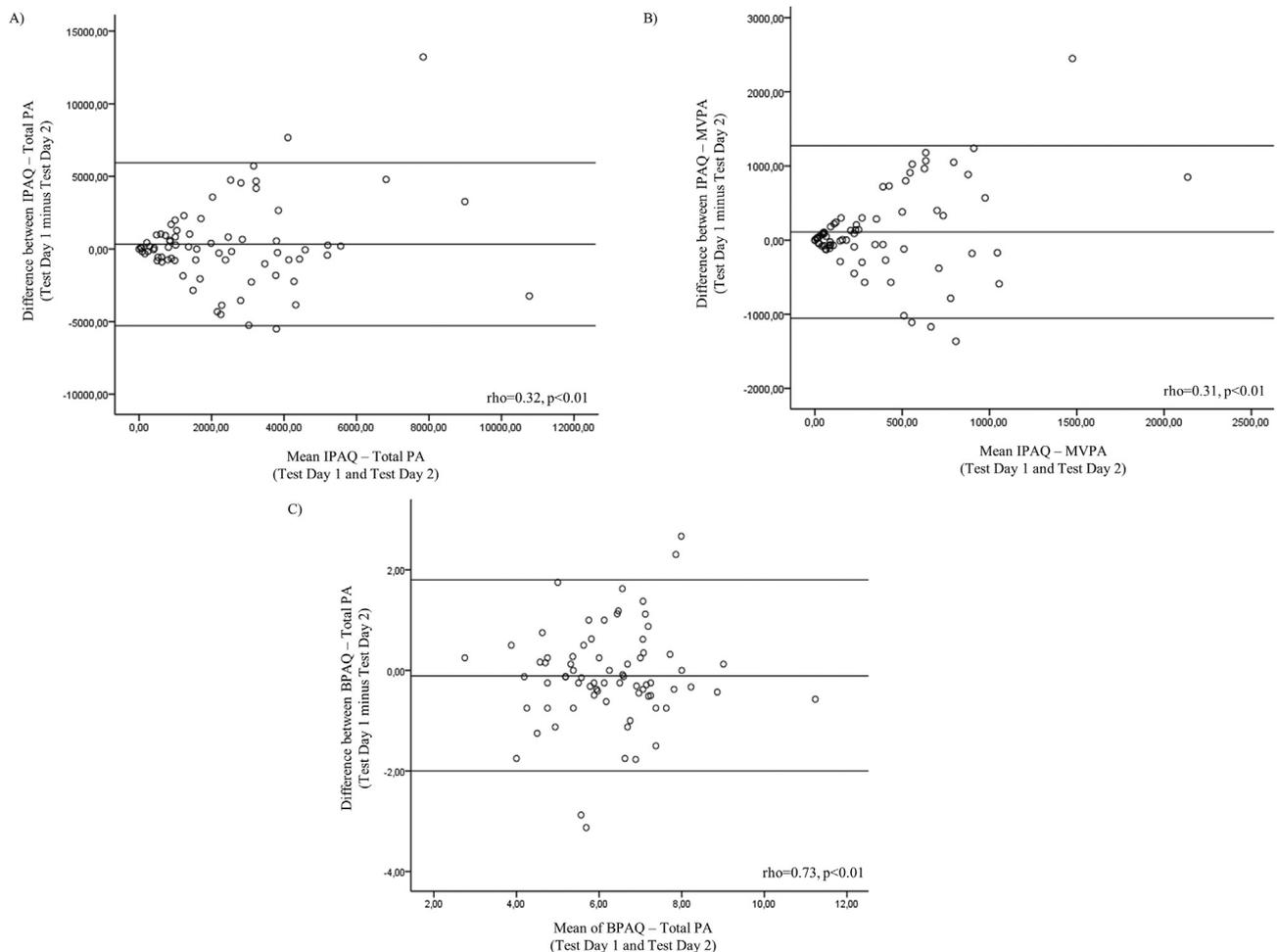
<sup>a</sup> In Brazil, primary schools provide education from the age of 5–11.

<sup>b</sup> In Brazil, secondary schools provide education from the age of 12–17.

**Table 2**  
Self-reported measures of physical activity and test-retest reliability (n = 73).

	Day 1		Day 2		ICC(2,1) (95%CI)
	Mean (SD)	Median [IQR]	Mean (SD)	Median [IQR]	
<b>IPAQ (METs-min/week)</b>					
Total PA (all domains)	2506.4 (2799.6)	1588.5 [3377.0]	2172.6 (2303.1)	1214.0 [3741.5]	0.37 (0.16–0.55)
Occupational	928.3 (2417.8)	0.0 [169.5]	582.6 (1510.6)	0.0 [34.7]	0.32 (0.11–0.51)
Transport	342.9 (520.4)	109.5 [396.0]	303.8 (461.5)	148.5 [420.8]	0.20 (–0.21 to 0.25)
Household	917.9 (1507.2)	240.0 [1141.3]	795.8 (1095.0)	360.0 [975.0]	0.40 (0.19–0.58)
Leisure	317.2 (654.4)	0.0 [468.5]	268.8 (766.2)	0.0 [192.0]	0.38 (0.17–0.56)
<b>IPAQ (min/week)</b>					
Walking	224.9 (401.9)	75.0 [270.0]	221.7 (410.5)	70.0 [280.0]	0.72 (0.58–0.81)
MVPA	431.2 (558.8)	180.0 [715.0]	321.8 (408.4)	140.0 [400.0]	0.25 (0.03–0.45)
<b>BPAQ</b>					
Total PA (all indices)	6.2 (1.5)	6.1 [1.8]	6.3 (1.4)	6.3 [1.9]	0.77 (0.65–0.84)
Work index	2.3 (0.9)	2.4 [1.3]	2.2 (0.8)	2.4 [1.1]	0.84 (0.75–0.89)
Sport index	1.8 (0.9)	1.5 [0.7]	1.9 (0.9)	1.8 [0.9]	0.83 (0.74–0.89)
Leisure index	2.1 (0.6)	2.0 [0.8]	2.2 (0.6)	2.0 [0.9]	0.61 (0.45–0.74)

IPAQ, International Physical Activity Questionnaire; BPAQ, Baecke Physical Activity Questionnaire; PA, physical activity; MVPA, moderate-to-vigorous physical activity; SD, Standard deviation; IQR, interquartile range; ICC, Intraclass Correlation Coefficient; 95%CI, 95% confidence interval.



**Fig. 1.** Bland and Altman plots for the comparison between data collected on two different days (one week apart). Measures are (A) IPAQ – Total PA, (B) IPAQ – MVPA and (C) BPAQ – Total PA. (BPAQ, Baecke Physical Activity Questionnaire; IPAQ, International Physical Activity Questionnaire; PA, Physical activity; MVPA, moderate and vigorous physical activity).

measures and the accelerometer data showed at best fair correlation. Taken together these findings challenge the validity of PA questionnaires when assessing PA level in this population and suggest caution in the interpretation of self-reported PA measures by researchers and clinicians. In addition, our data suggest that only

objective PA measures fairly correlate with clinical measures of disability but not pain in chronic non-specific LBP.

Structural differences between the two PA questionnaires might explain the discrepancies in the reliability results. While the IPAQ investigates the activities performed over the last seven days, the

**Table 3**  
Assessment of criterion-related validity.

	IPAQ - Total PA	IPAQ -MVPA	BPAQ -Total PA	Accelerometer data				Pain (0 -10)	Pain <sup>c</sup> (0 -10)	QBPD (0 -100)	QBPD <sup>c</sup> (0-100)
				Counts/min	VM counts/min	MVPAmin/day	Steps/day				
IPAQ - Total PA	–	0.75 <sup>b</sup>	0.46 <sup>b</sup>	0.33 <sup>b</sup>	0.33 <sup>b</sup>	0.22	0.37 <sup>b</sup>	–0.00	–0.00	–0.06	–0.10
IPAQ -MVPA		–	0.44 <sup>b</sup>	0.18	0.21	0.23	0.25 <sup>a</sup>	0.09	0.19	–0.11	–0.22
BPAQ -Total PA			–	0.18	0.26 <sup>a</sup>	0.17	0.27 <sup>a</sup>	–0.14	–0.16	–0.01	–0.01
Accelerometer data											
Counts/min				–	0.86 <sup>b</sup>	0.78 <sup>b</sup>	0.78 <sup>b</sup>	–0.18	–	–0.28 <sup>a</sup>	–
VM counts/min					–	0.58 <sup>b</sup>	0.81 <sup>b</sup>	–0.12	–	–0.19	–
MVPA min/day						–	0.71 <sup>b</sup>	–0.16	–	–0.27 <sup>a</sup>	–
Steps/day							–	–0.14	–	–0.27 <sup>a</sup>	–

IPAQ, International Physical Activity Questionnaire; BPAQ, Baecke Physical Activity Questionnaire; PA, physical activity; QBPD, Quebec back pain disability scale; MVPA, moderate-to-vigorous physical activity.

<sup>a</sup> Correlation is significant at the 0.05 level.

<sup>b</sup> Correlation is significant at the 0.01 level.

<sup>c</sup> Adjusted correlation for the effect of unreliable measures.

BPAQ collects information about activities performed during a typical week. The recall period, last week and typical week, used in each questionnaire might have led the respondents to answer the questions referring to different periods of time. Although a previous study with the IPAQ has shown that these recall periods (i.e. past seven days versus typical week) performed similarly in the general adult population (Craig et al., 2003), we found that this might not be the case in patients with chronic LBP. Given that patients with chronic LBP exhibit daily variation in PA (van Weering et al., 2009), it is possible that a similar variation can occur on a weekly basis. If this is the case, the fluctuations in PA might explain the poor reliability found with the IPAQ when assessing two consecutive weeks. Nevertheless, the influence of PA fluctuations are less evident when respondents are asked to recall PA levels from a typical week as in the BPAQ.

Sample characteristics such as education level and chronicity are factors that might have also affected the IPAQ reliability results. The IPAQ questions are somewhat challenging cognitively because respondents are asked to calculate time spent in different PA domains and PA intensities over the week. Given that our sample included a high proportion (51%) of patients with lower education levels, it is reasonable to assume that these patients might have had difficulty in estimating PA behaviour. Furthermore, there is evidence to suggest that people suffering from chronic pain have a deficit in working memory (i.e. memory function necessary for making decision, reasoning and planning) (Berryman et al., 2013), which, in turn, might contribute to the discrepancies between days found for the IPAQ in our test re-test assessments. The Bland and Altman plots for the IPAQ-Total PA and IPAQ-MVPA revealed a pattern in which more physically active participants reported greater disagreement between days. One possible explanation is that those who were more physically active started to pay more attention to their daily routine after the first assessment and reported a different PA behaviour on the second assessment.

The correlations coefficients between self-reported PA measures and the accelerometer data were at best fair. The magnitude of association between IPAQ and accelerometer data is somewhat similar to the ones reported in a study in healthy populations (Craig et al., 2003) and studies on specific disease, such as ankylosing spondylitis (Arends et al., 2013) and systemic lupus (Ahn et al., 2015). A previous study (van Weering et al., 2011) in a similar population, investigating whether the BPAQ correlates with accelerometer data (i.e. mean acceleration per minute) found no correlation. Our study using double the sample size showed significant and fair correlation of BPAQ-total PA with number of steps and vector magnitude. However, even when the correlations were

significant, they were lower than the recommended criteria ( $r \geq 0.50$ ) to assess validity of PA questionnaires (Terwee et al., 2010). Our findings suggest that the LBP field faces a similar disparity between self-report and objective PA assessment method reported in studies with general adult population (Prince et al., 2008).

Our findings showed that none of the PA measures were significantly associated with pain. For the association between PA measures and disability, we found fair negative correlation between objective PA measures and disability which is in accordance with the results of a previous systematic review (Lin et al., 2011). However, our results showed no relationship between self-reported PA measures and disability even after adjusting for the effect of unreliable measures. These findings suggest that systematic reviews should not combine self-reported and objective PA measures indiscriminately. We would advocate for the use of only one type of PA measure in systematic reviews (Oliveira et al., 2016).

The sample included in this study had an average pain intensity of 6.9 points (on a 0–10 pain scale), which might be considered higher compared to what is expected for a population of chronic LBP. We would argue that our sample comes from an outpatient physical therapy university clinic that provides care to those patients who normally cannot afford private health care in Brazil. While this may limit generalizability of our results, recent randomized clinical trials conducted with this population in similar settings in Brazil have shown similar values, with mean pain intensity ranging from 6.8 to 7.6 points (Added et al., 2016; Magalhaes et al., 2015; Parreira Pdo et al., 2014). Similarly high levels of fear of movement were found when comparing our sample with a sample from a previous study conducted with this population (Magalhaes et al., 2015). A potential limitation of the current study is that the hip-worn accelerometer may not accurately capture movements during activities such as cycling, carrying or lifting objects (Trost and O'Neil, 2014). Nevertheless, we would assume that the contribution of these activities to the overall physical activity is probably small in this population. Another limitation is that, although we instructed participants to maintain their typical weekly schedule, we did not monitor whether they followed this instruction. It is possible that some participants might have had a different weekly routine, which may interfere with PA levels measured in this study.

Our results revealed that, among the self-reported measures, the BPAQ showed better reliability than the IPAQ long-version in patients with chronic LBP. Both self-reported measures, however, did not demonstrate acceptable validity against accelerometer data.

These findings suggest that self-report and objective PA measures should not be used interchangeably in this population.

### Financial supporters

F. A. Carvalho is a honours student with a scholarship, grant number 2015/17093-7, from São Paulo Research Foundation (FAPESP), Brazil. M.R. Franco is a research fellow with a scholarship, grant number 2015/07704-9, from FAPESP, Brazil. P.K. Morelhão is a master student with a scholarship, grant number 2015/02744-2, from FAPESP, Brazil. R.Z. Pinto received a research grant, grant number 2014/14077-8, from FAPESP – Brazil. C. G. Maher holds a principal research fellowship funded by Australia's National Health and Medical Research Council.

### References

- Added, M.A., Costa, L.O., de Freitas, D.G., Fukuda, T.Y., Monteiro, R.L., Salomao, E.C., et al., 2016. Kinesio taping does not provide additional benefits in patients with chronic low back pain who receive exercise and manual therapy: a randomized controlled trial. *J. Orthop. Sports Phys. Ther.* 46, 506–513.
- Ahn, G.E., Chmiel, J.S., Dunlop, D.D., Helenowski, I.B., Semanik, P.A., Song, J., et al., 2015. Self-reported and objectively measured physical activity in adults with systemic lupus erythematosus. *Arthritis Care Res. Hob.* 67, 701–707.
- Arends, S., Hofman, M., Kamsma, Y.P., van der Veer, E., Houtman, P.M., Kallenberg, C.G., et al., 2013. Daily physical activity in ankylosing spondylitis: validity and reliability of the IPAQ and SQUASH and the relation with clinical assessments. *Arthritis Res. Ther.* 15, R99.
- Baecke, J.A., Burema, J., Frijters, J.E., 1982. A short questionnaire for the measurement of habitual physical activity in epidemiological studies. *Am. J. Clin. Nutr.* 36, 936–942.
- Beck, A.T., Ward, C.H., Mendelson, M., Mock, J., Erbaugh, J., 1961. An inventory for measuring depression. *Arch. Gen. Psychiatry* 4, 561–571.
- Berryman, C., Stanton, T.R., Jane Bowering, K., Tabor, A., McFarlane, A., Lorimer Moseley, G., 2013. Evidence for working memory deficits in chronic pain: a systematic review and meta-analysis. *Pain* 154, 1181–1196.
- Bland, J.M., Altman, D.G., 1999. Measuring agreement in method comparison studies. *Stat. Methods Med. Res.* 8, 135–160.
- Ciconelli, R.M., Ferraz, M.B., Santos, W., Meinão, I., Quaresma, M.R., 1999. Brazilian-Portuguese version of the SF-36. A reliable and valid quality of life outcome measure. *Rev. Bras. Reumatol.* 39, 143–150.
- Costa, L.O., Maher, C.G., Latimer, J., Ferreira, P.H., Ferreira, M.L., Pozzi, G.C., et al., 2008. Clinimetric testing of three self-report outcome measures for low back pain patients in Brazil: which one is the best? *Spine (Phila Pa 1976)* 33, 2459–2463.
- Craig, C.L., Marshall, A.L., Sjostrom, M., Bauman, A.E., Booth, M.L., Ainsworth, B.E., et al., 2003. International physical activity questionnaire: 12-country reliability and validity. *Med. Sci. Sports Exerc.* 35, 1381–1395.
- Eivazi, M., Abadi, L., 2012. Low back pain in diabetes mellitus and importance of preventive approach. *Health Promot Perspect.* 2, 80–88.
- Fleiss, J.L., 1986. *The Design and Analysis of Clinical Experiments*. Wiley, New York.
- Ha, I.H., Lee, J., Kim, M.R., Kim, H., Shin, J.S., 2014. The association between the history of cardiovascular diseases and chronic low back pain in South Koreans: a cross-sectional study. *PLoS One* 9, e93671.
- Hendrick, P., Milosavljevic, S., Hale, L., Hurlley, D.A., McDonough, S., Ryan, B., et al., 2011. The relationship between physical activity and low back pain outcomes: a systematic review of observational studies. *Eur. Spine J.* 20, 464–474.
- Kopec, J.A., Esdaile, J.M., Abrahamowicz, M., Abenhaim, L., Wood-Dauphinee, S., Lamping, D.L., et al., 1995. The Quebec back pain disability scale. Measurement properties. *Spine (Phila Pa 1976)* 20, 341–352.
- Lin, C.W., McAuley, J.H., Macedo, L., Barnett, D.C., Smeets, R.J., Verbunt, J.A., 2011. Relationship between physical activity and disability in low back pain: a systematic review and meta-analysis. *Pain* 152, 607–613.
- Magalhaes, M.O., Muzi, L.H., Comachio, J., Burke, T.N., Renovato Franca, F.J., Vidal Ramos, L.A., et al., 2015. The short-term effects of graded activity versus physiotherapy in patients with chronic low back pain: a randomized controlled trial. *Man. Ther.* 20, 603–609.
- Oliveira, C.B., Franco, M.R., Maher, C.G., Christine Lin, C.W., Morelhão, P.K., Araújo, A.C., Negrão Filho, R.F., Pinto, R.Z., 2016. Physical activity interventions for increasing objectively measured physical activity levels in chronic musculoskeletal pain: systematic review. *Arthritis Care Res (Hoboken)* 68 (12), 1832–1842.
- Parreira Pdo, C., Costa Lda, C., Takahashi, R., Hespanhol Junior, L.C., Luz Junior, M.A., Silva, T.M., et al., 2014. Kinesio taping to generate skin convolutions is not better than sham taping for people with chronic non-specific low back pain: a randomised trial. *J. Physiother.* 60, 90–96.
- Philippaerts, R.M., Westerterp, K.R., Lefevre, J., 2001. Comparison of two questionnaires with a tri-axial accelerometer to assess physical activity patterns. *Int. J. Sports Med.* 22, 34–39.
- Pinto, R.Z., Ferreira, P.H., Kongsted, A., Ferreira, M.L., Maher, C.G., Kent, P., 2014. Self-reported moderate-to-vigorous leisure time physical activity predicts less pain and disability over 12 months in chronic and persistent low back pain. *Eur. J. Pain* 18, 1190–1198.
- Portney, L.G., Watkins, M.P., 2000. *Foundations of Clinical Research: Applications to Practice*, second ed. Prentice-Hall, Upper Saddle River, New Jersey.
- Prince, S.A., Adamo, K.B., Hamel, M.E., Hardt, J., Connor Gorber, S., Tremblay, M., 2008. A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. *Int. J. Behav. Nutr. Phys. Act.* 5, 56.
- Rodrigues, M.F., Michel-Crosato, E., Cardoso, J.R., Traebert, J., 2009. Psychometric properties and cross-cultural adaptation of the Brazilian Quebec back pain disability scale questionnaire. *Spine (Phila Pa 1976)* 34, E459–E464.
- Ross, R.L.P., 1997. Clinical assessment of pain. In: van Dieen, J.H. (Ed.), *Assessment in Occupational Therapy and Physical Therapy*. WB Saunders, Philadelphia, pp. 123–133.
- Sallis, R.E., 2009. Exercise is medicine and physicians need to prescribe it! *Br. J. Sports Med.* 43, 3–4.
- Shiri, R., Karppinen, J., Leino-Arjas, P., Solovieva, S., Viikari-Juntura, E., 2010. The association between obesity and low back pain: a meta-analysis. *Am. J. Epidemiol.* 171, 135–154.
- Smeets, R.J., Wade, D., Hidding, A., Van Leeuwen, P.J., Vlaeyen, J.W., Knottnerus, J.A., 2006. The association of physical deconditioning and chronic low back pain: a hypothesis-oriented systematic review. *Disabil. Rehabil.* 28, 673–693.
- Terwee, C.B., Mokkink, L.B., van Poppel, M.N., Chinapaw, M.J., van Mechelen, W., de Vet, H.C., 2010. Qualitative attributes and measurement properties of physical activity questionnaires: a checklist. *Sports Med.* 40, 525–537.
- Troiano, R.P., Berrigan, D., Dodd, K.W., Masse, L.C., Tilert, T., McDowell, M., 2008. Physical activity in the United States measured by accelerometer. *Med. Sci. Sports Exerc* 40, 181–188.
- Trost, S.G., O'Neil, M., 2014. Clinical use of objective measures of physical activity. *Br. J. Sports Med.* 48, 178–181.
- van Middelkoop, M., Rubinstein, S.M., Verhagen, A.P., Ostelo, R.W., Koes, B.W., van Tulder, M.W., 2010. Exercise therapy for chronic nonspecific low-back pain. *Best. Pract. Res. Clin. Rheumatol.* 24, 193–204.
- van Weering, M.G., Vollenbroek-Hutten, M.M., Hermens, H.J., 2011. The relationship between objectively and subjectively measured activity levels in people with chronic low back pain. *Clin. Rehabil.* 25, 256–263.
- van Weering, M.G., Vollenbroek-Hutten, M.M., Tonis, T.M., Hermens, H.J., 2009. Daily physical activities in chronic lower back pain patients assessed with accelerometry. *Eur. J. Pain* 13, 649–654.
- Vlaeyen, J.W., Kole-Snijders, A.M., Boeren, R.G., van Eek, H., 1995. Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance. *Pain* 62, 363–372.