As accepted for publication in Journal of Aging and Physical Activity,

https://doi.org/10.1123/japa.2016-0207

Validation of the IPAQ against different accelerometer cut-points in older cancer survivors and adults at risk of cancer

Abstract

Introduction: The present study investigated the convergent validity of an interview-administered IPAQ long version (IPAQ-L) in an older population by comparison with objective accelerometry movement data.

Methods: Data from 52 participants (mean age 67.9 years, 62% male) were included in the analysis. Treadmill derived (TM-ACC: 1952-5724 cpm) and free-living physical activity (PA) derived (FL-ACC: 760-5724 cpm) accelerometer cut-points were used as criterion.

Results: IPAQ-L measures (total PA, leisure-time, walking-time, sedentary time) were significantly correlated with accelerometry ($P \leq 0.05$). Differences in sex were observed. Bland-Altman Limits of Agreement analysis showed that the IPAQ-L overestimated PA in relation to accelerometry.

Conclusion: Our results show that an interview-administered IPAQ-L shows low to moderate convergent validity with objective PA measures in this population but there may be differences between males and females which should be further investigated.

Keywords: physical activity, elderly, older adults, cancer, International Physical Activity Questionnaire, measurement

INTRODUCTION

The role of physical activity (PA) in maintaining health and vitality in older age has been well documented (Nelson et al., 2007). Despite this, PA levels show a decline with advancing age (Department of Health, 2011) and evidence for the long-term effectiveness of PA interventions in older people is lacking (Department of Health, 2011). However, valid PA measures are needed to assess the effectiveness of interventions targeted at this population.

The International Physical Activity Questionnaire long version (IPAQ-L) was developed to measure PA across ages and countries and to enable international comparisons (Craig et al., 2003). Acceptable validity (using accelerometry as criterion measure) has been reported for people aged 18-65 (Craig et al., 2003; Hagstromer, Oja, & Sjostrom, 2006; Macfarlane, Chan, & Cerin, 2011) but to the best of our knowledge only two studies have assessed the validity of the IPAQ-L against accelerometry in older populations and small to moderate correlation coefficients were reported (Cerin et al., 2012; Van Holle, De Bourdeaudhuij, Deforche, Van Cauwenberg, & Van Dyck, 2015). However, both of these studies compared the IPAQ-L (which measures PA across different lifestyle PA domains) to accelerometer cut-points that were calibrated during treadmill walking (Freedson, Melanson, & Sirard, 1998; Copeland & Esliger, 2009). One would expect these thresholds would have higher validity for walking than free-living activities. Accelerometer cut-points using free-living activities have been derived (Hendelman, Miller, Baggett, Debold, & Freedson, 2000; Matthews, 2005), but there is currently no consensus on the optimal cut-points for these activities or this population (Swartz et al., 2000, Copeland & Esliger, 2009; Miller, Strath, Swartz, & Cashin, 2010, Hall, Howe, Rana, Martin, & Morey, 2013).

The aim of this study was to examine the convergent validity of an interview-administered IPAQ-L in an elderly population by comparison with commonly used cut-points developed during treadmill walking and accelerometry cut-points derived from free-living activities. Treadmill-derived accelerometer cut-points for moderate to vigorous intensity PA were defined by Freedson (1,952-5,724 cpm; Freedson et al., 1998), and free-living PA accelerometer cut-points by Matthews (760-5724 cpm; Matthews, 2005). Furthermore, differences in convergent validity between males and females were investigated and findings reported for the individual IPAQ-L domains separately. In addition to assessing the impact of accelerometer cut-point adjustment, associations between self-reported PA domains and accelerometer-derived data for total accumulated PA and bouts of ≥10-min (consistent with current recommendations) (Pollock et al., 1998) were also investigated.

METHODS

Participants

Participants were patients recruited for pilot intervention studies and baseline data were available from 58 participants (mean age= 67.9 years, range 60-88) who were diagnosed with either bowel polyps or were recovering from curative bowel cancer treatment (Dukes stages A-C, within 3 years of completed treatment for cancer). As part of the main trials, participants were screened for a history of co-morbid conditions that might preclude them from safely undertaking exercise. Conditions included a recent myocardial infarction, uncontrolled hypertension, or unstable angina. We did not collect data on other co-morbidities. None of the participants were physically restricted in carrying out moderate-intensity PA. Informed consent was obtained prior to entering the study, which was approved by the NRES East of England Ethics Committee. More details of the original studies can be found elsewhere (https://clinicaltrials.gov/ct2/show/NCT02724306, https://clinicaltrials.gov/ct2/show/NCT02751892).

Physical Activity Assessments

Participants presented themselves at the University of East Anglia and were fitted with an accelerometer which they were instructed to wear during waking hours until their next appointment at least 7 days later. At this second appointment accelerometer data were downloaded onto a computer and the IPAQ-L was completed in an interview setting to capture self-reported PA over the past seven days (corresponding with accelerometer wear-time). Before the interviews, the interviewer clarified the time period of interest and explained the different PA domains that were captured (see below). The interviewer further explained that only PA of at least 10 min continuous duration is captured by the questionnaire. All interviews were conducted by the same interviewer. The meaning of moderate and vigorous intensity PA were demonstrated with the 15-item BORG scale (range 6-20) (Borg, 1982), which was presented as a visual aid during each question. A rating of 11-13 on the BORG scale was considered moderate intensity PA and ratings of ≥14 as vigorous intensity PA (Pollock et al., 1998). Once the interviewer was satisfied that the participant understood the concept of the IPAQ-L, the questions were read out loud. Each response was probed to ensure that reported activities met the requirements for intensity and duration and that the same activities were not reported repeatedly.

IPAQ-L Scoring

The IPAQ-L is a 27-item questionnaire which identifies duration (hours and minutes per day), frequency (times per week) and intensity (moderate and vigorous) of PA within four different domains (occupation, transportation, household/house maintenance, leisure). Sedentary behaviour is also captured. The IPAQ-L was scored according to original guidelines (The IPAQ Group). PA was reported in minutes per week and vigorous intensity PA was not included in the analysis because only five participants reported being engaged in this type of PA. For analysis, the different PA domains were condensed into the following categories: (i) total PA minutes per week as the sum of all PA, including moderate and walking PA (TOTAL-IPAQ); (ii) total moderate PA as the sum of all moderate PA excluding walking (MOD-IPAQ); (iii) total leisure time PA including walking for leisure (LEISURE-IPAQ); (iv) total walking PA as the sum of the 'transportation' and 'walking' domains (WALK-IPAQ); and (v) the sum of occupational and household/house maintenance activities (OH-IPAQ). Household/housework PA and occupational PA were merged because most participants were retired and thus, did not report occupational PA.

Accelerometry Data

Participants were fitted with a GT3X accelerometer (Actigraph, Pensacola, FL, USA), which was worn on the right hip. The device is a tri-axial accelerometer measuring accelerations in a vertical (y-axis), antero-posterior (x-axis), and medio-lateral plane (z-axis). The output also provides vector magnitude which is a composite measure of all three axes. The epoch period was set at 1 minute as used in previous calibration studies (Freedson et al., 1998; Hendelman et al., 2000; Miller et al., 2010), and spike tolerance was set to 2 minutes. Moderate intensity PA was analysed using two different cut-point thresholds, one of which was treadmill-derived (TM) (Freedson et al., 1998) and the other free-living derived (FL) (Matthews, 2005). Two different PA duration criteria were applied as follows: (i) total moderate intensity PA in continuous bouts of ≥10 min, using TM cut-points 1952-5724 cpm (Freedson et al., 1998) (TM-10MIN); (ii) total moderate intensity PA in continuous bouts of ≥10 min, using 760-5724 cpm (Matthews, 2005) (FL-10MIN) (iii) total accumulated moderate intensity PA, using 1952-5724 cpm (Freedson et al., 1998) (TM-ACC) and (iv) total accumulated moderate intensity PA, using 760-5724 cpm (Matthews, 2005) (FL-ACC). Time spent sitting was defined as <100 cpm (Matthews et al, 2008). Step counts (SC) per week were also recorded and used in the analysis. Only moderate intensity PA recordings are reported here because of a lack of vigorous intensity PA in the accelerometry data (only one participant had recordings above 5724 cpm).

Data Analysis

On return of the device, data were downloaded onto a computer and examined for valid wear-time of at least 10 h per day on a minimum of 5 days per week, including a weekend day (Choi, Liu, Matthews, & Buchowski, 2011). Data that did not meet these criteria were excluded from the analysis. Physical activity diaries, which were kept by participants during the accelerometer-wear-period, were investigated for participants' engaging in water activities (e.g. swimming). Nobody was identified as having engaged in water activities, and therefore, no participant was excluded from the analysis for this reason.

Data were analysed with the Statistical Package for the Social Sciences (SPSS) for Windows, version 22 (Armonk, NY: IBM Corp). The Shapiro-Wilk test showed that PA data were non-normally distributed, and non-parametric tests were used for the analysis. Differences in PA behaviour between males and females were tested with the Mann-Whitney-U test and correlation statistics were performed with the Spearman rank correlation. The correlation coefficient (ρ) was interpreted according to Hopkins (0-0.1 trivial, >0.1-0.3 small, >0.3 to 0.5 moderate, >0.5-0.7 large, >0.7-0.9 very large, and >0.9-1 nearly perfect) (Hopkins, 2002). Correlations were calculated with the Fisher's 'z' transformation and differences between Bland-Altman plots were used to assess the limits of agreement between the two methods (Martin Bland & Altman, 1986). Therefore, % difference between the two methods was plotted with the Bland-Altman method: values closer to zero suggest greater limits of agreement, whereas more dispersed values represent greater differences between IPAQ and accelerometer data.

Power Calculation

The sample size calculation was based on the correlation between two measures rather than the mean difference between males and females. With n = 30 (males) the study would have more than 80% power to detect a correlation between any two measurements of 0.5; and with n=20 (females) the study would have more than 80% power to detect a correlation between any two measurements.

Results

After exclusion of six participants for whom accelerometer wear-time was invalid, data from 52 participants were available for analysis. Of those 38% (n=20) and 62% (n=32) were females and males, respectively. Participants were on average 67.9 (range 60-80) years old and had a BMI of 28.7 kg/m^2 (standard deviation SD±4.7) (Table 1). Females were on average 4 years younger than males ($P<0.014$). There were no other significant differences between sexes. Physical activity levels from IPAQ and accelerometry are reported in Table 2.

Table 3 presents correlations between the different accelerometer cut-points and domains of the IPAQ for the overall sample population. Overall, the strongest correlations were observed between accelerometry and WALK-IPAQ (ρ = 0.34- 0.57, $P\leq0.01$), followed by moderate correlations with LEISURE-IPAQ (ρ = 0.30- 0.45, $P\leq0.01$) and TOTAL-IPAQ (ρ = 0.38-0.43, $P\leq0.01$) and small but non-significant correlations with MOD-IPAQ (ρ = 0.16- 0.27, $P\geq$ 0.05) and OH-IPAQ (ρ =-0.08- 0.27, $P\geq0.05$). Correlations between the TM-10MIN criterion and IPAQ variables were strongest for TOTAL-IPAQ (ρ =0.43, P≤0.01) and WALK-IPAQ (ρ =0.57, P≤0.001). Correlations between the FL-10MIN criterion and IPAQ data, we found strongest correlations for MOD-IPAQ (ρ =0.23, P≥0.05) and OH-IPAQ (ρ =0.25, P≥0.05) but these were not significant. Sedentary time for the two measurement methods was moderately correlated (ρ =0.33, P≤0.05). Correlations between accelerometer step count data and IPAQ measures were moderate and significant for WALK-IPAQ (ρ =0.34, P≤0.05) and LEISURE-IPAQ (ρ =0.33, P≤0.05) and small but not significant for TOTAL-IPAQ (ρ =0.27, P≥0.05), MOD-IPAQ (ρ =0.14, P≥0.05), and OH-IPAQ (ρ =0.12, P≥0.05). Finally, correlations between vector magnitude were moderate for WALK-IPAQ (ρ =0.34, P≥0.05) and LEISURE-IPAQ (ρ =0.32, P≥0.05).

In general, correlations were higher for females than for males (Table 4) and this was significant for several PA criteria. TOTAL-IPAQ correlations with TM-ACC (0.71 vs 0.24, P ≤0.05) and FL-ACC (0.71 vs 0.19, P≤0.05) were significantly stronger in females than in males. Furthermore, significant sex differences were observed for correlations between WALK-IPAQ and TM-ACC (0.84 vs 0.42, P≤0.05) and TM-10MIN (0.81 vs 0.40, P≤0.05), between MOD-IPAQ and LOW-ACC (0.61 vs 0.10, P≤0.05) and LOW-10M (0.58 vs 0.08, P≤0.05), and finally between OH-IPAQ and LOW-10M (0.6 vs 0.01, P≤0.05). It should also be noted, that all of IPAQ-L domains were significantly correlated with LOW-ACC and FL-10MIN in females, but not in males.

The agreement between the two methods is displayed as Bland-Altman plots (Figure 1) and the plots revealed a high level of heteroscedasticity. The plots present the percent difference between methods and show largest bias between the TOTAL-IPAQ and the TM-10MIN criterion, followed by TOTAL-IPAQ and TM-ACC. Differences between TOTAL-IPAQ and FL-ACC showed the lowest bias (23%). In summary, the IPAQ overestimated PA compared to all four accelerometer criteria, and this overestimation was largest for TM-10MIN and lowest for FL-ACC. The bias between IPAQ sedentary time and accelerometer derived sedentary time was 49.9%

Differences between methods were also explored to investigate whether treadmill derived cut-points (TM-ACC) were similar to walking from IPAQ, and free-living derived cut-points (FL-ACC) were similar to total PA from the IPAQ (walking + other activities). There were no significant differences between TOTAL-IPAQ and FL-ACC (P= 0.11), and between WALK-IPAQ and TM-ACC (P= 0.07). In contrast, significant differences were found between WALK-IPAQ and FL-ACC ($P \leq 0.05$), and between TOTAL-IPAQ and TM-ACC ($P \leq 0.05$).

Discussion

This study is novel in several ways and addresses the limitations of previous IPAQ-L validation studies in older people. First, the IPAQ-L was administered by interview to prevent misinterpretation of common PA terms such as 'duration', 'frequency', and 'intensity' in older populations and all interviews were carried out by the same interviewer, thus eliminating inter-rater bias. Second, cut-points which may be more appropriate for classifying free-living moderate intensity PA in older people were included in the analysis. Third, both total PA and total PA as continuous bouts of ≥10 min, deemed important for health benefits (Pollock et al., 1998), were compared between the two instruments.

In contrast with other IPAQ-L validation studies, this study demonstrated stronger correlations for WALK-PA and sedentary time against accelerometry-derived data, whereas total PA and time spent in moderate intensity PA were similar to previous findings (Craig et al., 2003; Hagstromer, Ainsworth, Oja, & Sjostrom, 2010; Macfarlane et al., 2011; Cerin et al., 2012; Van Holle et al., 2015). This indicates that an interview-administered IPAQ may more accurately capture the PA domains walking and sedentary time than the self-administered IPAQ. Analysing accelerometer data as continuous bouts of ≥10 min did not yield stronger correlations with self-reported PA. In the overall sample, the applied accelerometer criteria (FL-10MIN, TM-10MIN, FL-ACC, TM-ACC) were significantly correlated with TOTAL-IPAQ, WALK-IPAQ, LEISURE-IPAQ and sedentary time, but not with MOD-IPAQ or OH-IPAQ. Lack of correlation with the latter two variables may reflect limitations of accelerometry as an accurate measure of upper-body activities, such as gardening and household tasks that are recorded within the moderate PA and occupational/household domains of the IPAQ-L, respectively (Hendelman et al., 2000). Furthermore, accelerometers are generally unable to distinguish between different walking conditions, such as uphill walking or carrying heavy loads, and have been shown to underestimate activities such as cycling, and resistance exercise (Swartz et al., 2000; Hansen et al., 2013). However, significant correlations between accelerometer criteria and MOD-IPAQ and OH-IPAQ were demonstrated in females. Differences between the sexes maybe attributable to higher levels of OH-IPAQ minutes in females and differences between males and females in movement patterns and/or occupational/household activities.

Lower accelerometry cut-points than the commonly used cut-points of Freedson et al. (Freedson et al., 1998) have been recommended for older adults (Swartz et al., 2000; Matthews et al., 2005; Copeland & Esliger, 2009). In the present study however, correlations for IPAQ-L measures with the FL accelerometry cut-points (760-5724cpm) were not different from correlations with TM accelerometry cut-points (1952-5724cpm) when male and female data were combined. Two other validation studies in an elderly population compared different accelerometry cut-points (≥1952cpm, ≥100cpm, and ≥1,041cmp) to capture moderate intensity PA (Cerin et al., 2012; Van Holle et al., 2015), and both reported stronger correlations between the lower cut-point data and interview-administered IPAQ responses. In contrast with our data, these findings suggest that lower accelerometry cut-points more accurately reflect self-reported moderate intensity PA in the elderly. Although, we observed slightly stronger non-significant correlations for both MOD-IPAQ and OH-IPAQ and the lower accelerometry cut-points, it is unclear why our results differ from previous research. One explanation might be differences in age of the study participants, as the minimum age of participants in the aforementioned studies was older than those recruited to the present study (≥65y vs ≥60y). Furthermore, participants in these other studies were healthy (Cerin et al., 2012; Van Holle et al., 2015). Only one other study was identified using a clinical population (people diagnosed with coronary artery disease) to compare different accelerometer cut-points with self-reported PA (Prince et al, 2015). Cut-points were developed with a coronary artery disease population and younger seemingly healthy adults. Their findings demonstrated no superior correlation between self-report and lower cut-points compared to higher cut-points. However, the lowest cut-point threshold for moderate intensity PA applied in this study was 1800cpm, which is similar to the more conservative TM cut-points applied in the present study. These thresholds were also developed using treadmill walking. In light of the evidence, it may be that there is a threshold age or level of physical function at which lower accelerometer cut-points more accurately reflect moderate intensity PA. However, stronger correlations for MOD-IPAQ and OH-IPAQ were observed for FL accelerometry cut-points in our female participants, indicating that any such thresholds may be sex-specific, which warrants further research.

Overall, correlation coefficients for all accelerometer criteria were stronger for females than for males and this was significant for some of the PA criteria. Only one other validation study with the IPAQ-L stratified the results by sex and observed stronger correlations in males compared to females (Hagstromer et al., 2010), although p-values for this relationship were not reported. The observed sex differences in this study may be indicative of more accurate self-reporting by the female participants, although this is not consistent with previous systematic review evidence (Prince et al., 2008). Comparing outcomes of self-reported PA to accelerometery, females were found to over-report PA to a larger degree than males (Prince et al., 2008) but sex of the interviewer might influence responses. It was shown previously that males report a higher perceived exertion during cycling exercise in the presence of a female versus male observer (Winchester et al., 2012) and despite this contextual difference, the potential for sex effects needs to be taken into account (Janz, 2006). In the present study, as the interviewer was a female, there may have been less likelihood of over-reporting by female participants. Nevertheless, given our small sample size, there is a need for further investigations of sex differences in self-reported PA.

The Bland-Altman Limits of Agreement (LoA) analysis showed that overall, IPAQ-L over-estimates PA in relation to accelerometry-derived data in this population and this is in agreement with previous findings. In their systematic review, Prince et al., (2008) found that self-reported PA estimates are generally higher than estimates from objective measures, in the range of -78% to 500%. In this study we found that over-reporting of the IPAQ-L was less pronounced for the lower cut-point criteria (FL criteria), consistent with another study of older people which employed lower cut-points for moderate intensity PA (Van Holle et al., 2015). This is still a large difference between the measures, but indicates that using FL cut-points might be more suitable to measure a wider range of movements such as household activities in this population. Because the accelerometers used in this study were worn on the hip, the application of FL cut-points might not capture the whole range of movements of participants. There is evidence that different wear sites (waist, ankle, wrist, upper arm) may be more or less appropriate to capture particular movements at different speeds (Kim et al, Park, & Joo, 2014). However, it is unclear whether the over-estimation of self-reported PA measures compared to accelerometry is due to over-reporting or accelerometry limitations and this warrants further investigation in this population. The findings that total IPAQ data (walking + other activities) were not different from the FL cut-point but different from the TM cut-points also demonstrates that treadmill-derived cut-points may not be suitable to measure free-living PA in an elderly population.

The IPAQ-L records only PA that is carried out for at least 10 minutes or longer, and disregards any PA that does not meet this minimum PA duration criterion. It is therefore surprising that correlations between IPAQ-L and the 10 MIN bouts accelerometry criteria were not stronger than correlations between IPAQ-L and total accelerometer minutes. Data from Bland-Altman plots show a larger bias between the IPAQ-L and the 10 MIN accelerometer bouts for each of the cut-points applied. Again, this may be due to the limitation of accelerometers to record upper-body movement, and in light of this limitation, the total PA minutes recorded by accelerometers may be more reflective of the actual PA performed by the participants. This warrants further investigation.

This study had a number of limitations. The modest sample size means that the results can only be interpreted with caution. Despite the observed sex differences, further validation studies with larger sample sizes, including both males and females, are needed to confirm our findings. In addition, it is acknowledged that accelerometers cannot accurately measure varying intensities of some activities, e.g. walking at an incline or carrying heavy loads, cycling, swimming, upper-body activities, etc. (Welk, 2002; Kozey, Lyden, Howe, Staudenmayer, & Freedson, 2010; Hansen et al., 2013), which may confound interpretation of the data. It should also be noted that the characteristics of the study population are different from the populations of studies that developed the applied cut-points. The present sample was older (mean age 69 years compared to 22.9 to 42.0 years), had a higher BMI (28.7 kg/m^2 compared to 24.4 - 26.2 kg/m^2), and included participants that had been diagnosed with polyps or bowel cancer compared to healthy populations in the other studies (Freedson et al, 1998, Matthews et al, 2005). These differences could have influenced the classification of PA intensity from the accelerometers, as they were developed for a different population. Furthermore, participants could have been suffering from other co-morbidities, which were not screened for, reducing the accuracy of the accelerometers for similar reasons. Participants were excluded from the main trials if they presented with conditions that would prevent them from exercising safely. If none were reported, the authors did not collect additional information. Finally, administering the IPAQ-L in an interview-format could also be construed as a limitation, as social desirability might contribute to the over-reporting of PA (Janz, 2006). Nevertheless, our results show that convergent validity of an interview-administered IPAQ-L for the assessment of TOTAL-IPAQ and different sub-domains of PA in older people is comparable to previous validation studies.

In conclusion, our findings suggest that an interview-administered IPAQ-L may be more accurate than the self-administered IPAQ when recording WALK-PA and sedentary time in older populations. Although correlations between IPAQ measures and the FL accelerometry cut-points were not superior to correlations for TM accelerometry cut-points, the FL cut-points were associated with narrower limits of agreement (versus accelerometry data) and yielded stronger correlations in our female participants.

Table 1. Participant's characteristics

Characteristics	N=52
Sex (M/F)	32/20
Age in years	67.9 ± 6.6
Colorectal cancer survivors (N)	23
Time since diagnosis (years)	13.3 ± 9.4
Diagnosed with colorectal polyps (N)	29
Body weight (kg)	83.3 ± 16.9
BMI kg/m2	28.7 ± 4.6
Body fat (%)	30.9 ± 7.6
Waist-hip-ratio	0.93 ± 0.09

Data is shown in means (standard deviation) unless indicated otherwise, BMI= Body Mass Index

Table 2 Physical activity levels from IPAQ and accelerometry

Variable	All (n=52) Mean ± SD	Men (n=32) Mean ± SD	Women (n=20) Mean ± SD	P-Value
IPAQ (min·wk^{-1})				
Total IPAQ	441±301	431±297	456±316	0.93
Mod IPAQ	264±212	250±224	288±193	0.23
Walk IPAQ	176±199	182±211	168±182	0.82
leisure IPAQ	120±152	134±168	96±121	0.43
OH IPAQ	239±231	230±240	254±219	0.56
Sedentary	3025±1392	3193±1514	2742±1139	0.82
Accelerometry				
VM counts·min^{-1}	190±95	191±101	189±89	0.91
TM-ACC (min·wk^{-1})	120±110	100±99	153±122	0.14
TM-10MIN (min·wk^{-1})	53±81	46±85	64±76	0.19
FL-ACC (min·wk^{-1})	497±90	449±254	574±334	0.09
FL--10MIN (min·wk^{-1})	168±169	143±147	209±197	0.39
Steps·wk^{-1}	39939±12700	43711±5659	3872±14432	0.57
Sedentary time (min·wk^{-1})	3919±1380	4051±805	3708±757	0.21

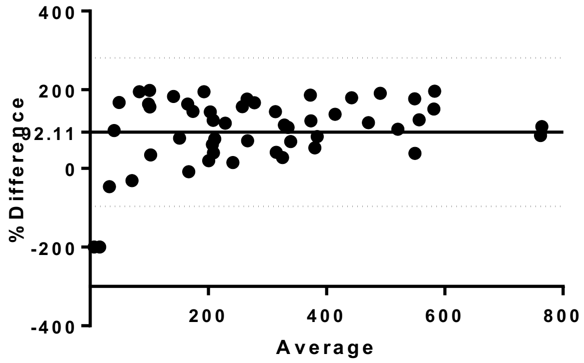
BMI= Body Mass Index, IPAQ= International Physical Activity Questionnaire, Mod= moderate intensity PA, OH= Occupational and Household related PA, ACC= total accumulated PA, VM= Vector magnitude, TM-10MIN and TM-ACC corresponds to bouts of ≥10min or total accumulated PA at 1952-5724cpm, FL-10MIN and FL-ACC corresponds to bouts of ≥10min or total accumulated PA at 760-5724cpm

Table 3 Spearman correlation coefficients (r) between IPAQ-L and accelerometer-based Measures in overall sample**.**

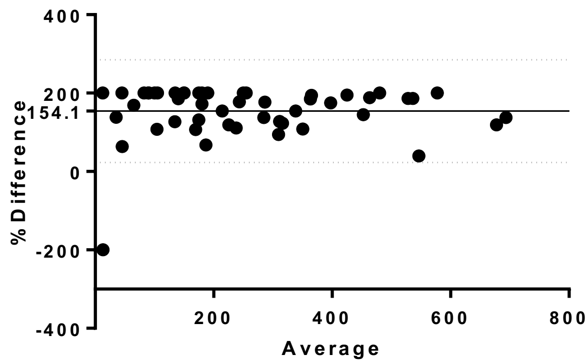
	TM-ACC	**TM-10MIN**	**FL-ACC**	**FL-10MIN**	**Step count**	**VM**	**Sedentary time**
Total IPAQ	**.39**[b]	**.43**[b]	**0.40**[b]	**.38**[b]	**.46**[a]	.27	
Mod IPAQ	.16	.16	0.27	.23	.23	.14	
Walk IPAQ	**.54**[c]	**.57**[c]	**0.38**[b]	**.41**[b]	**.49**[a]	**.34**[a]	
Leisure IPAQ	**.45**[b]	**.44**[b]	**0.30**[a]	**.44**[b]	**.47**[a]	**.32**[a]	
OH IPAQ	-.08	.19	0.27	.25	.31	.12	
Sedentary							**.33**[a]

[a]P≤0.05, [b]P≤0.01, [c]P≤0.001, PA= physical activity, , IPAQ= International Physical Activity Questionnaire, Mod= moderate intensity PA, OH= Occupational and Household related PA, ACC= total accumulated PA, VM= Vector magnitude, TM-10MIN and TM-ACC corresponds to bouts of ≥10min or total accumulated PA at 1952-5724cpm, FL-10MIN and FL-ACC corresponds to bouts of ≥10min or total accumulated PA at 760-5724cpm

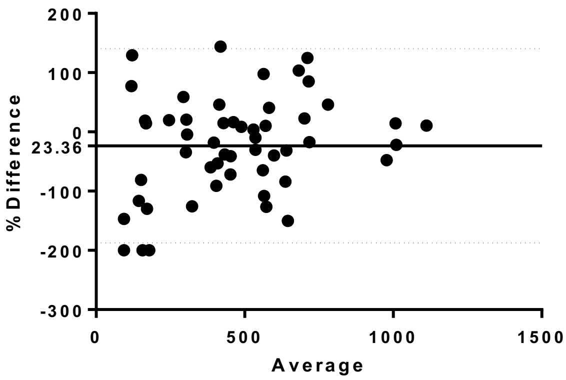
Figure 1 %Differences between total self-reported physical activity and accelerometry (TM-ACC, FL-ACC, TM-10MIN, and FL-10MIN)



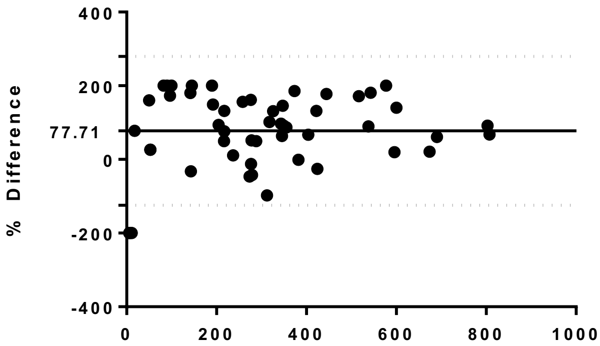
A



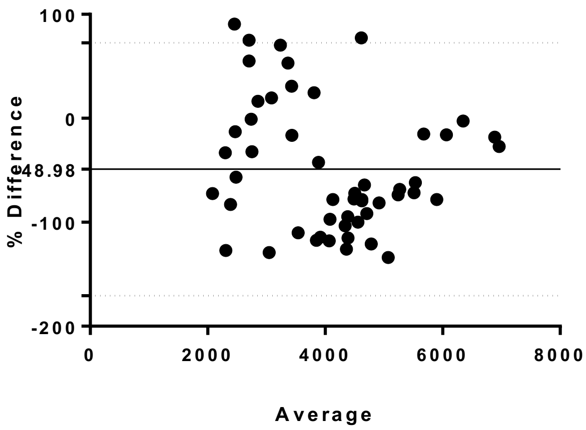
B



C



D



E

A) %Difference vs. average: Total IPAQ – TM-ACC, **B)** %Difference vs. average: Total IPAQ – TM-10MIN, **C)** %Difference vs. average: Total IPAQ – FL-ACC, **D)** %Difference vs. average: Total IPAQ – FL-10MIN, **E)** %Difference vs. average: IPAQ sedentary time –

Accelerometer sedentary time, IPAQ= International Physical Activity Questionnaire, Mod= moderate intensity PA, OH= Occupational and Household related PA, ACC= total accumulated PA, VM= Vector magnitude, TM-10MIN and TM-ACC corresponds to bouts of ≥10min or total accumulated PA at 1952-5724cpm, FL-10MIN and FL-ACC corresponds to bouts of ≥10min or total accumulated PA at 760-5724cpm, Dotted lines represent limits of agreement; black line represents %bias

Table 4 Spearman correlation coefficients between IPAQ-L and accelerometer-based measures by gender

	TM-ACC		**P-value**	**TM-10MIN**		**P-value**	**FL-ACC**		**P-value**	**FL-10MIN**		**P-value**	**Sedentary time**		**P-value**
	M	**W**		**M**	**W**		**M**	**W**		**M**	**W**		**M**	**W**	
IPAQ															
Total	.24	**.71**[b]	**.04**	.32	**.58**[b]	.28	0.19	**0.71**[c]	**.01**	.19	**.62**[b]	.08	-.20	-.17	.92
Mod IPAQ	.02	**.50**[a]	.08	.10	.33	.42	0.10	**0.61**[b]	**.05**	.08	**.58**[a]	**.05**	-.25	-.15	.73
Walk IPAQ	**.42**[a]	**.84**[c]	**.01**	**.40**[a]	**.81**[b]	**.02**	0.25	**0.62**[b]	.08	.29	**.60**[b]	.20	.05	-.21	.39
Leisure IPAQ	**.41**[a]	**.73**[b]	.11	**.37**[a]	**.57**[a]	.39	0.20	**0.58**[a]	.13	**.36**[a]	**.65**[b]	.20	.16	-.18	.26
OH IPAQ	.02	**.46**[a]	.12	.17	.26	.75	0.10	**0.58**[a]	.06	.01	**.60**[b]	**.02**	**-.37**[a]	-.14	.42
Sedentary													**.40**[a]	.11	.30

[a]P≤0.05, [b]P≤0.01, [c]P≤0.001, NS= not significant, M=Men, W=Women, I, IPAQ= International Physical Activity Questionnaire, Mod= moderate intensity PA, OH= Occupational and Household related PA, ACC= total accumulated PA, VM= Vector magnitude, TM-10MIN and TM-ACC corresponds to bouts of ≥10min or total accumulated PA at 1952-5724cpm, FL-10MIN and FL-ACC corresponds to bouts of ≥10min or total accumulated PA at 760-5724cpm

References

Bland, J.M., & Altman, D.G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet, 327*(8476), 307-310.

Borg, G.A. (1982). Psychophysical bases of perceived exertion. *Medicine & Science in Sports & Exercise, 14*(5), 377-381.

Cerin, E., Barnett, A., Cheung, M., Sit, P., Cindy, H., Macfarlane, D.J., & Chan, W.. (2012). Reliability and validity of the IPAQ-L in a sample of Hong Kong urban older adults: does neighborhood of residence matter? *Journal of Aging & Physical Activity, 20*(4), 402-420.

Choi, L., Liu, Z., Matthews, C.E., & Buchowski, M.S.. (2011). Validation of accelerometer wear and nonwear time classification algorithm. *Medicine and Science in Sports and Exercise, 43*(2), 357-364.

Copeland, J.L & Esliger, D.W. (2009). Accelerometer assessment of physical activity in active, healthy older adults. *Journal of Aging & Physical Activity, 17*(1), 17-30.

Craig, C.L., Marshall, A.L., Sjostrom, M., Bauman, A.E., Booth, M.L., Ainsworht, B.E., Oja, P. (2003). International physical activity questionnaire: 12-country reliability and validity. *Medicine & Science in Sports & Exercise, 195*(9131/03), 3508-1381.

Department of Health. (2011). Start Active, Stay Active: A report on physical activity for health from the four home countries'. Retrieved from London: www.bhfactive.org.uk

Freedson, P.S., Melanson, E., & Sirard, J. (1998). Calibration of the Computer Science and Applications, Inc. accelerometer. *Medicine and Science in Sports and Exercise, 30*(5), 777-781.

Hagstromer, M., Ainsworth, B.E., Oja, P., & Sjostrom, M. (2010). Comparison of a subjective and an objective measure of physical activity in a population sample. *Journal of Physical Activity & Health, 7*(4), 541-550.

Hagstromer, M., Oja, P., & Sjostrom, M.. (2006). The International Physical Activity Questionnaire (IPAQ): a study of concurrent and construct validity. *Public Health Nutrition, 9*(6), 755-762. Retrieved from

Hall, K.S., Howe, C.A., Rana, S.R., Martin, C.L., & Morey, M.C. (2013). METs and accelerometry of walking in older adults: standard versus measured energy cost. *Medicine and Science in Sports and Exercise, 45*(3), 574-582.

Hansen, B.H., Børtnes, I., Hildebrand, M., Holme, I., Kolle, E., & Anderssen, S.A.. (2013). Validity of the ActiGraph GT1M during walking and cycling. *Journal of Sports Sciences, 32*(6), 510-516.

Hendelman, D., Miller, K., Baggett, C., Debold, E., & Freedson, P. (2000). Validity of accelerometry for the assessment of moderate intensity physical activity in the field. *Medicine and Science in Sports and Exercise, 32*(9; SUPP/1), S442-449.

Hopkins, W. (2002). A New View of Statistics: Correlation Coefficient. http://www.sportsci.org/resource/stats/index.html, accessed on 10.02.2016

Janz, K.F. (2006). Physical activity in epidemiology: moving from questionnaire to objective measurement. *British Journal of Sports Medicine, 40*(3), 191-192.

Kim, D.Y., Jung, Y., Park, R., & Joo, N.. (2014). Different location of triaxial accelerometer and different energy expenditures. *Yonsei Medical Journal*, 55(4), 1145-1151

Kozey, S.L., Lyden, K., Howe, C.A., Staudenmayer, J.W., & Freedson, P.S. (2010). Accelerometer output and MET values of common physical activities. *Medicine and Science in Sports and Exercise, 42*(9), 1776-1784.

Macfarlane, D., Chan, A., & Cerin, E.. (2011). Examining the validity and reliability of the Chinese version of the International Physical Activity Questionnaire, long form (IPAQ-LC). *Public Health Nutrition, 14*(03), 443-450.

Matthews, C.E. (2005). Calibration of accelerometer output for adults. *Medicine and Science in Sports and Exercise, 37*(11), S512-522.

Matthews, C.E., Ainsworth, B.E., Hanby, C., Pate, R.R., Addy, C., Freedson, P.S., Macera, C.A. (2005). Development and testing of a short physical activity recall questionnaire. *Medicine and Science in Sports and Exercise, 37*(6), 986-994.

Matthews, C.E, Chen, K., Freedson, P.S., Buchowski, M.S., Beech, B.M., Pate, R.R., & Troiano, R.P. (2008). Amount of time spent in sedentary behaviors in the United States, 2003–2004. *American Journal of Epidemiology*, 167(7), 875-881

Miller, N.E., Strath, S.J., Swartz, A.M., & Cashin, S. E. (2010). Estimating absolute and relative physical activity intensity across age via accelerometry in adults. *Journal of Aging and Physical Activity, 18*(2), 158-170.

Nelson, M.E., Rejeski, W., Blair, S.N., Duncan, P.W., Judge, J.O., King, A.C., Castaneda-Sceppa, C. (2007). Physical activity and public health in older adults: recommendation from the American College of Sports Medicine and the American Heart Association. *Medicine & Science in Sports & Exercise, 39*(8), 1435-1445.

Pollock, M.L., Gaesser, G.A., Butcher, J.D., Després, J., Dishman, R.K., Franklin, B.A., & Garber, C.Ewing. (1998). ACSM position stand: the recommended quantity and quality of exercise for developing and maintaining cardiorespiratory and muscular fitness, and flexibility in healthy adults. *Medicine & Science in Sports & Exercise, 30*(6), 975-991.

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. *International Journal of Behavioral Nutrition and Physical Activity, 5*(56). doi: 10.1186/1479-5868-5-56

Prince, S.A., Reed, J.L., Mark, A.E., Blanchard, C.M., Grace, S.L., & Reid, R.D. (2015). A comparison of accelerometer cut-points among individuals with coronary artery disease. *PLoS One [Electronic Resource],* 10(9), e0137759.

Swartz, A.M., Strath, S.J., Bassett, D.R., O Brien, W.L., King, G.A., & Ainsworth, B.E. (2000). Estimation of energy expenditure using CSA accelerometers at hip and wrist sites. *Medicine and Science in Sports and Exercise, 32*(9; SUPP/1), S450-S456.

The IPAQ Group. IPAQ scoring protocol. www.ipaq.ki.se, Accessed on 01.08.2014

Van Holle, V., De Bourdeaudhuij, I., Deforche, B., Van Cauwenberg, J., & Van Dyck, D.. (2015). Assessment of physical activity in older Belgian adults: validity and reliability of an adapted interview version of the long International Physical Activity Questionnaire (IPAQ-L). *BMC Public Health, 15*(1), 433-447.

Welk, G.J. (2002). Use of accelerometry-based activity monitors to assess physical activity. *Physical Activity Assessments for Health Related Research*, 125-141.

Winchester, R., Turner, L. A., Thomas, K., Ansley, L., Thompson, K. G., Micklewright, D., & St Clair Gibson, A. (2012). Observer effects on the rating of perceived exertion and affect during exercise in recreationally active males. *Perceptual & Motor Skills, 115*(1), 213-227. doi:10.2466/25.07.05.pms.115.4.213-227