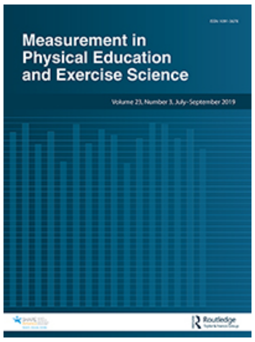

Measurement in Physical Education and Exercise Science

ISSN: 1091-367X (Print) 1532-7841 (Online) Journal homepage: https://www.tandfonline.com/loi/hmpe20

Comparison of Self-Reported Sedentary Time on Weekdays with an Objective Measure (activPAL)

Kaja Kastelic & Nejc Šarabon

To cite this article: Kaja Kastelic & Nejc Šarabon (2019) Comparison of Self-Reported Sedentary Time on Weekdays with an Objective Measure (activPAL), Measurement in Physical Education and Exercise Science, 23:3, 227-236, DOI: 10.1080/1091367X.2019.1603153

To link to this article: https://doi.org/10.1080/1091367X.2019.1603153

Published online: 18 Apr 2019.

MEASUREMENT IN PHYSICAL EDUCATION AND EXERCISE SCIENCE
2019, VOL. 23, NO. 3, 227–236
https://doi.org/10.1080/1091367X.2019.1603153

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

Comparison of Self-Reported Sedentary Time on Weekdays with an Objective Measure (activPAL)

Kaja Kastelic [a,b] and Nejc Šarabon [a,b,c]

[a]Andrej Marušič Institute, Department of Health Study, University of Primorska, Koper, Slovenia; [b]S2P, Science to Practice, Ltd., Laboratory for Motor Control and Motor Behavior, Ljubljana, Slovenia; [c]Faculty of Health Sciences, University of Primorska, Izola, Slovenia

ABSTRACT
Self-reports are commonly used tools for obtaining sedentary behaviors. The aim of our study was to assess agreement between two self-reports of sedentary time and a gold standard sedentary time objective monitor. A worksite sample (n = 42) completed the Slovenian version of the Global Physical Activity Questionnaire (GPAQ), the Slovenian version of the Sedentary Behaviour Questionnaire (SBQ for weekdays) and wore an objective physical activity monitor (activPAL) for up to five consecutive working days. Results revealed that GPAQ and SBQ consistently underestimated the total sedentary time, with the mean bias of −165 min/day and −181 min/day, respectably. Wide limits of agreement showed poor precision and intraclass correlation revealed a low level of agreement. GPAQ and SBQ are not recommended to be used in studies seeking for relationships with health outcomes, nor in studies where detecting the behavioral change is of interest. Objective measurement should be the preferred choice when possible.

KEYWORDS
Validity; error; bias; concordance; sedentary lifestyle

Introduction

In the last two decades, sedentary behavior became an important topic in the field of physical activity for health. It has been shown that being excessively sedentary is an independent risk factor for several non-communicable diseases and all-cause mortality (de Rezende, Rodrigues Lopes, Rey-Lopez, Matsudo, & Luiz Odo, 2014). Sedentary behavior became a priority public health concern, since it is ubiquitous in contemporary society. Many are sitting for prolonged periods of time while at the workplace, during commuting with motorized or railed vehicles, at home using electronic entertainment technologies, etc. Several health organizations and national policies have already issued the recommendations for avoiding sedentary behavior (Davies, Burns, Jewell, & Mcbride, 2011; Garber et al., 2011).

For scientific research, as well as for population surveillance, quality measurement of sedentary time and patterns of sedentary behavior is essential. The most pragmatic and commonly used tools are questionnaires, which offer an insight not only into the total time spent sedentary but also into the other aspects, such as the type and the context of behavior (Dall, Coulter, Fitzsimons, Skelton, & Chastin, 2017). Questionnaires rely on the respondents' recall ability and honesty in providing the answers (Kang & Rowe, 2015), which might importantly compromise the validity of the results. In order to evaluate the accuracy and precision of self-reported time spent in sedentary behavior, a comparison with a criterion measure (i.e. the gold standard) is to be performed (Kelly, Fitzsimons, & Baker, 2016).

Until recently, hip-worn accelerometers (e.g. ActiGraph) have been a preferred choice for being a criterion measure of sedentary time (Dall et al., 2017). However, their important limitation of erroneous classification of all (quasi-) static postures (which are characterized with a lack of movement) as sedentary behavior (Janssen & Cliff, 2015), has been overcome by a new generation of accelerometers, which were designed to distinguish between body postures. In particular, the thigh-worn activPAL measures the thigh inclination in order to predict time spent sitting/lying, standing, and stepping (Sellers, Dall, Grant, & Stansfield, 2016). Inter-device reliability for sitting/lying time was shown to be very high (Grant, Ryan, Tigbe, & Granat, 2006). Also, the activPAL has perfect correlation and excellent agreement with direct observation for a total sitting/lying time in the adult population (Edwardson et al., 2017). The device is currently regarded as the gold standard for the objective measurement of sedentary behavior (Chastin et al., 2018; Kang & Rowe, 2015; Kim, Barry, & Kang, 2015).

CONTACT Nejc Šarabon nejc.sarabon@fvz.upr.si Faculty of Health Sciences, University of Primorska, Polje 42, Izola SI-6310, Slovenia

To date, only a few sedentary behavior questionnaires have been validated against the activPAL (Dall et al., 2017). In particular, a World Health Organisation's Global Physical Activity Questionnaire (GPAQ) which consists of 15 questions about physical activity and one question about sedentary behavior (Armstrong & Bull, 2006), have been validated only against motion-based accelerometers (Bull, Maslin, & Armstrong, 2009; Chu, Ng, Koh, & Müller-Riemenschneider, 2018; Cleland et al., 2014; Herrman, Heumann, Der Ananian, & Ainsworth, 2013; Trinh, Do Nguyen, van der Ploeg, Dibley, & Bauman, 2009; Wanner et al., 2017). However, the choice of objective monitor in those studies seems prudent, since motion-based accelerometers are a preferred choice when measuring physical activity, but not sedentary behavior (Kang & Rowe, 2015). There is a need for validation of the GPAQ sitting item against the activPAL. Likewise, the Sedentary Behaviour Questionnaire (SBQ) which measures time spent in different sedentary behaviors (Rosenberg et al., 2010), has never been solely validated against posture-based accelerometers. The aim of this study was to assess agreement between the GPAQ sitting item and the SBQ against the gold standard sedentary time objective monitor (activPAL) in the population of sedentary workers.

Methods

Participants recruitment

A convenience sample of sedentary workers (crane operators and office workers) from a Slovenian logistics company (Port of Koper) were recruited via an oral invitation. Workers included in the study were scheduled for participation. All participants considered themselves as healthy and able to work and mobilize normally and could read, speak and understand the Slovenian language. They have signed an informed consent statement prior to the experiment. The study was conducted in line with the Helsinki declaration and approved by the National Medical Ethics Committee (approval number: 0120-557/2017/4).

According to the published recommendations, a sample size of 30 participants is sufficient to ensure adequate statistical power (80%, $\alpha = 0.05$) to identify a moderate correlation ($r = 0.50$) between self-reported and objective measures (Prince, LeBlanc, Colley, & Saunders, 2017). Besides that, we also carried out a pilot study ($n = 9$) and calculated sample size for correlation analyses and the paired t-test. We used G*Power 3.1 software (Faul, Erdfelder, Lang, & Buchner, 2007) and $n > 28$ turned out to be sufficient.

Self-reported measures

The Global Physical Activity Questionnaire (version 2; Armstrong & Bull, 2006) contains 15 questions, asking about physical activity and one question about sedentary behavior. In this study, we used only the sedentary behavior question (item 16). The question asks respondents to estimate the usual total number of hours and minutes on a typical day spent sitting or reclining (at work, at home, getting to and from places, or with friends; not including time spent sleeping).

The Sedentary Behaviour Questionnaire (Rosenberg et al., 2010) asks about time spent doing nine behaviors (watching television, playing computer/video games, sitting while listening to music, sitting and talking on the phone, doing paperwork or office work, sitting and reading, playing a musical instrument, doing arts and crafts, sitting and driving/riding in a car, bus, or train) on a typical weekday and weekend day, separately. Response options are none, 15 min or less, 30 min, 1 hr, 2 hr, 3 hr, 4 hr, 5 hr, or 6 hr or more. In this study, only the questions asking about weekdays were evaluated.

Both questionnaires are written in the English language. Since our participants were Slovenian, the questionnaires were translated into Slovenian language by a researcher who is bilingual and familiar with the area in which the questionnaire is used. Then, we asked five people (two researchers and three non-experts), who are also bilingual and were previously not familiar with the GPAQ and SBQ, for reverse translation. All five translators stated that the questionnaires were understandable and easy to translate. No modifications were found.

The participants were also asked to record a diary of daily activities and sleep time. This information was used in the process of isolating waking hours, which is described in the Data management sub-section.

Objective measure

The activPAL3™ micro (PAL Technologies Ltd., Glasgow, Scotland) is a light and small (9 g; 55 × 25 x 5 mm) physical activity monitor that quantifies postural allocation. Placed on the subject's anterior thigh it measures thigh inclination directly and provides robust identification of sitting/lying, standing and stepping periods (Sellers et al., 2016). It is currently regarded as the gold standard for objective measurement of sedentary time (Chastin et al., 2018; Kang & Rowe, 2015; Kim et al., 2015).

Study design

On day one, participants were introduced to the study and equipped with an activPAL to obtain objective and continuous measures of sitting/lying time. For the purpose of this study, the acquisition settings of the activPAL were set as default (20 Hz sampling frequency, 10 s criteria to register a new event). The activPAL was attached by the researcher on the participants' right thigh (anterior aspect of the midline of the thigh, half the distance between the knee and the hip) using flexible waterproof sleeves and waterproof medical grade adhesive dressing (Tegaderm^{TM}). They were requested to wear the activPAL continuously 24 h per day and instructed how to act with it (dressing up with caution, showering rather than bathing, avoid swimming). They were also introduced to the diary of daily activities and sleeping time. Self-reported demographic information (sex, age, height, weight) were collected. On the third day, participants were invited to meet the researcher again. The activPALs' adhesive dressing was checked and changed if needed. At that occasion, the participants also fulfilled the sitting item (item 16) from the GPAQ (interview-administered) and SBQ (for weekdays only). The researcher reviewed the questionnaires in the presence of the participant. In case of missing data, the participant was asked to answer the question. On the fifth day (at the end of their working hours), participants returned the measurement device and diaries.

Data management

Results from both questionnaires (GPAQ and SBQ) were entered into the computer manually. Values in all the answers were converted into minutes. A response of "15 min or less" from SBQ was converted to 15 min and a response of "6 hr or more" into 360 min (Rosenberg et al., 2010). Total sedentary time from SBQ was calculated as a sum of behaviors.

The data from the activPAL were downloaded using the activPAL3 software (version 7.2.32, PAL Technologies Ltd., Glasgow, Scotland) and exported as .csv files in a pre-classified form (15 s epoch summary file of sitting/lying, standing, stepping) by each measured calendar day. The software package Microsoft Excel 2007 was used for the further processing steps. The person-oriented day approach was used to define a day (Stephens et al., 2014). A valid day consists of continuous activPAL data from one wake time to the next day wake time. Wear time was recorded by a researcher who had placed and removed the activPAL from the participants' thigh (before the removal the adhesive dressing was inspected visually for an existence of any traces, which might indicate preterm removal). Additionally, the data were visually inspected via heat maps to ensure that no non-wear time exists (Edwardson et al., 2017). Waking hours were considered to be from the time a person got up from the bed to the time going to bed. Any reported daytime naps were categorized as sleep time, whereas any wakefulness during the bedtime was categorized as sleep-related behavior (unless some specific activity has been reported in the diary) and was not considered as wake time (Barone Gibbs & Kline, 2018).

The end and start of the bedtime were visually (manually) identified from the 15-s epoch summary file, starting around self-reported times. The end of the bedtime was defined as the first epoch containing some standing or walking after >1 h of exclusively sitting/lying epochs and start of bedtime as the first exclusively sitting/lying epoch followed by >2 h of sitting/lying epochs (a similar approach was applied in Chastin, Culhane, & Dalet, 2014; Dowd, Harrington, Bourke, Nelson, & Donnelly, 2012). Total sitting/lying time while awake (i.e. sedentary time) was calculated by summing up the duration of sitting/lying times within wake time epochs (also, the time of self-reported daytime naps were deducted, if exists). Participants with at least three valid weekdays of activPAL data (Pontt, Rowlands, & Dollman, 2015) and fulfilled GPAQ and SBQ, were included in the further analysis.

Statistical analysis

Statistical analyses were performed using the statistical package SPSS, version 18.0 (SPSS Inc., Illinois, USA) and version 25.0 with Python extension. For each participant, average sitting/lying time (in minutes) was calculated from at least three valid weekdays of activPAL data. Descriptive statistics were performed for all measures of sitting/lying time (GPAQ, SBQ, activPAL). The agreement analysis between self-reports against the activPAL was performed according to the recommendations (Watson & Petrie, 2010). The difference between each of the n-pairs (GPAQ – activPAL and SBQ – activPAL) was calculated and tested for normality using Shapiro–Wilk test. Then, a paired sample t-test was applied to test for systematic differences. Since the activPAL was regarded as a gold standard, Krouwer's (2008) method was used to construct the Bland–Altman diagrams. An acceptable difference between the self-reports and objective measure was defined to be 1 hr (based on the current findings of the dose–response relationship between sedentary behavior and health outcome (Patterson et al., 2018)). The two-way mixed model intraclass correlation coefficient (ICC (2,1)) was also used to measure the extent of the agreement between self-reports and the activPAL. Additionally, categories of tertiles and quartiles of

Table 1. Comparison of the GPAQ sitting item and SBQ (for weekdays) with activPAL data of average daily sedentary time (670 ± 101 min/day).

Self-reports	Mean ± SD [mins/day]	Mean difference [95% CI] [mins/day]	SD of difference [mins/day]	Intraclass correlation	Pearson/Spearmans' rho
GPAQ	489 ± 107	−165 [−207, −123]	135	0.207 (p = .058)	0.317 (p = .041)
SBQ	505 ± 118	−181 [−227, −136]	146	0.014 (p = .454)	0.018 (p = .910)

GPAQ: Global Physical Activity Questionnaire; SBQ: Sedentary Behaviour Questionnaire; SD: standard deviation; 95% CI: 95% confidence interval.

sedentary time were calculated and Weighted Kappa applied to determine the agreement between the categorized levels (GPAQ – activPAL and SBQ – activPAL). Normality for all measures of sitting/lying time was tested using the Shapiro–Wilk test. Pearson's correlation (or Spearman's rho in case of non-normal data distribution) was used to test the linear (monotonic) relationship between total sitting/lying time obtained by the self-reports and the activPAL. Statistical significance was set at $p < .05$, and effect sizes (ES) were calculated.

Results

Participants

Sixty-four sedentary workers participated in the study. The inclusion criteria of at least three valid weekdays of measurement (activPAL data) and completed GPAQ and SBQ were not met by 22 participants, who were excluded from the analysis. Forty-two participants (29 crane operators and 13 office workers) achieved the inclusion criteria and were included into further analysis (37 males (38 ± 8 years, 180 ± 7 cm, 89 ± 15 kg), 5 females (50 ± 7 years, 167 ± 3 cm, 63 ± 6 kg)).

Comparison of GPAQ sitting item and activPAL

The mean difference between each of the n-pairs (GPAQ – activPAL) was −165 min/day, which indicates poor accuracy between the measures at the individual level. Confidence intervals revealed that the GPAQ sitting item consistently underestimated sedentary time (Table 1). The difference between the n-pairs was normally distributed. A paired t-test ($p < .001$) showed a systematic difference (systematic error) between the two methods (95% CI −207, −123; ES = −1.2). Since activPAL is regarded as the gold standard, the systematic difference implies that there was a bias in the mean. The standard deviation of the difference (random error) was 135 min/day, which showed poor precision of the GPAQ. The intraclass correlation coefficient (ICC = 0.207, 95% CI −0.181, 0.514) displayed a low level of agreement between GPAQ sitting item and activPAL. Additionally, a weak monotonic relationship using Spearman's rho ($\rho = 0.317$, $p = .041$, ES = 0.100) was found. The agreement between GPAQ and activPAL was fair when ranking sitting time into tertiles (47.6%, k = 0.316 (95% CI 0.088, 0.543), p = .01) and poor for quartiles (28.6%, k = 0.143 (95% CI −0.056, 0.342), p = .193).

In a Bland–Altman diagram (Figure 1) using Krouwer's method, the difference between a pair is plotted on the vertical axis and the criterion measure on the horizontal axis. The diagram confirmed that there was a constant underestimation by the GPAQ sitting item. Limits of agreement were very wide, ranging from −429 min/day to 99 min/day, showing poor precision. The regression line ($y = -0.712\,x + 312.06$; $R^2 = 0.266$) showed negative proportional bias ($\beta = -0.533$, $p < .001$). However, if the difference between the methods was expressed relatively, then no bias exists ($\beta = 0.130$, $p = .411$). Negative proportional bias appeared to be relative to the amount of sitting.

Comparison of SBQ (for weekdays) and activPAL

Total sedentary time obtained by the SBQ was also consistently underestimated (on average −181 min/day). The t-test revealed that there was a systematic difference ($p < .001$) between the SBQ and the

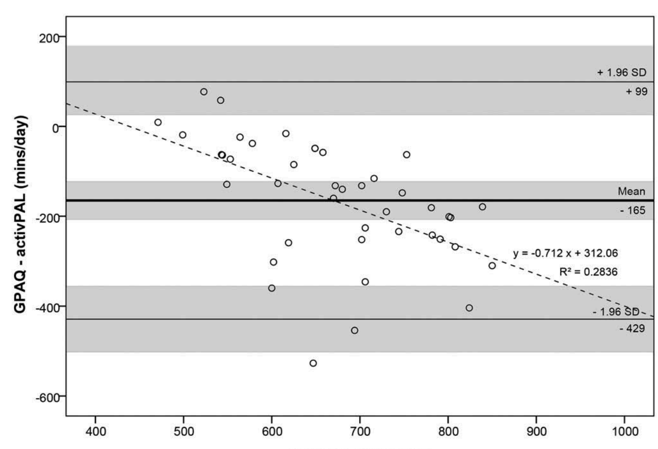

Figure 1. Bland–Altman diagram for the average sedentary time (min/day) measured by GPAQ and the activPAL. Bold line shows the mean difference, black lines show the limits of agreement at 95% level and dashed line shows the linear regression trend. Shaded areas present the 95% confidence intervals for the mean and agreement limits.

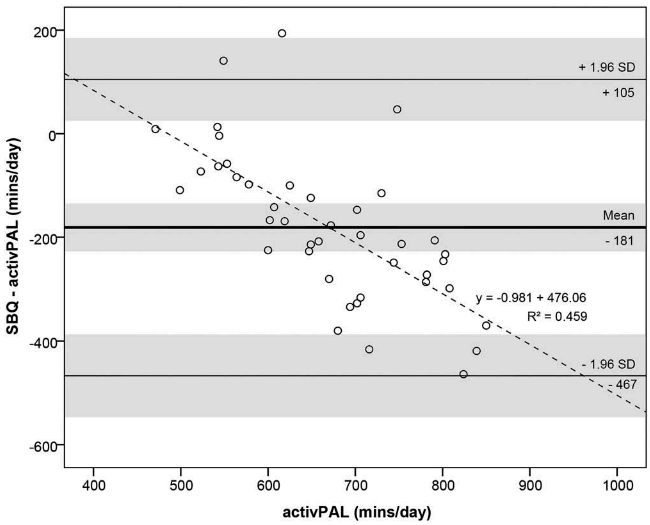

Figure 2. Bland–Altman diagram for the average sedentary time (mins/day) measured by SBQ (for weekdays) and the activPAL. Bold line shows the mean difference, black lines show the limits of agreement at 95% level and dashed line shows the linear regression trend. Shaded areas present the 95% confidence intervals for the mean and agreement limits.

activPAL (95% CI −227, −136; ES = −1.2), which shows on the presence of bias. The standard deviation of the difference (146 min/day) and wide limits of agreement (which can be seen on the Bland–Altman diagram (Figure 2)) indicating that the precision of the SBQ is poor. There is also negative proportional bias (β = −0.678, $p < .001$). No monotonic relationship (r = 0.018, p = .910, ES < 0.000) and low level of agreement (ICC = 0.014, 95% CI −0.212, 0.269) was found. Ranking participants into tertiles (31.0%, k = 0.018 (95% CI −0.212, 0.248), p = .880) and quartiles (31.0%, k = 0.105 (95% CI −0.108, 0.319), p = .344) of sedentary time reviled poor agreement between SBQ and the activPAL.

Discussion

This study aimed to assess the agreement between the GPAQ sitting item and the SBQ (for weekdays) against the gold standard sedentary time objective monitor (activPAL) in the population of sedentary workers. The results demonstrated a low level of agreement between self-reports and the activPAL. The GPAQ and the SBQ underestimated total sedentary time for an average of nearly 3 hr per day. According to the objective measure (activPAL), our cohort of sedentary workers spent on average more than 11 hr per day in sedentary behavior, making them a highly sedentary group.

GPAQ sitting item

Asking about the time one spends sitting or reclining on a typical day is probably the simplest way to evaluate total sedentary time. Such a single item question is part of the World Health Organization's questionnaire about physical activity and sedentary behavior (GPAQ). The questionnaire was developed as a surveillance tool for evaluation and comparisons between culturally diverse populations worldwide (Armstrong & Bull, 2006). To date, several validation studies including different populations have been performed (Alkahtani, 2016; Aguilar-Farias & Leppe Zamora, 2017; Bull et al., 2009; Chu et al., 2018; Cleland et al., 2014; Herrman et al., 2013; Hoos, Espinoza, Marshall, & Arredondo, 2012; Trinh et al., 2009; Wanner et al., 2017;). Overall, the studies show an acceptable reliability of the GPAQ sitting item, whereas validity was shown to be poor. Most often, a low correlation with motion-based accelerometers (e.g. ActiGraph) has been reported. The GPAQ tends to underestimate total sedentary time by an average of 88 min/day to 420 min/day, depending on the population tested (note that some difference might also be attributed to objective data collection and processing decisions (Janssen & Cliff, 2015)). Additionally, the precision of the GPAQ and the level of agreement between the two methods was also reported to be low (Alkahtani, 2016; Aguilar-Farias & Leppe Zamora, 2017; Chu et al., 2018; Cleland 2014; Wanner et al., 2017). However, using the posture-based accelerometer (activPAL) instead of the ActiGraph would present a more valid criterion measure (Kim et al., 2015).

To our knowledge, this study was the first to report an agreement between the GPAQ sitting item with the activPAL. Including the Slovenian population of sedentary workers, our findings were in line with previous studies. The GPAQ has been only rarely validated in the European countries and findings somehow differ. In 2014, Cleland and co-workers performed a study including the Irish population and concluded that GPAQ is not a valid measure of time spent in sedentary behavior or change in sedentary behavior. They found a poor correlation (r = 0.187, $p > .05$) between GPAQ sitting item and the ActiGraph and a significant underestimation (for an average of 349 min/day with wide limits of agreement) of sedentary time as reported in the GPAQ. More recently, Wanner et al. (2017) performed a similar study including Swiss, German, Italian and French populations and reported the moderate correlation (r = 0.47, $p < .001$) and an underestimation of 88 min/day (ActiGraph produces only 1.2 times higher sitting time than GPAQ). The authors assume that substantially higher correlation that was reported in other studies (Aguilar-Farias & Leppe Zamora, 2017; Cleland et al., 2014; Mumu, Ali, Barnett, & Merom, 2017) might be due to the use of self-administered format of GPAQ. Indeed, this assumption has been supported recently (Chu et al., 2018). In our study, we used the interview-administered Slovenian version of the GPAQ and obtained a low correlation ($\rho = 0.317$, $p < .05$) with the activPAL. Previous reports indicate that female-only shows higher correlation results than male-only (Cleland et al., 2014), therefore a sex-balanced sample may have produced somewhat higher correlation in our case (our sample consisted of substantially more male than females). Also, the correlation might be higher if self-administered version would be used.

Our GPAQ data underestimated sedentary time by 165 min/day (activPAL produces 1.33 times higher sitting time than GPAQ). Underestimation might be a consequence of recall ability or social desirability error. Since sedentary behavior is not a structured activity, rather it occurs persistently in periods throughout the day, it is likely that recall ability is substantially compromised. Significant underestimation of the sedentary time when using the GPAQ affects the population estimates. In order to provide more accurate estimates or to compare the results from the GPAQ with the results obtained with other tools, a correction factor (set equal to systematic error) can be used at the group or population level. Chastin et al. (2018) have recently provided a correction factor of 250.6 min to be used on single item questionnaires (such as the GPAQ sitting item). The cohort of Chastin on which the correction factor was calculated included only older adults; thus, the results might not be generalizable to other populations. We propose that a correction factor of +33% (which is 165 min in our case) should be used on similar populations like in our study (young- and middle-aged adult population from Europe).

Due to low correlation and large random error, using the GPAQ in epidemiological studies where the relationship between sedentary behavior and health outcomes is of interest will omit or overlook the significance of the exposure on health (Chastin et al., 2018). This might also be true in case of categorizing participants into ranks of sedentary time. Aguilar-Farias and Leppe Zamora (2017) reported that agreement between GPAQ sitting item and ActiGraph for categorizing individuals into tertiles or quartiles of sedentary time was poor. Based on our results, the agreement between GPAQ and activPAL for tertiles was fair and for quartiles poor. In both cases, less than 50% of the participants were correctly classified, which make the GPAQ less appropriate to such use.

The Bland–Altman analysis revealed that the limits of agreement were wide, ranging from −429 min/day to 99 min/day, indicating high individual variability. Due to poor precision to capture sitting time on the individual level, it is less likely that GPAQ would be responsive to detect changes over two time points (as in intervention studies) when working on a group level. A substantially higher sample size when working on the population level will lead to somewhat higher precision. Thus, the GPAQ sitting item might potentially be able to detect population trends in sedentary behavior. However, a decreasing trend of a populational sedentary behavior followed by the (inter)national campaign aimed to raise awareness of health risks arising from being highly sedentary, might be contributed to change in social desirability error, rather than actual sedentary behavior. Using objective measures would overcome this issue.

Our results also show negative proportional bias, which turns out to be relative to the amount of sitting. In other words, highly sedentary participants were more likely to under-report their sedentary time absolutely, but not relatively. The finding of the negative proportional bias (when considering absolute values) is in line with Chastin et al. (2018), but in contrast with Cleland et al. (2014), who reported an opposite trend. The explanation for those discrepancies may derive from the selection of the method to construct the Bland–Altman plot. While in the study of Chastin and in our study, a criterion measure was regarded as the gold standard and Krouwer's method was applied. Clelands' study did not use a gold standard as a criterion measure, thus constructing the Bland–Altman as originally proposed.

SBQ (for weekdays)

In addition to total time, the context and type of sedentary behavior are also of interest in research and surveillance. Objective motion sensors are very limited in providing such information, so researchers mainly rely on the questionnaires, asking about the sedentary behavior in different domains or in different behaviors. Such composite questionnaires are difficult to validate. In most cases, the sum of behaviors (domains) is calculated for a validation of total time or comparison of each behavior with an objective total time (Bond et al., 2013; Munguia-Izquierdo et al., 2013; Rosenberg et al., 2010).

The SBQ has rarely been validated. According to Rosenberg et al. (2010), the reliability of the SBQ for overweight adults was acceptable for all items and the total score (ICC = 0.51 to 0.93), while the correlations with total sedentary time using ActiGraph were weak (r = 0.001 to 0.26). Bond et al. (2013) reported comparable estimates of total sedentary time between the SBQ and an objective physical activity monitor (SenseWear Armband) in overweight adults, but only at the group level. Agreement between the measures (for weekdays) at the individual level were poor (ICC = 0.22, p = .06) with the mean absolute difference between measures of 4.05 ± 2.94 hr/day. Our results also demonstrated a low level of agreement (ICC = 0.014) and no relationship (r = 0.018). Even when ranking the participants into tertiles and quartiles of sedentary time, the agreement was shown to be poor (in both cases, only 31% of the participants were correctly classified).

Chastin et al. (2018) reported that composite measures (sum of behaviors) are the only type of questionnaires (according to TASST framework) where total sedentary time is consistently overestimated. For such questionnaires with an unanchored recall period (like in SBQ), a correction factor of −219.8 min was proposed. However, the precision was shown to be very low (95% limits of agreement ranged from −725 to 286 min). Contrary to Chastin et al. (2018), we found overestimation of sedentary time using the SBQ only in 12% of the participants. Our results show a constant underestimation of the total sedentary time (for 3 hr on average). This discrepancy might be age-related since the study of Chastin included only older adults, who might more often perform different sedentary behaviors simultaneously (i.e. reading newspaper and listening to music, doing crafts and watching television), which could lead to overestimation when summing the behaviors from the SBQ. This was less likely to be the case in our study, where the average duration of sedentary behaviors, other than office work, driving and watching television, was low. However, this age-related assumption needs further investigation.

Underestimation of the sedentary time in our study may be partly attributed to the SBQ measurement scale, particularly the maximum category of ≥6 hr. For the purpose of the analysis, the category is recorded as 6 hr (Rosenberg et al., 2010), which consequently leads to underestimation. It is very likely that in a highly sedentary population like in ours, individuals spent substantially more than 6 hr in a single sedentary behavior. Almost 80% of our participants chose that answer at one point. Another consideration is related to rapid technological development, from which new sedentary behaviors arise, which are not covered by the SBQ.

Strengths and limitations

The main strength of our study was the use of a gold standard objective monitor for sedentary behavior (activPAL) as a criterion measure. The agreement analysis for the GPAQ sitting item and the SBQ was performed according to recommendations (Watson & Petrie, 2010). A person-oriented day approach deemed to be a behaviorally relevant approach (Stephens et al., 2014) was used. By applying a strict valid day criterion to the activPAL data, by which a valid day constitutes activPAL data from one wake time to the next day wake time, we leave no space for an underestimation of objective sedentary time which may result from applying a criterion of minimum wear time (Kang & Rowe, 2015).

A convenience worksite sample included only sedentary workers, more male than female, which makes the generalization of our results limited. Similarly, the results should be interpreted with some caution because of the relatively small sample size. Also, we adopted a criterion of at least three valid days, while longer duration would provide more reliable estimates of individual sedentary behavior (Edwardson et al., 2017). Another issue arises from the determination of wake time. We decided to adopt a combination of two common approaches – diary and criteria to classifying the first and the last standing event of the individuals' day. However, it has been reported that no waking wear identification method is universally acceptable and that the methods used should be reported in detail (Edwardson et al., 2017). Finally, due to feasibility limitations in our study, only weekdays were objectively assessed – consequently, a part of the SBQ about the weekend days was not a subject of evaluation.

Future research, policy, and practice

A consensus statement on the research priority of understanding and exploring the context of sedentary behavior has been endorsed recently (Dogra et al., 2017). This has strengthened the need for a high-quality validation of a composite questionnaire assessing sedentary behavior in different behaviors/domains. To date, a common practice in validation studies was that a single behavior/domain and their sum were correlated against the objective total sedentary time. Future studies should find a way to objectively measure sedentary behavior in each behavior/domain separately. That would rise the validation of composite questionnaires to a new level and provide valuable information for the development of new composite questionnaires or adaptations to the existing ones. A promising tool which could be used (solely or in combination with the activPAL) as a criterion measure in those studies might be a wearable camera (Kelly et al., 2014). New questionnaires should avoid using answers with unlimited scores. That would allow to perform a quality validation of the questionnaire and notably because the risk of being sedentary increases exponentially with each hour spent sedentary (Patterson et al., 2018).

Based on our observations, the level of the awareness that sedentary behavior is ubiquitous is still somehow low in our country. That might partly explain a significant underestimation of the self-reported sedentary time. We propose a study with the aim to test if an intervention of raising awareness about sedentary behavior would improve the accuracy and precision of self-reported sedentary time. We believe it might be good that the intervention (e.g. sending a flyer to the participants) is performed one week before completing the questionnaire, leaving the time for self-observation. If the results were promising, the proposed intervention could be applied even in large-scale studies, since additional burden and costs would be minimal.

In general, questionnaires are not a reliable tool for measuring sedentary behavior (Chastin et al., 2018), thus objective measurement should be the preferred choice when possible. When an accurate estimate of individual sedentary time is to be obtained, using highly reliable and valid tools (i.e. objective monitor, especially activPAL) would be the only option. However, if the group/population estimates are of interest, then questionnaire can also be used. The accuracy of the group/population mean can be improved by applying a correction factor, but the precision can only be improved by increasing the sample size. Thus, the precision of the questionnaire would be always higher when working on a population level as compared to the group level. This fact makes a questionnaire potentially useful to detect populational trends in sedentary time, while it is less likely that a questionnaire would detect changes on a group level. However, a low correlation with the objective measure indicates that using a questionnaire in epidemiological studies, where the relationship between sedentary behavior and health outcomes is of interest, will omit or overlook the significance of the exposure on health (Chastin et al., 2018).

Conclusions

To our knowledge, this is the first study providing agreement analysis between the SBQ and the activPAL and between the GPAQ sitting item and the activPAL in an active population. The GPAQ and the SBQ consistently underestimated the total sedentary time and agreed poorly with the current gold standard, both at the group and individual level. A correction factor should be considered in order to obtain more accurate population estimates. Due to low precision and low correlation against the reference measure (activPAL), the GPAQ and the SBQ are not a valid choice in epidemiological studies where correlations with health outcomes are of interest. It is also less likely that the GPAQ or SBQ would detect changes in sedentary behavior over time. Objective measurement should be the preferred choice when possible.

Key points

- When using GPAQ sitting item on a similar population like in our study, a correction factor of +33% was proposed for more accurate group/population estimates.
- The GPAQ sitting item and the SBQ are not recommended to be used in epidemiological studies that aimed to assess the relationships with health outcomes, nor in studies where changes in the behavior at the group level aimed to be observed.
- Due to measurement scale properties, the SBQ may be less suitable to use in a highly sedentary population.

Acknowledgments

The authors gratefully acknowledge the Horizon 2020 Framework Programme of the European Union; H2020 WIDESPREAD-2-Teaming: #739574 and the Republic of Slovenia for investment funding from the European Union of the European regional Development Fund. We also thank Mike Burnard for proofreading the manuscript and Bojan Leskošek for statistical support.

ORCID

Kaja Kastelic http://orcid.org/0000-0003-4570-8064
Nejc Šarabon http://orcid.org/0000-0003-0747-3735

References

Aguilar-Farias, N., & Leppe Zamora, J. (2017). Is a single question of the Global Physical Activity Questionnaire (GPAQ) valid for measuring sedentary behaviour in the Chilean population? *Journal of Sports Sciences*, *35*(16), 1652–1657. doi:10.1080/02640414.2016.1229010

Alkahtani, S. A. (2016). Convergent validity: Agreement between accelerometry and the global physical activity

questionnaire in college-age Saudi men. *BMC Research Notes*, *9*(1), 436. doi:10.1186/s13104-016-2242-9
Armstrong, T., & Bull, F. (2006). Development of the world health organization Global Physical Activity Questionnaire (GPAQ). *Journal of Public Health*, *14*(2), 66–70. doi:10.1007/s10389-006-0024-x
Barone Gibbs, B., & Kline, C. E. (2018). When does sedentary behavior become sleep? A proposed framework for classifying activity during sleep-wake transitions. *International Journal of Behavioral Nutrition and Physical Activity*, *15*(1), 81. doi:10.1186/s12966-018-0712-2
Bond, D. S., Thomas, J. G., Unick, J. L., Raynor, H. A., Vithiananthan, S., & Wing, R. R. (2013). Self-reported and objectively measured sedentary behavior in bariatric surgery candidates. *Surgery for Obesity and Related Diseases*, *9*(1), 123–128. doi:10.1016/j.soard.2012.09.008
Bull, F. C., Maslin, T. S., & Armstrong, T. (2009). Global Physical Activity Questionnaire (GPAQ): Nine Country reliability and validity study. *Journal of Physical Activity and Health*, *6*(6), 790–804. doi:10.1123/jpah.6.6.790
Chastin, S. F., Culhane, B., & Dall, P. M. (2014). Comparison of self-reported measure of sitting time (IPAQ) with objective measurement (activPAL). *Physiological Measurement*, *35*(11), 2319–2328. doi:10.1088/0967-3334/35/11/2319
Chastin, S. F. M., Dontje, M. L., Skelton, D. A., Cukic, I., Shaw, R. J., Gill, J. M. R., … Dall, P. M. (2018). Systematic comparative validation of self-report measures of sedentary time against an objective measure of postural sitting (activPAL). *International Journal of Behavioral Nutrition and Physical Activity*, *15*(1), 21. doi:10.1186/s12966-018-0652-x
Chu, A. H. Y., Ng, S. H. X., Koh, D., & Müller-Riemenschneider, F. (2018). Domain-specific Adult Sedentary Behaviour Questionnaire (ASBQ) and the GPAQ single-item question: A reliability and validity study in an Asian population. *International Journal of Environmental Research and Public Health*, *15*(4), 739. doi:10.3390/ijerph15040739
Cleland, C. L., Hunter, R. F., Kee, F., Cupples, M. E., Sallis, J. F., & Tully, M. A. (2014). Validity of the Global Physical Activity Questionnaire (GPAQ) in assessing levels and change in moderate-vigorous physical activity and sedentary behaviour. *BMC Public Health*, *14*(1), 1255. doi:10.1186/1471-2458-14-1255
Dall, P. M., Coulter, E. H., Fitzsimons, C. F., Skelton, D. A., & Chastin, S. (2017). TAxonomy of self-reported sedentary behaviour Tools (TASST) framework for development, comparison and evaluation of self-report tools: Content analysis and systematic review. *BMJ open*, *7*(4), e013844. doi:10.1136/bmjopen-2016-013844
Davies, S., Burns, H., Jewell, T., & Mcbride, M. (2011). *Start active, stay active: A report on physical activity from the four home countries*. London, UK: UK Department of Health Chief Medical Officers.
de Rezende, L. F., Rodrigues Lopes, M., Rey-Lopez, J. P., Matsudo, V. K., & Luiz Odo, C. (2014). Sedentary behavior and health outcomes: An overview of systematic reviews. *PloS one*, *9*(8), e105620. doi:10.1371/journal.pone.0105620
Dogra, S., Ashe, M. C., Biddle, S. J. H., Brown, W. J., Bauman, M. P., Chastin, S., … Copeland, J. L. (2017). Sedentary time in older men and women: An international consensus statement and research priorities. *British Journal of Sports Medicine*, *51*(21), 1526–1532. doi:10.1136/bjsports-2016-097209
Dowd, K. P., Harrington, D. M., Bourke, A. K., Nelson, J., & Donnelly, A. E. (2012). The measurement of sedentary patterns and behaviors using the activPAL Professional physical activity monitor. *Physiological Measurement*, *33*(11), 1887–1899. doi:10.1088/0967-3334/33/11/1887
Edwardson, C. L., Winkler, E. A. H., Bodicoat, D. H., Yates, T., Davies, M. J., Dunstan, D. W., & Healy, G. N. (2017). Considerations when using the activPAL monitor in field-based research with adult populations. *Journal of Sport and Health Science*, *6*(2), 162–178. doi:10.1016/j.jshs.2016.02.002
Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. doi:10.3758/bf03193146
Garber, C. E., Blissmer, B., Deschenes, M. R., Franklin, B. A., Lamonte, M. J., Lee, I. M., … Swain, D. P. (2011). American college of sports medicine position stand. Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: Guidance for prescribing exercise. *Medicine and Science in Sports and Exercise*, *43*(7), 1334–1359. doi:10.1249/MSS.0b013e318213fefb
Grant, P. M., Ryan, C. G., Tigbe, W. W., & Granat, M. H. (2006). The validation of a novel activity monitor in the measurement of posture and motion during everyday activities. *British Journal of Sports Medicine*, *40*(12), 992–997. doi:10.1136/bjsm.2006.030262
Herrmann, S. D., Heumann, K. J., Der Ananian, C. A., & Ainsworth, B. E. (2013). Validity and reliability of the Global Physical Activity Questionnaire (GPAQ). *Measurement in Physical Education and Exercise Science*, *17*(3), 221–235. doi:10.1080/1091367X.2013.805139
Hoos, T., Espinoza, N., Marshall, S., & Arredondo, E. M. (2012). Validity of the Global Physical Activity Questionnaire (GPAQ) in adult latinas. *Journal of Physical Activity and Health*, *9*(5), 698–705. doi:10.1123/jpah.9.5.698
Janssen, X., & Cliff, D. P. (2015). Issues related to measuring and interpreting objectively measured sedentary behavior data. *Measurement in Physical Education and Exercise Science*, *19*(3), 116–124. doi:10.1080/1091367X.2015.1045908
Kang, M., & Rowe, D. A. (2015). Issues and challenges in sedentary behavior measurement. *Measurement in Physical Education and Exercise Science*, *19*(3), 105–115. doi:10.1080/1091367X.2015.1055566
Kelly, P., Doherty, A., Mizdrak, A., Marshall, S., Kerr, J., Legge, A., … Foster, C. (2014). High group level validity but high random error of a self-report travel diary, as assessed by wearable cameras. *Journal of Transport & Health*, *1*(3), 190–201. doi:10.1016/j.jth.2014.04.003
Kelly, P., Fitzsimons, C., & Baker, G. (2016). Should we reframe how we think about physical activity and sedentary behaviour measurement? Validity and reliability reconsidered. *International Journal of Behavioral Nutrition and Physical Activity*, *13*(1), 32. doi:10.1186/s12966-016-0351-4
Kim, Y., Barry, V. W., & Kang, M. (2015). Validation of the ActiGraph GT3X and activPAL accelerometers for the

assessment of sedentary behavior. *Measurement in Physical Education and Exercise Science*, *19*(3), 125–137. doi:10.1080/1091367X.2015.1054390

Krouwer, J. S. (2008). Why Bland-Altman plots should use X, not (Y+X)/2 when X is a reference method. *Statistics in Medicine*, *27*(5), 778–780. doi:10.1002/sim.3086

Mumu, S. J., Ali, L., Barnett, A., & Merom, D. (2017). Validity of the global physical activity questionnaire (GPAQ) in Bangladesh. *BMC Public Health*, *17*(1), 650. doi:10.1186/s12889-017-4666-0

Munguia-Izquierdo, D., Segura-Jimenez, V., Camiletti-Moiron, D., Alvarez-Gallardo, I. C., Estevez-Lopez, F., Romero, A., … Delgado-Fernandez, M. (2013). Spanish adaptation and psychometric properties of the sedentary behaviour questionnaire for fibromyalgia patients: The al-Andalus study. *Clinical and Experimental Rheumatology*, *31*(6), S22–33.

Patterson, R., McNamara, E., Tainio, M., de Sa, T. H., Smith, A. D., Sharp, S. J., … Wijndaele, K. (2018). Sedentary behaviour and risk of all-cause, cardiovascular and cancer mortality, and incident type 2 diabetes: A systematic review and dose response meta-analysis. *European Journal of Epidemiology*, *33*(9), 811–829. doi:10.1007/s10654-018-0380-1

Pontt, J. L., Rowlands, A. V., & Dollman, J. (2015). Comparison of sedentary behaviours among rural men working in offices and on farms. *Australian Journal of Rural Health*, *23*(2), 74–79. doi:10.1111/ajr.12143

Prince, S. A., LeBlanc, A. G., Colley, R. C., & Saunders, T. J. (2017). Measurement of sedentary behaviour in population health surveys: A review and recommendations. *PeerJ*, *5*, e4130. doi:10.7717/peerj.4130

Rosenberg, D. E., Norman, J. G., Wagner, N., Patrick, K., Calfas, K. J., & Sallis, J. F. (2010). Reliability and validity of the Sedentary Behavior Questionnaire (SBQ) for adults. *Journal of Physical Activity and Health*, *7*(6), 697–705. doi:10.1123/jpah.7.6.697

Sellers, C., Dall, P., Grant, M., & Stansfield, B. (2016). Validity and reliability of the activPAL3 for measuring posture and stepping in adults and young people. *Gait & Posture*, *43*, 42–47. doi:10.1016/j.gaitpost.2015.10.020

Stephens, S. K., Winkler, E. A., Trost, S. G., Dunstan, D. W., Eakin, E. G., Chastin, S. F., & Healy, G. N. (2014). Intervening to reduce workplace sitting time: How and when do changes to sitting time occur? *British Journal of Sports Medicine*, *48*(13), 1037–1042. doi:10.1136/bjsports-2014-093524

Trinh, O. T. H., Do Nguyen, N., van der Ploeg, H. P., Dibley, M. J., & Bauman, A. (2009). Test-retest repeatability and relative validity of the global physical activity questionnaire in a developing country context. *Journal of Physical Activity and Health*, *6*(s1), S46–S53. doi:10.1123/jpah.6.s1

Wanner, M., Hartmann, C., Pestoni, G., Martin, B. W., Siegrist, M., & Martin-Diener, E. (2017). Validation of the Global Physical Activity Questionnaire for self-administration in a European context. *BMJ Open Sport & Exercise Medicine*, *3*(1), e000206. doi:10.1136/bmjsem-2016-000206

Watson, P. F., & Petrie, A. (2010). Method agreement analysis: A review of correct methodology. *Theriogenology*, *73*(9), 1167–1179. doi:10.1016/j.theriogenology.2010.01.003