Accepted author's manuscript. Maturitas.
Publisher DOI: https://doi.org/10.1016/j.maturitas.2019.08.004

Comparison of the Physical Activity Frequency Questionnaire (PAFQ) with accelerometry in middle-aged and elderly: the CoLaus study

Sanne Verhoog[1]; Cédric Gubelmann[2]; Idris Guessous[3]; Arjola Bano[1,4]; Oscar H. Franco [1] and Pedro Marques-Vidal[2]

1 Institute of Social and Preventive Medicine, University of Bern, Switzerland
2 Department of Medicine, Internal Medicine, Lausanne University Hospital (CHUV), Lausanne, Switzerland
3 Division of primary care medicine, Department of primary care medicine, Geneva university hospitals, Geneva, Switzerland
4 Department of Cardiology, Inselspital, Bern University Hospital, Bern, Switzerland

Authors' emails:

Sanne Verhoog: sanne.verhoog@ispm.unibe.ch
Cédric Gubelmann: cedric.gubelmann@gmail.com
Idris Guessous: idris.guessous@hcuge.ch
Arjola Bano: arjola.bano@ispm.unibe.ch
Oscar H. Franco: oscar.franco@ispm.unibe.ch
Pedro Marques-Vidal: pedro-manuel.marques-vidal@chuv.ch

Address for correspondence and reprints

Sanne Verhoog
Mittelstrasse 43
3012 Bern, Switzerland
Phone: +41 31 631 56 75
Email: sanne.verhoog@ispm.unibe.ch

Funding:
The authors report no conflict of interest.

Word count: 6811 with references and tables, 4472 without references and tables

Number of tables: 3 **Figures**: 2 **References**: 54

Supplementary material: 4 tables, 2 additional files

ABSTRACT

Objective: The Physical Activity Frequency Questionnaire (PAFQ) has been used in several studies, but its validation dates from 1998. We aimed to compare the PAFQ with accelerometry data for measuring activity.

Design: Cross-sectional analysis conducted among individuals aged 35-75 years within the prospective CoLaus study in Switzerland. All participants were requested to complete the PAFQ and wear a wrist-worn accelerometer for 14 consecutive days. Spearman correlation, Lin’s concordance coefficient and Bland-Altman plots were performed to compare PAFQ and accelerometry data.

Results: 1752 participants (50.7% female, age range 45.2-87.1 years) had complete information on the PAFQ and accelerometer. Compared to the accelerometer, the PAFQ overestimated total, light, moderate and vigorous activity by a median [interquartile range] of 143 [34.5; 249], 72 [12; 141], 23 [-46; 100] and 13 [-1; 41] minutes/day, respectively, and underestimated sedentary behaviour by 123 [14; 238] minutes/day. Spearman’s correlation coefficients ranged from 0.171 for vigorous PA and 0.387 for total PA and sedentary behaviour, and Lin’s concordance coefficients ranged from 0.044 for vigorous PA and 0.254 for moderate-to-vigorous PA, indicating a poor correlation and concordance. The difference between PAFQ and accelerometer increased with increasing time spent in each activity level.

Conclusion: There is limited agreement between the measures of activity obtained by PAFQ compared to accelerometers, suggesting that these tools measure activity through different angles. Although there is some degree of comparability, they should be considered as complementary tools to obtain comprehensive information of individual and population activity levels.

Keywords: physical activity; sedentary behaviour; questionnaire; accelerometry; comparison; Switzerland.

Abstract word count: 219

Publisher DOI: https://doi.org/10.1016/j.maturitas.2019.08.004

INTRODUCTION

Physical activity (PA) is an important modifiable risk factor in the prevention of non-communicable diseases, including cardiovascular disease, diabetes mellitus, several types of cancer and depression [1, 2]. Several methods exist for assessing PA and sedentary behaviour (SB) [3]. Self-reported measures (e.g. questionnaires) have been widely used in epidemiological studies because of their low cost and low burden to participants [4]. Recently, accelerometers have become more accessible and have allowed measuring the intensity and duration of PA objectively in large samples [5-7]. However, discrepancies emerged when comparing PA data collected by questionnaire and by accelerometry in large populations [8, 9]. Correlations are low, and self-reported measures either over- or underestimate PA relative to accelerometry. Also, SB has been recently suggested to impact cardiovascular disease independent of PA [10-13]. Still, few PA questionnaires have been validated or compared with accelerometers for the assessment of SB [14-16].

The physical activity frequency questionnaire (PAFQ) was developed in 1998 to measure total and activity-specific energy expenditure in the population of Geneva, Switzerland [17]. The PAFQ focuses mainly on light and moderate PA and has been used in several studies to quantify physical activity [18-23], including studies as CoLaus [24] and Bus Santé [25]. Validation of the PAFQ was performed in 41 volunteers using a heart rate monitor in 1998 [17], which was captured in a large systematic review comparing direct and self-reported measures of activity [9]. The validation study reported a good correlation ($r = 0.76$) between total energy estimated by the PAFQ and by a heart rate monitor, and did not find statistically significant differences between estimates from the PAFQ and the heart rate monitor [17]. Since then, no further validations have been performed and it has not been compared with alternative methods such as accelerometry. In the CoLaus study, accelerometers were recently added as an additional tool to measure PA. How the PAFQ compares to PA measured with accelerometers, and whether these two methods are interchangeable or not remains unclear.

The assessment of physical activity can be influenced by an individuals' characteristics, as well as their geographical and cultural context. The reliability and validity of self-reported measures is limited by recall bias or social desirability [26]. In particular, light intensity physical activities (LPA) such as household chores, are harder to recall than activities of higher intensity [26]. Recall bias and reporting errors may also occur more often in older populations, in whom cognitive impairment is more likely [27]. Social desirability is another factor that might influence physical activity measurement, especially for women and individuals with a higher body mass index (BMI) [28]. In addition, individuals in retiring age (i.e. 65 years in Switzerland) have a different activity behaviour than working individuals.

Therefore, in a large population of middle-aged and older adults, we aimed to compare activity measured with the PAFQ to activity measured by accelerometry, and to investigate whether these comparisons differed by categories of gender, age and BMI.

METHODS

Participants

The detailed process of the recruitment of the CoLaus study and the follow-up procedures have been described previously [24]. Briefly, the CoLaus study is a population-based cohort study exploring the biological, genetic, and environmental determinants of cardiovascular disease. A non-stratified, representative sample of the population of Lausanne (Switzerland) was recruited between 2003 and 2006 based on the following inclusion criteria: i) age 35–75 years and ii) willingness to participate. The second follow-up occurred 10.9 years after the baseline survey and included an optional module assessing the participant's PA for 14 days with an accelerometer. For this study, we performed a cross-sectional analysis using data of the second follow-up only.

Accepted author’s manuscript. Maturitas.
Publisher DOI: https://doi.org/10.1016/j.maturitas.2019.08.004

Exclusion criteria

Participants were excluded if they i) did not fill in the PAFQ, ii) did not wear the accelerometry, or iii) had an insufficient number of valid days for accelerometry (less than 5 weekdays or 2 weekend days).

Measurements

Physical activity frequency questionnaire

The PAFQ is a self-administered measure of the total and activity specific energy expenditure. Detailed information on the PAFQ can be found elsewhere [17, 29]. Briefly, the questionnaire lists 70 types of PA from various domains (e.g. occupational, housework, leisure time, sports, etc.). The participants indicated the number of days in the past week (0 to 7) and the duration per day (0 to 10h, 15 min increment) for each activity.

For the purpose of this study, each type of activity was categorized into SB (<2 metabolic equivalent of tasks – METs), LPA (2 to <3 METs), MPA (3-6 METs) and VPA (>6 METs) according to the compendium of physical activities[30, 31] (see **Additional file 1)**. Total PA was defined as the sum of LPA, MPA and VPA. For each item of the PAFQ, the time spent per week was computed as average hours per day×number of days performing the activity. For each category of item (i.e. corresponding to SB, LPA, MPA or VPA), the times were summed up and divided by 7 to estimate an average daily time. For example, if the participant spent 2 hours/day housekeeping (MPA) and performed this activity for 3 days per week, then the total time was 2×3=6 hours/week; if the participant also spent 1 hour/day sewing and ironing (MPA) and performed this activity every day, then the total time was 1×7=7 hours/week, and the average daily time spent in MPA activities would be (6+7)/7=1.85 hours or 111 minutes.

Usual sleep time (in minutes) was assessed by asking the participants when they went to bed and when they woke up, and the number of minutes on non-sleep time (NST) was computed as

NST=1440–sleep time. As participants tended to under or overestimate time spent in the different activities, a standardization was performed as follows: First, we calculated *T*, the total amount of time spent in SB, LPA; MPA and VPA activities. Second, we computed the percentage of time dedicated to each type of activity, i.e. P_{SB}=time spent in SB/T. Third, we computed the standardized time spent on each type of activity by multiplying the non-sleep time by the percentage of time spent in each activity $ST_{SB}=T\times P_{SB}$. A detailed example of calculation is provided in **Additional file 2**.

Accelerometer Physical Activity

Accelerometry-based PA was assessed using a wrist-worn triaxial accelerometer (*GENEActive*, Activinsights Ltd, UK). This device has been validated against reference methods [32]. The intra- and inter-instrument coefficients of variation were 1.4% and 2.1%; and high correlations with reference methods such as mechanical shaking (r=0.98) and indirect calorimetry (r=0.83) have been reported [32]. The accelerometers were pre-programmed with a 50 Hz sampling frequency and subsequently attached to the participants' right wrist. Participants were requested to wear the device continuously for 14 days in their free-living conditions. Accelerometry data were downloaded using the GENEActiv software version 2.9 (GENEActiv, Activinsights Ltd., United Kingdom) and collapsed into 60 s epoch files.

Data were analysed using the GENEActiv macro file "General physical activity" version 1.9 based on intensity cutoffs validated among middle-aged adults: SB (<241 g.min), light (241–338 g.min), moderate (339-1131 g.min) and vigorous (≥1132 g.min) PA.[32] Total PA was defined as the sum of light, moderate and vigorous PA. Conversely, no information was available regarding the criteria used for non-wear time (proprietary). Based upon a previous study [33], a valid day was defined as ≥10 hours (i.e. 600 min) and ≥8 hour (i.e. 480 min) of diurnal wear-time on weekdays and weekend days, respectively. For each participant, the proportion of time (in percentage) spent in SB, LPA, MPA and VPA was averaged for all valid days.

Accepted author's manuscript. Maturitas.
Publisher DOI: https://doi.org/10.1016/j.maturitas.2019.08.004

Other covariates

Gender was self-reported. Age at the time of the examination was rounded to the nearest year and further categorized into two groups: <65 and ≥65 years, to reflect the working and retired individuals. Body weight and height were measured with participants barefoot and in light indoor clothes. Body weight was measured in kilograms to the nearest 100 g using a Seca® scale (Hamburg, Germany). Height was measured to the nearest 5 mm using a Seca® (Hamburg, Germany) height gauge. Body mass index (BMI) was calculated by weight(kg)/height(m)^2, and categorized into low (<25 kg/m^2) and high (≥25 kg/m^2) BMI.

Statistical analysis

Statistical analyses were performed using Stata version 15.1 for Windows (Stata Corp, College Station, Texas, USA). Descriptive results were expressed as number of participants (percentage) for categorical data, and average ± standard deviation (SD), or median [interquartile range (IQR)] for continuous data. Lin's concordance correlation coefficients and corresponding 95% confidence intervals (CI) were used to measure the agreement between PAFQ and accelerometer data [34]. Interpretation of Lin's concordance coefficients was as follows: <0.90 poor; 0.90 to 0.95 moderate; 0.95 to 0.99 substantial; >0.99 almost perfect [35]. Spearman correlations were used to associate PAFQ and accelerometer data; 95% CIs were obtained by bootstrapping with replacement, using 1000 iterations and bias-corrected values. Bland–Altman plots were used to visualize the extent of (dis)agreement between the two measures. Interpretation of Spearman correlation was as follows: 0.0 to 0.3 negligible; 0.3 to 0.5 low; 0.5 to 0.7 moderate; 0.7 to 0.9 high; 0.9 to 1.0 very high correlation [36].

As a sensitivity analysis, we excluded participants with extreme sleep durations (<4 or >10 hours/day).

Simple linear regressions were performed to assess whether differences between the PAFQ and accelerometer regarding SB or PA were associated with gender (men vs. women), age (<65 vs. ≥65

Publisher DOI: https://doi.org/10.1016/j.maturitas.2019.08.004

years) and BMI (low vs. high) groups, for each group separately. Groups were coded as dichotomous (0/1) variables, whereby the categories of women, age <65 years and low BMI were considered as reference.

$$\Delta = \alpha + \beta . group$$

Where Δ is the difference between the PAFQ and the accelerometer regarding time spent in SB and the different levels of PA. A statistically significant value of $\alpha<0$ indicates overestimation by the accelerometer, and underestimation otherwise. A statistically significant value of β indicates that Δ differs between groups. In addition, multivariable linear regressions were performed including gender, age and BMI groups as independent variables, allowing adjustment for each other. Statistical significance was considered for a two-sided test with $p<0.05$.

RESULTS

Characteristics of participants

Of the initial 4881 participants, 1752 (35.9%) were eligible for the analysis. The selection procedure is shown in **Figure 1** and the characteristics of included and excluded participants are presented in **Supplementary table 1**. Included participants were younger, less likely to be female and had a lower BMI. The mean age (SD) of the included participants was 60.5 (9.4) and 50.7% of them were female.

Activity levels, correlations and concordance

Time spent in SB and different levels of PA (in minutes per day or as percentage of time) according to the PAFQ and accelerometer are presented in **Table 1** as median [interquartile range]. Compared to the accelerometer, the PAFQ overestimated total PA, LPA, MPA and VPA by 143 [34.5; 249], 72 [12; 141], 23 [-46; 100] and 13 [-1; 41] minutes, respectively, and underestimated SB by 123 [14; 138] minutes. When the results were expressed as percentage of time, the PAFQ overestimated

LPA, MPA and VPA by 7.5% [1.1; 14.6], 1.7% [-5.3; 10.1] and 1.3% [-0.1; 4.2], respectively, and underestimated SB by 14.7% [4.1; 25.7].

Spearman correlation and Lin's concordance coefficients of SB and different levels of PA between the PAFQ and the accelerometer are presented in **Table 2**. When the results were expressed in minutes, Spearman coefficients (95% CI) ranged between 0.171 (0.124; 0.221) for VPA and 0.370 (0.325; 0.409) for SB, indicating a neglectible to low correlation, whereas total PA had a correlation of 0.373 (0.331; 0.417), indicating a low correlation. Lin's concordance coefficients ranged between 0.044 (0.035; 0.053) for VPA and 0.250 (0.218; 0.282) for MVPA, indicating a poor concordance. When the results were expressed as percentage of time, Spearman coefficients ranged between 0.171 (0.126; 0.217) for VPA and 0.387 (0.346; 0.426) for SB, indicating a neglectible to low correlation. Lin's coefficients ranged between 0.045 (0.036; 0.055) for VPA and 0.254 (0.221; 0.286) for MVPA, indicating a poor concordance. All observed Spearman correlations and Lin concordance coefficients were statistically significant at $p<0.001$.

The Bland-Altman plots illustrating the absolute differences in time estimated with the PAFQ and the accelerometer are presented in **Figure 2**. Both for absolute time and proportion of time, there were positive mean differences for LPA, MPA, VPA and MVPA, indicating that the PAFQ overestimates these behaviours compared to the accelerometer, and a negative mean difference for SB indicating that SB is underestimated in the PAFQ compared to accelerometery. The limits of agreement were wide for all activity levels, meaning that there are large differences in estimated time spent in all activity levels comparing accelerometer and PAFQ. An increased discrepancy between accelerometer and PAFQ was seen at more minutes of time spent in LPA, MPA; VPA and MVPA, while no clear trend was visible for SB.

Sensitivity analysis

Results after excluding 40 participants with extreme sleeping times (<4 or >10 hours/day) are presented in **Supplementary tables 2 and 3**. Overall, no changes in the differences between PAFQ and accelerometry (**Supplementary table 2**) and in the Spearman correlation and Lin's concordance coefficients (**Supplementary table 3**) were observed.

Differences between gender, age and BMI groups

Results according to gender, age and BMI groups are presented in **Table 3**. Compared to accelerometry data, both men and women underestimated SB and MPA, and overestimated LPA, VPA and total PA when using the PAFQ. Women overestimated LPA and underestimated MPA to a larger extent than men, while no differences were found for total PA, VPA and SB.

Both age groups (<65 and ≥65 years) underestimated SB and MPA, and overestimated LPA, VPA and total PA when using the PAFQ. Participants aged <65 years underestimated MPA and overestimated LPA and VPA to a larger extent than participants aged ≥65 years. Participants aged ≥65 years underestimated SB to a larger extent than participants aged <65 years, while no differences were found for total PA.

Both BMI groups (low and high) underestimated SB and MPA, and overestimated LPA, VPA and total PA when using the PAFQ. No differences in over- or underestimation were found between the groups.

Multivariate linear regression including gender, age and BMI groups showed that the high BMI group overestimated LPA to a larger extent than low BMI group, while no differences were found for total PA, MPA, VPA and SB (**Supplementary table 4**).

DISCUSSION

In this study comparing levels of PA measured with the PAFQ to activity measured with accelerometers, we observed that the PAFQ overestimated all PA levels while underestimating SB. Our findings also showed moderate overall correlations and concordance coefficients between the PAFQ and accelerometry data for total PA and SB. Differences in over- or underestimation were found between gender for LPA, MPA and MVPA, while between the age groups differences in over- or

Accepted author's manuscript. Maturitas.
Publisher DOI: https://doi.org/10.1016/j.maturitas.2019.08.004

underestimation were found for all levels of activity, except total PA. No differences in over- or underestimation were found between the BMI groups.

Comparison with previous literature

A systematic review identified 100 different PA questionnaires with their respective validity against objective measures [37]. For PA, median correlation coefficients (0.25 to 0.41) were comparable to our findings (0.171 to 0.387). For SB, a median Spearman correlation of 0.12 was reported across all questionnaires between reported SB and inactivity calculated from objective measures, which was lower than in our study (0.370 to 0.387). They also found that newly developed physical activity questionnaires do not seem to perform better than existing ones. An explanation for the higher correlation of SB in our study might be that the PAFQ was developed with a focus on low intensity activities which otherwise may have been overlooked as important contributors to energy output [17]. At the development of the questionnaire, about 50 percent of the total energy expenditure was spent sleeping, siting quitly, eating, and office work was the most important work related activity [17]. In addition, the GENEActiv accelerometer is shown to have a good validity for measuring adult sedentary time. A study comparing the GENEActiv accelerometer with an activPAL device estimating total sedentary time during waking hours reported an Intraclass Correlation Coefficient of 0.80 (95% CI: 0.68-0.88) [38].

The IPAQ (International Physical Activity Questionnaire) has been widely used to assess PA. A recent review summarized 23 validation studies of the short form of the IPAQ (IPAQ-SF) and showed poor correlations when compared with objective measurements [39]. The correlations ranged from 0.09 to 0.39 for total PA, which is comparable to the findings in our study (0.373 to 0.387). Additionally, in agreement with our study, participants tended to report more PA with the IPAQ-SF than as assessed by the accelerometer. Over-reporting of PA remains a key limitation of most self-reported measures of PA [28] and is often due to misreporting of frequency, intensity and/or duration of activities [40].

Factors contributing to misreporting are social desirability and recall bias, which can be influenced by population characteristics such as gender, age and health status [41, 42].

In our study, results from the Bland-Altman plots showed that the differences between the PAFQ and accelerometry data increased with increasing time spent in all types of PA, similar to what has been seen in other studies [43, 44]. This suggests that the more time spent in each activity level, the higher the under- or overestimation. Therefore, the PAFQ might be a good estimator of activity for individuals with average time spent in each activity level, while for individuals that spent more time in a specific level of PA the PAFQ might lead to considerable overestimation compared to accelerometers.

Results from our study showed that the differences between PAFQ and accelerometer data for SB, VPA and total PA were generally similar for men and women. However, women overestimated LPA and underestimated MPA to a larger extent than men. Women tend to engage more in light and moderate PA, which are the most challenging types of activity to recall because they are most dominant in daily life as, for example, in household activities [45]. In our study, participants aged ≥65 years underestimated SB to a larger extent than participants aged <65 years. This is similar to results from another study, comparing the IPAQ with accelerometer measurements among 1751 adults [44]. Possible explanations is that elderly people are more likely to engage in activities that are most inaccurately assessed by questionnaires [46]. No differences in under- or overestimation were found between BMI groups, which is similar to findings in other studies [44, 47]. This is important as it has been suggested that obese people tend to overestimate PA due to social desirability [48, 49]. In summary the differences between PAFQ and accelerometer are similar in both genders and BMI groups for SB and total PA, and in the age groups for total PA only. Conversely, the differences between PAFQ and accelerometer in both genders and age groups are different for specific levels of PA, which might lead to differential estimations according to individual characteristics.

The PAFQ was developed in Geneva, Switzerland on the basis of a population survey using an approach similar to that proposed by Block et al. [50] for food frequency questionnaires [17]. The data

Publisher DOI: https://doi.org/10.1016/j.maturitas.2019.08.004

obtained from 24-hour recalls in the target population were used to establish a list of the major contributors in order to explain 95 percent of the engery expended by the total sample and by the male and female subsamples [17]. Therefore, activities that are typically related to Switzerland (e.g. Alpine skiing, stacking wood or angling) were captured in the questionnaire. The differences that were found between the questionnaire and the accelerometer could not be the result of typical Swiss activities that were missing in the questionnaire.

Implications for public health

The findings of our study may have implications for current and future practice on the measurement of PA. The PAFQ showed similar correlations and concordance with accelerometer measurements as found in similar studies validating other PA questionnaires. These findings suggest that the PAFQ is as valid as other questionnaires in SB and PA estimation. However, the Bland-Altman plots showed that the longer the time spent in a PA level, the higher the overestimation in the PAFQ compared to accelerometer. Hence, active people as assessed by PAFQ might be less active than they claim, and the associations between PA and health as obtained by questionnaires might actually be valid to lesser time spent in PA levels than thought, due to the overestimation of PA in the active individuals. In addition, under- or overestimation appears to differ between genders and age groups for different levels of PA, but not between BMI groups. Results from questionnaires might need either stratification, or adjustment on gender and age.

Questionnaires and accelerometers measure physical activity through different perspectives. The accelerometer is increasingly used in current research and designed to objectively measure the intensity and duration of activity and their results are comparable irrespective of the device used [51]. However, they also have limitations. Depending on the attachment site of the hip or the wrist, they measure mainly lower or upper body movement, respectively [7]. Accelerometers also do not capture information on body posture (i.e. sitting or standing still) or whether a person is carrying weight. Moreover, they cannot capture the context in which the activities take place and they cannot distinguish between types of activities. The questionnaire, on the other hand, is reliable on an individuals's answering abitlity but able to capture the context and type of activity, which is a major strength of physical activity questionnaires. Considering their advantages and disadvantages, questionnaires and accelerometers should be considered as complementary tools rather than a replacement for one another. Whenever possible, future studies should use both measurement methods in order to obtain comprehensive information of individual and population SB and PA levels.

Strengths and limitations

One of the major strengths of our study is that we took into account the differences between gender, age and BMI when comparing activity measured by the PAFQ and accelerometer. In addition, we compared SB, four different levels of PA and total PA. Another strength is our large sample size. Most validation studies summarized in the review by Helmerhorst et al [37] included fewer than 300 participants, while our study comprised 1752 participants. Moreover, the use of a waterproof wrist-worn accelerometer ensured high compliance. Participants were requested to wear the accelerometer for 14 consecutive days, which far exceeds the 3-5 days required to assess a daily estimate of the individual's habitual activity [52]. Finally, the sensitivity analyses whereby participants with extreme sleep durations were excluded provided consistent findings.

The findings of this study must be considered in the context of some limitations. First, the SB and PA estimates from PAFQ and accelerometer were not obtained at the same period. Participants filled in the questionnaire and brought it to the study centre, where they were given the accelerometer for 14 consecutive days. Nevertheless, it has been reported that physical activity at middle age tend to be relatively constant overtime [53]. Therefore the effect might not be of a large magnitude nor would it be expected to differ by levels of PA. Second, included participants were younger than excluded participants, which might affect the generalizability among the elderly. Furthermore, the findings of this study are based on the difference between the PAFQ and the GENEActiv accelerometer. As such, our findings might not be generalizable to other questionnaires, since other questionnaires measure

Publisher DOI: https://doi.org/10.1016/j.maturitas.2019.08.004

PA in a different way. Finally, it is reported that the GENEActiv accelerometer tends to overestimate PA and underestimate SB [54]. Therefore, the differences observed in this study might actually be underestimated.

Conclusion

The present study shows large variations between the PAFQ and accelerometer measured PA and SB, and the difference increased with more time spent in PA. Sex and age, but not BMI, influenced these variations. Questionnaires and accelerometers measure activity through different perspectives. Although there is a certain degree of comparability, they should be considered as complementary tools to obtain comprehensive information of individual and population physical activity levels.

DECLARATIONS

Ethicals approval and consent to participate

The institutional Ethics Committee of the University of Lausanne, which afterwards became the Ethics Commission of Canton Vaud (www.cer-vd.ch) approved the baseline CoLaus study (reference 16/03, decisions of 13th January and 10th February 2003). The approval was renewed for the first (reference 33/09, decision of 23rd February 2009) and the second (reference 26/14, decision of 11th March 2014) follow-ups. The full decisions of the CER-VD can be obtained from the authors upon request. The study was performed in agreement with the Helsinki declaration and its former amendments, and in accordance with the applicable Swiss legislation. All participants gave their signed informed consent before entering the study.

Availability of data and material

Due to the sensitivity of the data and the lack of consent for online posting, individual data cannot be made accessible. Only metadata will be made available in digital repositories. Metadata requests can also be performed via the study website www.colaus-psycolaus.ch

Competing interest

The authors report no conflict of interest.

Funding

The CoLaus study was and is supported by research grants from GlaxoSmithKline, the Faculty of Biology and Medicine of Lausanne, and the Swiss National Science Foundation [grant numbers 33CSCO-122661, 33CS30-139468 and 33CS30-148401].

Authors contributions

SV made part of the statistical analyses and wrote most of the article; CG collected; PMV made part of the statistical analysis and wrote part of the article; OHF wrote part of the article; CG, PMV, OHF, IG and AB revised the article for important intellectual content. PMV had full access to the data and is the guarantor of the study.

REFERENCES

[1] F.W. Booth, C.K. Roberts, M.J. Laye, Lack of exercise is a major cause of chronic diseases, Comprehensive Physiology 2(2) (2012) 1143-1211.
[2] World Health Organisation, Global recommendations on physical activity for health, Geneva; 2010.
[3] D. Ndahimana, E.K. Kim, Measurement Methods for Physical Activity and Energy Expenditure: a Review, Clin Nutr Res 6(2) (2017) 68-80.
[4] L.G. Sylvia, E.E. Bernstein, J.L. Hubbard, L. Keating, E.J. Anderson, A Practical Guide to Measuring Physical Activity, Journal of the Academy of Nutrition and Dietetics 114(2) (2014) 199-208.
[5] S. Skender, J. Ose, J. Chang-Claude, M. Paskow, B. Brühmann, E.M. Siegel, K. Steindorf, C.M. Ulrich, Accelerometry and physical activity questionnaires - a systematic review, BMC Public Health 16 (2016) 515.
[6] J.E. Sasaki, K.S. da Silva, B. Gonçalves Galdino da Costa, D. John, Chapter 2 - Measurement of Physical Activity Using Accelerometers, in: J.K. Luiselli, A.J. Fischer (Eds.), Computer-Assisted and Web-Based Innovations in Psychology, Special Education, and Health, Academic Press, San Diego, 2016, pp. 33-60.
[7] I.M. Lee, E.J. Shiroma, Using Accelerometers to Measure Physical Activity in Large-Scale Epidemiologic Studies: Issues and Challenges, British journal of sports medicine 48(3) (2014) 197-201.
[8] K. Kowalski, R. Rhodes, P.-J. Naylor, H. Tuokko, S. MacDonald, Direct and indirect measurement of physical activity in older adults: a systematic review of the literature, The international journal of behavioral nutrition and physical activity 9 (2012) 148-148.
[9] S.A. Prince, K.B. Adamo, M.E. Hamel, J. Hardt, S. Connor Gorber, M. Tremblay, A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review, The international journal of behavioral nutrition and physical activity 5 (2008) 56.
[10] M.P. Buman, E.A. Winkler, J.M. Kurka, E.B. Hekler, C.M. Baldwin, N. Owen, B.E. Ainsworth, G.N. Healy, P.A. Gardiner, Reallocating time to sleep, sedentary behaviors, or active behaviors: associations with cardiovascular disease risk biomarkers, NHANES 2005-2006, American journal of epidemiology 179(3) (2014) 323-34.
[11] G.N. Healy, C.E. Matthews, D.W. Dunstan, E.A. Winkler, N. Owen, Sedentary time and cardio-metabolic biomarkers in US adults: NHANES 2003-06, European heart journal 32(5) (2011) 590-7.
[12] G.N. Healy, K. Wijndaele, D.W. Dunstan, J.E. Shaw, J. Salmon, P.Z. Zimmet, N. Owen, Objectively measured sedentary time, physical activity, and metabolic risk: the Australian Diabetes, Obesity and Lifestyle Study (AusDiab), Diabetes care 31(2) (2008) 369-71.
[13] Q. Qi, G. Strizich, G. Merchant, D. Sotres-Alvarez, C. Buelna, S.F. Castaneda, L.C. Gallo, J. Cai, M.D. Gellman, C.R. Isasi, A.E. Moncrieft, L. Sanchez-Johnsen, N. Schneiderman, R.C. Kaplan, Objectively Measured Sedentary Time and Cardiometabolic Biomarkers in US Hispanic/Latino Adults: The Hispanic Community Health Study/Study of Latinos (HCHS/SOL), Circulation 132(16) (2015) 1560-9.
[14] S. Scholes, N. Coombs, Z. Pedisic, J.S. Mindell, A. Bauman, A.V. Rowlands, E. Stamatakis, Age- and sex-specific criterion validity of the health survey for England Physical Activity and Sedentary Behavior Assessment Questionnaire as compared with accelerometry, American journal of epidemiology 179(12) (2014) 1493-502.
[15] F.A. Barwais, T.F. Cuddihy, T. Washington, L.M. Tomson, E. Brymer, Development and validation of a new self-report instrument for measuring sedentary behaviors and light-intensity physical activity in adults, Journal of physical activity & health 11(6) (2014) 1097-104.
[16] S.R. Gomersall, T.G. Pavey, B.K. Clark, A. Jasman, W.J. Brown, Validity of a Self-Report Recall Tool for Estimating Sedentary Behavior in Adults, Journal of physical activity & health 12(11) (2015) 1485-91.
[17] M. Bernstein, D. Sloutskis, S. Kumanyika, A. Sparti, Y. Schutz, A. Morabia, Data-based approach for developing a physical activity frequency questionnaire, American journal of epidemiology 147(2) (1998) 147-54.

[18] M.C. Costanza, S. Beer-Borst, A. Morabia, Achieving energy balance at the population level through increases in physical activity, Am J Public Health 97(3) (2007) 520-525.
[19] Y. Henchoz, F. Bastardot, I. Guessous, J.-M. Theler, J. Dudler, P. Vollenweider, A. So, Physical activity and energy expenditure in rheumatoid arthritis patients and matched controls, Rheumatology 51(8) (2012) 1500-1507.
[20] P. Marques-Vidal, G. Waeber, P. Vollenweider, I. Guessous, Socio-demographic and lifestyle determinants of dietary patterns in French-speaking Switzerland, 2009-2012, BMC Public Health 18(1) (2018) 131.
[21] M. Aligisakis, P. Marques-Vidal, I. Guessous, P. Vollenweider, Did Dumbo suffer a heart attack? independent association between earlobe crease and cardiovascular disease, BMC cardiovascular disorders 16 (2016) 17-17.
[22] S. Khalatbari-Soltani, F. Imamura, S. Brage, E. De Lucia Rolfe, S.J. Griffin, N.J. Wareham, P. Marques-Vidal, N.G. Forouhi, The association between adherence to the Mediterranean diet and hepatic steatosis: cross-sectional analysis of two independent studies, the UK Fenland Study and the Swiss CoLaus Study, BMC Med 17(1) (2019) 19-19.
[23] I. Guessous, M. Bochud, J.-M. Theler, J.-M. Gaspoz, A. Pechère-Bertschi, 1999-2009 Trends in prevalence, unawareness, treatment and control of hypertension in Geneva, Switzerland, PloS one 7(6) (2012) e39877-e39877.
[24] M. Firmann, V. Mayor, P.M. Vidal, M. Bochud, A. Pecoud, D. Hayoz, F. Paccaud, M. Preisig, K.S. Song, X. Yuan, T.M. Danoff, H.A. Stirnadel, D. Waterworth, V. Mooser, G. Waeber, P. Vollenweider, The CoLaus study: a population-based study to investigate the epidemiology and genetic determinants of cardiovascular risk factors and metabolic syndrome, BMC cardiovascular disorders 8 (2008) 6.
[25] A. Morabia, M. Bernstein, S. Heritier, A. Ylli, Community-based surveillance of cardiovascular risk factors in Geneva: methods, resulting distributions, and comparisons with other populations, Prev Med 26(3) (1997) 311-9.
[26] R. Shephard, A. Vuillemin, Limits to the measurement of habitual physical activity by questionnaires, British journal of sports medicine 37(3) (2003) 197-206.
[27] G. Rait, A. Fletcher, L. Smeeth, C. Brayne, S. Stirling, M. Nunes, E. Breeze, E.S. Ng, C.J. Bulpitt, D. Jones, A.J. Tulloch, Prevalence of cognitive impairment: results from the MRC trial of assessment and management of older people in the community, Age and ageing 34(3) (2005) 242-8.
[28] J.F. Sallis, B.E. Saelens, Assessment of physical activity by self-report: status, limitations, and future directions, Research quarterly for exercise and sport 71 Suppl 2 (2000) 1-14.
[29] M.S. Bernstein, M.C. Costanza, A. Morabia, Association of physical activity intensity levels with overweight and obesity in a population-based sample of adults, Preventive Medicine 38(1) (2004) 94-104.
[30] B.E. Ainsworth, W.L. Haskell, A.S. Leon, D.R. Jacobs, Jr., H.J. Montoye, J.F. Sallis, R.S. Paffenbarger, Jr., Compendium of physical activities: classification of energy costs of human physical activities, Medicine and science in sports and exercise 25(1) (1993) 71-80.
[31] B.E. Ainsworth, W.L. Haskell, M.C. Whitt, M.L. Irwin, A.M. Swartz, S.J. Strath, W.L. O'Brien, D.R. Bassett, Jr., K.H. Schmitz, P.O. Emplaincourt, D.R. Jacobs, Jr., A.S. Leon, Compendium of physical activities: an update of activity codes and MET intensities, Medicine and science in sports and exercise 32(9 Suppl) (2000) S498-504.
[32] D.W. Esliger, A.V. Rowlands, T.L. Hurst, M. Catt, P. Murray, R.G. Eston, Validation of the GENEA Accelerometer, Medicine and science in sports and exercise 43(6) (2011) 1085-93.
[33] C. Gubelmann, P. Vollenweider, P. Marques-Vidal, Of weekend warriors and couch potatoes: Socio-economic determinants of physical activity in Swiss middle-aged adults, Prev Med 105 (2017) 350-355.
[34] L.I. Lin, A concordance correlation coefficient to evaluate reproducibility, Biometrics 45(1) (1989) 255-68.
[35] G.B. McBride, A proposal for strength-of-agreement criteria for Lin's concordance correlation coefficient, 2005.

[36] D.E. Hinkle, W. Wiersma, S.G. Jurs, Applied statistics for the behavioral sciences, (1988).
[37] H.J.F. Helmerhorst, S. Brage, J. Warren, H. Besson, U. Ekelund, A systematic review of reliability and objective criterion-related validity of physical activity questionnaires, The international journal of behavioral nutrition and physical activity 9 (2012) 103-103.
[38] T.G. Pavey, S.R. Gomersall, B.K. Clark, W.J. Brown, The validity of the GENEActiv wrist-worn accelerometer for measuring adult sedentary time in free living, Journal of science and medicine in sport 19(5) (2016) 395-9.
[39] P.H. Lee, D.J. Macfarlane, T.H. Lam, S.M. Stewart, Validity of the international physical activity questionnaire short form (IPAQ-SF): A systematic review, The international journal of behavioral nutrition and physical activity 8 (2011) 115-115.
[40] L.B. Rasmussen, J. Matthiessen, A. Biltoft-Jensen, I. Tetens, Characteristics of misreporters of dietary intake and physical activity, Public health nutrition 10(3) (2007) 230-7.
[41] S.A. Adams, C.E. Matthews, C.B. Ebbeling, C.G. Moore, J.E. Cunningham, J. Fulton, J.R. Hebert, The effect of social desirability and social approval on self-reports of physical activity, American journal of epidemiology 161(4) (2005) 389-98.
[42] S. Sabia, V.T. van Hees, M.J. Shipley, M.I. Trenell, G. Hagger-Johnson, A. Elbaz, M. Kivimaki, A. Singh-Manoux, Association between questionnaire- and accelerometer-assessed physical activity: the role of sociodemographic factors, American journal of epidemiology 179(6) (2014) 781-90.
[43] S.E. Bonn, P. Bergman, Y. Trolle Lagerros, A. Sjölander, K. Bälter, A Validation Study of the Web-Based Physical Activity Questionnaire Active-Q Against the GENEA Accelerometer, JMIR Research Protocols 4(3) (2015) e86.
[44] S.M. Dyrstad, B.H. Hansen, I.M. Holme, S.A. Anderssen, Comparison of self-reported versus accelerometer-measured physical activity, Med Sci Sports Exerc 46(1) (2014) 99-106.
[45] C.M. Friedenreich, K.S. Courneya, H.K. Neilson, C.E. Matthews, G. Willis, M. Irwin, R. Troiano, R. Ballard-Barbash, Reliability and validity of the Past Year Total Physical Activity Questionnaire, American journal of epidemiology 163(10) (2006) 959-70.
[46] R.A. Washburn, A.M. Jette, C.A. Janney, Using Age-Neutral Physical Activity Questionnaires in Research with the Elderly, Journal of Aging and Health 2(3) (1990) 341-356.
[47] U. Ekelund, H. Sepp, S. Brage, W. Becker, R. Jakes, M. Hennings, N.J. Wareham, Criterion-related validity of the last 7-day, short form of the International Physical Activity Questionnaire in Swedish adults, Public health nutrition 9(2) (2006) 258-65.
[48] M.S. Buchowski, K.M. Townsend, K.Y. Chen, S.A. Acra, M. Sun, Energy expenditure determined by self-reported physical activity is related to body fatness, Obesity research 7(1) (1999) 23-33.
[49] C. Howitt, S. Brage, I.R. Hambleton, K. Westgate, T.A. Samuels, A.M. Rose, N. Unwin, A cross-sectional study of physical activity and sedentary behaviours in a Caribbean population: combining objective and questionnaire data to guide future interventions, BMC public health 16(1) (2016) 1036.
[50] G. Block, A.M. Hartman, C.M. Dresser, M.D. Carroll, J. Gannon, L. Gardner, A data-based approach to diet questionnaire design and testing, American journal of epidemiology 124(3) (1986) 453-69.
[51] A.V. Rowlands, E.M. Mirkes, T. Yates, S. Clemes, M. Davies, K. Khunti, C.L. Edwardson, Accelerometer-assessed Physical Activity in Epidemiology: Are Monitors Equivalent?, Medicine and science in sports and exercise 50(2) (2018) 257-265.
[52] S.G. Trost, K.L. McIver, R.R. Pate, Conducting accelerometer-based activity assessments in field-based research, Med Sci Sports Exerc 37(11 Suppl) (2005) S531-43.
[53] C.J. Caspersen, M.A. Pereira, K.M. Curran, Changes in physical activity patterns in the United States, by sex and cross-sectional age, Med Sci Sports Exerc 32(9) (2000) 1601-9.
[54] M.E. Rosenberger, M.P. Buman, W.L. Haskell, M.V. McConnell, L.L. Carstensen, Twenty-four Hours of Sleep, Sedentary Behavior, and Physical Activity with Nine Wearable Devices, Medicine and science in sports and exercise 48(3) (2016) 457-65.

Accepted author's manuscript. Maturitas.
Publisher DOI: https://doi.org/10.1016/j.maturitas.2019.08.004

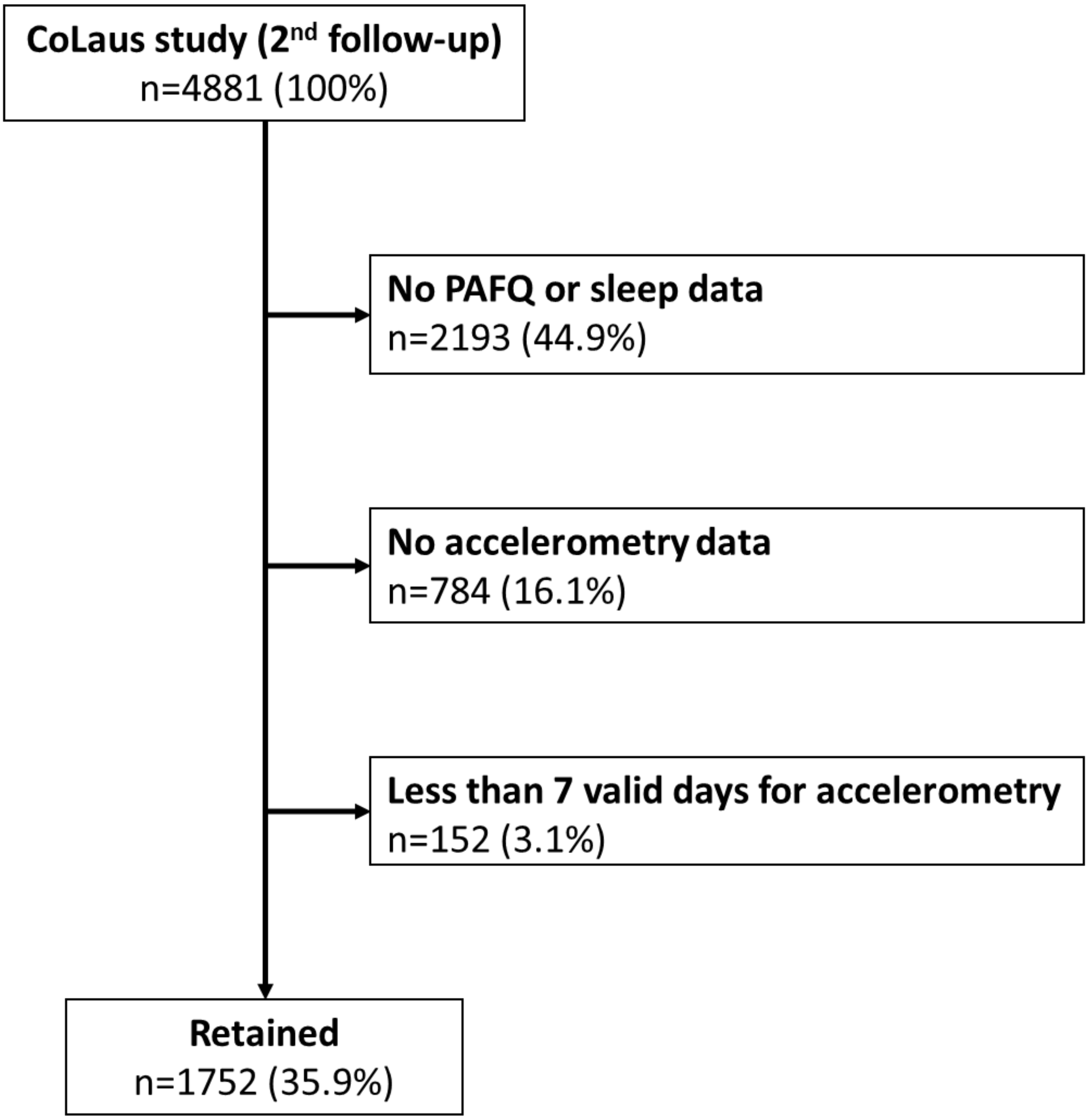
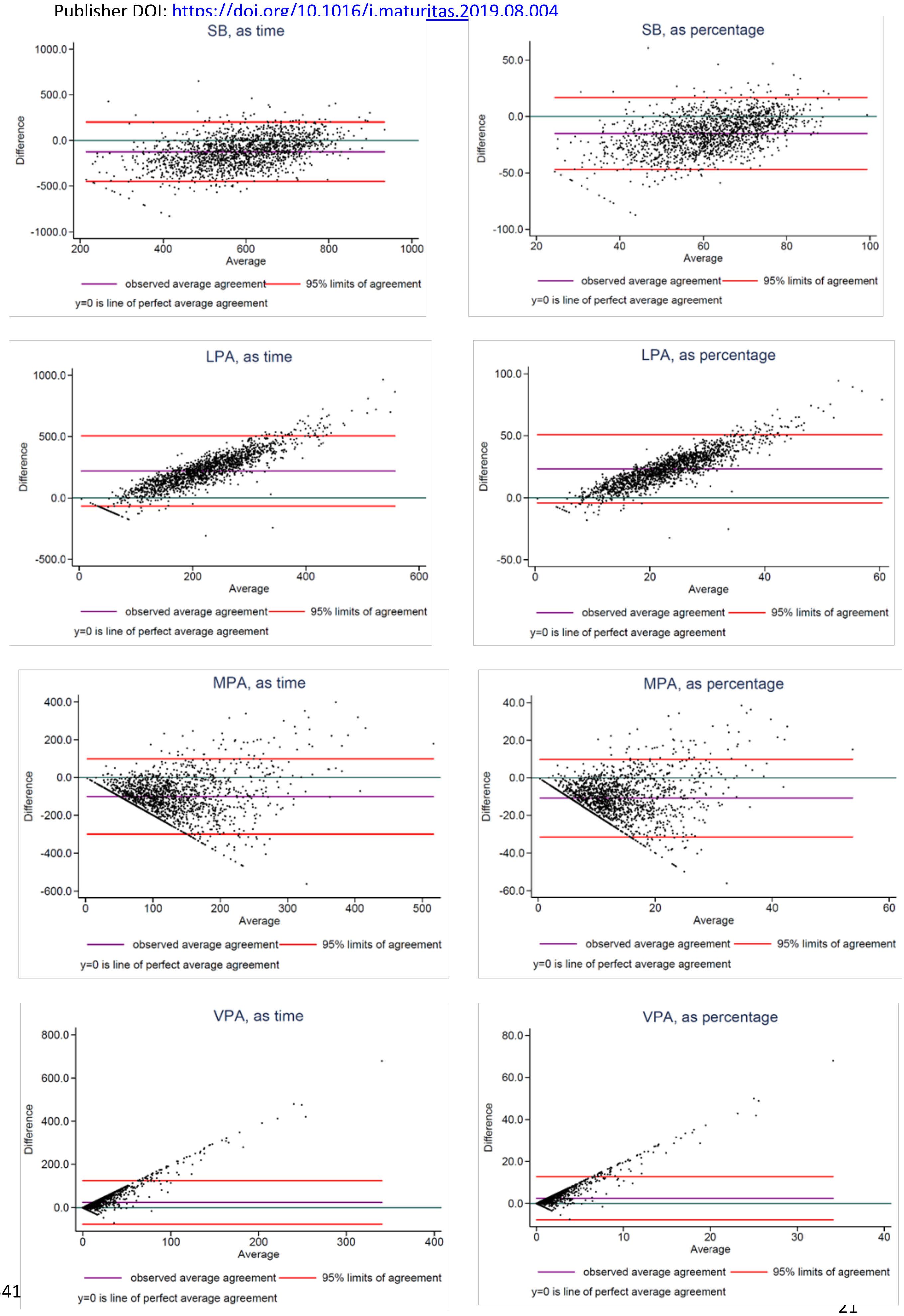


Figure 1. Flowchart of participants

Accepted author's manuscript. Maturitas.
Publisher DOI: https://doi.org/10.1016/j.maturitas.2019.08.004



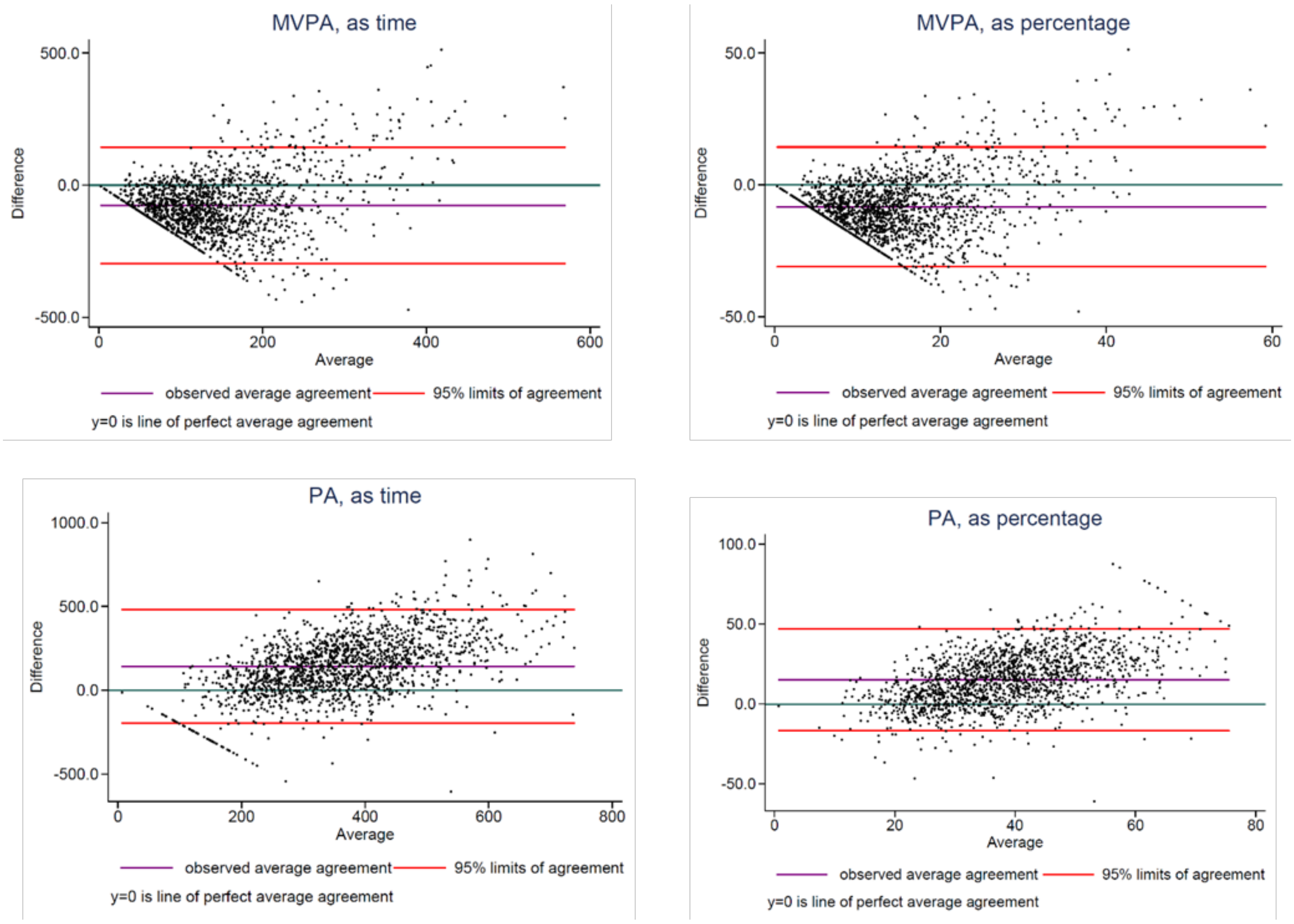


Figure 2. Bland-Altman plots as time and percentage of time for SB, LPA, MPA, VPA, MVPA, and total PA. x-axis depicts the average of the PAFQ and accelerometer; y-axis depicts difference between PAFQ and accelerometer

SB, sedentary behaviour; LPA, light physical activity; MPA, moderate physical activity; MVPA, moderate-to-vigorous physical activity; PA, physical activity.

Accepted author's manuscript. Maturitas.
Publisher DOI: https://doi.org/10.1016/j.maturitas.2019.08.004

Table 1. Activity levels according to PAFQ and accelerometer in 1752 participants.

	PAFQ	Accelerometer	Difference
As time (minutes/day)			
Sedentary behaviour	524 [410; 645]	652 [578; 724]	-123 [-238; -14]
Light PA	180 [119; 254]	106 [84; 130]	72 [12; 141]
Moderate PA	186 [120; 264.5]	162 [119; 219.5]	22.5 [-46; 100]
Vigorous PA	16 [0; 46]	2 [0; 5]	13 [-1; 41]
Moderate + vigorous PA	213 [143.5; 315]	166 [121.5; 226]	46 [-25; 141]
Total PA	424 [312; 541]	277 [215.5; 354]	143 [34.5; 249]
As % of time			
Sedentary behaviour	54.8 [42.9; 66.7]	70.4 [62.7; 76.8]	-14.7 [-25.7; -4.1]
Light PA	18.8 [12.6; 26.3]	11.3 [9.2; 13.7]	7.5 [1.1; 14.6]
Moderate PA	19.2 [12.7; 27.8]	17.5 [12.8; 23.4]	1.7 [-5.3; 10.1]
Vigorous PA	1.8 [0; 4.7]	0.2 [0; 0.6]	1.3 [-0.1; 4.2]
Moderate + vigorous PA	22.3 [14.8; 33.0]	18.1 [13.0; 24.2]	4.1 [-3.3; 14.2]
Total PA	45.2 [33.3; 57.1]	29.5 [23.2; 37.3]	14.7 [4.1; 25.7]

PA, physical activity; PAFQ, physical activity frequency questionnaire. Difference indicates PAFQ – accelerometer. Results are expressed as median [interquartile range]. Between-group comparisons are performed using Wilcoxon signed-rank test. All differences are significant at $p<0.001$.

Publisher DOI: https://doi.org/10.1016/j.maturitas.2019.08.004

Table 2. Correlation of activity levels between PAFQ and accelerometer.

	Spearman	Lin
As time (minutes/day)		
Sedentary behaviour	0.370 (0.325 – 0.409)	0.238 (0.208 – 0.268)
Light PA	0.205 (0.157 – 0.248)	0.074 (0.055 – 0.093)
Moderate PA	0.285 (0.241 – 0.331)	0.239 (0.201 – 0.276)
Vigorous PA	0.171 (0.124 – 0.221)	0.044 (0.035 – 0.053)
Moderate + vigorous PA	0.329 (0.289 – 0.372)	0.250 (0.218 – 0.282)
Total PA	0.373 (0.331 – 0.417)	0.209 (0.183 – 0.236)
As % of time		
Sedentary behaviour	0.387 (0.346 – 0.426)	0.218 (0.192 – 0.244)
Light PA	0.215 (0.165 – 0.260)	0.075 (0.057 – 0.094)
Moderate PA	0.279 (0.233 – 0.325)	0.239 (0.201 – 0.277)
Vigorous PA	0.171 (0.126 – 0.217)	0.045 (0.036 – 0.055)
Moderate + vigorous PA	0.326 (0.283 – 0.367)	0.254 (0.221 – 0.286)
Total PA	0.387 (0.348 – 0.430)	0.218 (0.192 – 0.244)

PA, physical activity. Results are expressed as Spearman rank correlation or Lin concordance coefficients (95% confidence intervals). For Spearman correlation coefficients, 95% confidence intervals are bootstrapped and bias-corrected values are presented. All coefficients are statistically significant at $p<0.001$. Results for sedentary behaviour and total PA expressed as % of time are similar as values are complementary.

Accepted author's manuscript. Maturitas.
Publisher DOI: https://doi.org/10.1016/j.maturitas.2019.08.004

Table 3. Associations of difference in time spent in sedentary behaviour and different levels of physical activity between PAFQ and accelerometer with gender, age and body mass index groups.

	Gender (men vs. women)		**Age (<65 vs. ≥65 years)**		**BMI (low vs. high)**	
	Constant	**Coefficient**	**Constant**	**Coefficient**	**Constant**	**Coefficient**
As time (mins/day)						
Sedentary behaviour	-130.7*	12.7 (-3.0; 28.5)	-113.9*	-32.4 (-49.1; -15.6)*	-117.2*	-13.0 (-28.8; 2.9)
Light PA	240.9*	-43.5 (-57.0; -30.1)*	227.4*	-24.5 (-39.0; -10.0)*	216.5*	5.3 (-8.5; 19.0)
Moderate PA	-118.6*	36.7 (27.3; 46.1)*	-111.6*	34.3 (24.6; 44.4)*	-102.7*	4.0 (-5.7; 13.6)
Vigorous PA	21.5*	4.6 (-0.2; 9.5)	25.9*	-6.5 (-11.7; -1.4)*	23.5*	0.6 (-4.2; 5.5)
Moderate + vigorous PA	-97.0*	41.4 (31.0; 51.7)*	-85.7*	27.8 (16.6; 39.0)*	-79.3*	4.6 (-6.0; 15.2)
Total PA	143.9*	-2.2 (-18.4; 14.0)	141.7*	3.3 (-14.0; 20.6)	137.3*	9.9 (-6.4; 26.1)

BMI, body mass index; PA, physical activity; PAFQ, physical activity frequency questionnaire. Results are expressed as coefficients (95% confidence intervals). All coefficients marked with * are statistically significant at $p<0.05$. The categories of women, age <65 years and low BMI are considered as reference.

Accepted author's manuscript. Maturitas.
Publisher DOI: https://doi.org/10.1016/j.maturitas.2019.08.004

Supplementary material

Supplementary table 1. Characteristics of excluded and included participants.

	Included (n=1752)	Excluded (n=3129)	p-value
Female (%)	889 (50.7)	1800 (57.5)	< 0.001
Age (years)	60.5 ± 9.4	64.3 ± 10.8	< 0.001
Age groups (%)			< 0.001
<65	1183 (67.5)	1665 (53.2)	
≥65	569 (32.5)	1464 (46.8)	
Body mass index (kg/m^2)	26.1 ± 4.6	26.6 ± 4.8	< 0.001
BMI categories (%)			< 0.001
Low	769 (43.9)	1088 (39.9)	
High	983 (56.1)	1641 (60.1)	

BMI, body mass index. Results are expressed as number of participants (column percentage) or as average ± standard deviation. Between-group comparisons are performed using chi-square for categorical variables or student's t-test for continuous variables. For BMI of excluded participants, number of participants does not add to 3129 due to the presence of missing values.

Accepted author's manuscript. Maturitas.
Publisher DOI: https://doi.org/10.1016/j.maturitas.2019.08.004

Supplementary table 2. Activity levels according to PAFQ and accelerometer, after exclusion of 40 participants with sleep time <4 or >10 hours/day (n = 1712)

	PAFQ	**Accelerometer**	**Difference**
	Median [IQR]	Median [IQR]	Median [IQR]
As time (minutes/day)			
Sedentary behaviour	526 [412; 648]	652 [578; 725]	-121 [-238; -12]
Light PA	181 [120; 254.5]	107 [84; 131]	73 [12; 142]
Moderate PA	187 [121; 267]	164 [120; 220]	22 [-47; 101]
Vigorous PA	16 [0; 46]	2 [0; 5]	13 [-1; 41]
Moderate + vigorous PA	213.5 [144; 317]	168.5 [123; 227]	46 [-25.5; 142]
Total PA	426 [314.5; 544]	278 [218; 355.5]	143.5 [34.5; 251.5]
As % of time			
Sedentary behaviour	54.8 [42.8; 66.7]	70.3 [62.6; 76.7]	-14.6 [-25.6; -4.1]
Light PA	18.8 [12.5; 26.4]	11.3 [9.2; 13.7]	7.5 [1.0; 14.6]
Moderate PA	19.2 [12.7; 27.8]	17.6 [12.9; 23.4]	1.6 [-5.3; 10.0]
Vigorous PA	1.8 [0; 4.8]	0.2 [0.1; 0.6]	1.4 [-0.1; 4.2]
Moderate + vigorous PA	22.3 [14.8; 33.0]	18.2 [13.1; 24.2]	4 [-3.4; 14.1]
Total PA	45.2 [33.4; 57.3]	29.7 [23.3; 37.4]	14.6 [4.1; 25.6]

PA, physical activity; PAFQ, physical activity frequency questionnaire. Results are expressed as median [interquartile range]. Between-group comparisons are performed using Wilcoxon signed-rank test. All differences are significant at $p<0.001$

Accepted author's manuscript. Maturitas.
Publisher DOI: https://doi.org/10.1016/j.maturitas.2019.08.004

Supplementary table 3. Correlation of activity levels between PAFQ and accelerometer, after exclusion of 40 participants with sleep time <4 or >10 hours/day excluded (n = 1712)

	Spearman	Lin
As time (minutes/day)		
Sedentary behaviour	0.373 (0.331 – 0.415)	0.240 (0.210 – 0.270)
Light PA	0.204 (0.156 – 0.249)	0.074 (0.054 – 0.093)
Moderate PA	0.282 (0.231 – 0.325)	0.234 (0.196 – 0.272)
Vigorous PA	0.169 (0.123 – 0.215)	0.044 (0.035 – 0.053)
Moderate + vigorous PA	0.326 (0.280 – 0.375)	0.246 (0.213 – 0.278)
Total PA	0.368 (0.324 – 0.410)	0.206 (0.179 – 0.232)
As % of time		
Sedentary behaviour	0.391 (0.347 – 0.430)	0.220 (0.193 – 0.246)
Light PA	0.216 (0.166 – 0.260)	0.076 (0.057 – 0.094)
Moderate PA	0.282 (0.239 – 0.326)	0.240 (0.201 – 0.278)
Vigorous PA	0.169 (0.122 – 0.214)	0.046 (0.036 – 0.055)
Moderate + vigorous PA	0.328 (0.283 – 0.373)	0.253 (0.220 – 0.286)
Total PA	0.391 (0.351 – 0.435)	0.219 (0.193 – 0.246)

PA, physical activity; PAFQ, physical activity frequency questionnaire. Results are expressed as Spearman rank correlation or Lin concordance coefficients (95% confidence intervals). For Spearman correlation coefficients, 95% confidence intervals are bootstrapped and bias-corrected values are presented. All coefficients are statistically significant at $p<0.001$. Results for sedentary behaviour and total PA expressed as % of time are similar as values are complementary.

Accepted author's manuscript. Maturitas.
Publisher DOI: https://doi.org/10.1016/j.maturitas.2019.08.004

Supplementary table 4. Associations of difference in time spent in SB and different levels of PA between PAFQ and accelerometer with gender, age and BMI groups, with adjustment for each other.

	Constant	Gender (men vs. women)	Age (<65 vs. ≥65 years)	BMI (low vs. high)
As time (minutes/day)				
Sedentary behaviour	-114.1*	14.2 (-1.7; 30.1)	-30.8 (-47.7; -14.0)*	-13.0 (-29.1; 3.1)
Light PA	242.8*	-46.8 (-60.4; -33.1)*	-27.2 (-41.6; -12.8)*	15.2 (1.4; 29.0)*
Moderate PA	-128.2*	38.5 (29.1; 47.9)*	35.9 (26.0; 45.8)*	-5.3 (-14.8; 4.2)
Vigorous PA	23.5*	4.4 (-0.5; 9.3)	-6.4 (-11.6; -1.3)*	0.4 (-4.5; 5.3)
Moderate + vigorous PA	-104.6*	42.9 (32.5; 53.4)*	29.5 (18.4; 40.5)*	-4.9 (-15.4; 5.7)
Total PA	138.2*	-3.8 (-20.3; 12.6)	2.3 (-15.0; 19.6)	10.3 (-6.3; 26.9)

PA, physical activity; PAFQ, physical activity frequency questionnaire. Results are expressed as coefficients (95% confidence intervals). All coefficients marked with * are statistically significant at $p<0.05$

Accepted author's manuscript. Maturitas.
Publisher DOI: https://doi.org/10.1016/j.maturitas.2019.08.004

Additional file 1: Categorization of the activities from the physical activity questionnaire

Occupational activities	Category
Office work, seated or standing (desk, PC, telephone, ...)	Sedentary
Writing at a black board, standing	Light
Housekeeping and cleaning (sweeping floors, vacuuming...)	Moderate
Cooking	Light
Sewing, ironing	Moderate
Washing, hanging out clothes	Light
Taking care of babies or toddlers	Moderate
Taking care of children calmly	Moderate
Carpentry, masonry, lock smith, handy work, electrician, etc.	Moderate
Heavy construction work (digging holes, carrying or pulling loads, etc.)	Vigorous
Auto mechanic	Moderate
Hand work in machine-tool, chemical, or electrical industry	Light
Warehouseman, storekeeper	Vigorous
Painting, decorating	Moderate
Laboratory work	Light
Gardening	Moderate
Harvesting fruits, vegetables, herbs	Moderate
Playing a musical instrument	Moderate
Ballroom or folk dancing	Moderate
Walking and Driving at Work (during Office Hours)	
Walking normally or walking slowly	Moderate
Walking quickly or walking uphill	Moderate
Walking while carrying or pulling a load	Vigorous
Climbing up stairs	Not coded
Driving a car or a truck	Light
Eating and Personal grooming	
Washing up (shower, bath, etc.), dressing, undressing	Sedentary
Meals and snacks	Sedentary
Nap or rest in bed	Sedentary
Domestic and Leisure Activities	
Sitting quietly (reading, television, playing cards)	Sedentary
Writing, using PC, typing	Sedentary
Standing quietly (for example waiting for something)	Light
Cooking, washing dishes	Light
Housekeeping and cleaning, washing and hanging out clothes	Moderate
Ironing, sewing	Light
Taking care of children calmly	Light
Taking care of babies or toddlers	Moderate
Leisure Time Activities	
Handy work standing up, auto mechanic, painting, wood work	Moderate
Gardening, racking leaves, pulling weeds, mowing the lawn...	Moderate

Accepted author's manuscript. Maturitas.
Publisher DOI: https://doi.org/10.1016/j.maturitas.2019.08.004

Working with a pitchfork	Vigorous
Spitting logs with an ax	Vigorous
Stacking fire place wood	Moderate
Singing	Light
Ballroom dancing	Moderate
Playing music (piano, violin, ...)	Light
Fishing	Light
Getting to work or elsewhere	
Walking normally	Moderate
Walking quickly or uphill	Moderate
Walking while carrying or pulling loads (cart on wheels, bags...)	Vigorous
Climbing up stairs	Not coded
Driving a car or a motorcycle	Light
Bicycling slowly (about 9 km/h)	Moderate
Bicycling fast (about 15 km/h)	Vigorous
Sports	
Bicycling (should not repeat questions 49-50)	Vigorous
Athletic walking	Vigorous
Running	Vigorous
Soccer	Vigorous
Basket-ball	Vigorous
Gymnastics	Moderate
Weight lifting	Moderate
Dancing (ballet, aerobics, rock)	Moderate
Swimming	Vigorous
Diving	Moderate
Skiing downhill, water skiing	Moderate
Cross-country skiing	Vigorous
Tennis (single), badminton	Vigorous
Tennis doubles	Vigorous
Squash	Vigorous
Golf, pulling a cart	Moderate
Golf (without pulling a cart)	Moderate
Judo, karate	Vigorous
Ice or roller skating	Vigorous

Accepted author's manuscript. Maturitas.
Publisher DOI: https://doi.org/10.1016/j.maturitas.2019.08.004

Additional file 2: Example of calculation of the standardized times spent in each type of activity

A participant reports sleeping an average of 10 hours (600 minutes) per day. He also reports SB activities of 600 minutes/day, LPA activities of 200 minutes/day, MPA activities of 90 minutes per day and VPA activities of 10 minutes/day.

Calculation of non-sleep time (1 day = $24 \times 60 = 1440$ minutes)

$$NST = 1440 - 600 = 840 \text{ minutes per day}$$

Calculation of total time spent in SB, LPA, MPA and VPA

$$T = 600 + 200 + 90 + 10 = 900 \text{ minutes/day} \qquad \text{here, } T > NST$$

Calculation of the percentage of time spent in each type of activity

P_{SB} =600/900 =0.667 or 66.7%
P_{LPA} =200/900 =0.222 or 22.2%
P_{MPA} =90/900 =0.100 or 10.0%
P_{VPA} =10/900 =0.011 or 1.1%

The standardized times are thus

$ST_{SB} = 840 \times 0.667 \cong 561$ minutes/day
$ST_{LPA} = 840 \times 0.222 = 186$ minutes/day
$ST_{MPA} = 840 \times 0.100 = 84$ minutes/day
$ST_{VPA} = 840 \times 0.011 = 9$ minutes/day

The calculation are similar if $T < NST$