

Physical activity derived from questionnaires and wrist-worn accelerometers: comparability and the role of demographic, lifestyle, and health factors among a population-based sample of older adults

Chantal M Koolhaas¹
Frank JA van Rooij¹
Magda Cepeda¹
Henning Tiemeier¹⁻³
Oscar H Franco¹
Josje D Schoufour¹

¹Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands; ²Department of Child and Adolescent Psychiatry, Erasmus Medical Center, Rotterdam, The Netherlands; ³Department of Psychiatry, Erasmus Medical Center, Rotterdam, The Netherlands

Background: Agreement between questionnaires and accelerometers to measure physical activity (PA) differs between studies and might be related to demographic, lifestyle, and health characteristics, including disability and depressive symptoms.

Methods: We included 1,410 individuals aged 51–94 years from the population-based Rotterdam Study. Participants completed the LASA Physical Activity Questionnaire and wore a wrist-worn accelerometer on the nondominant wrist for 1 week thereafter. We compared the Spearman correlation and disagreement (level and direction) for total PA across levels of demographic, lifestyle, and health variables. The level of disagreement was defined as the absolute difference between questionnaire- and accelerometer-derived PA, whereas the direction of disagreement was defined as questionnaire PA minus accelerometer PA. We used linear regression analyses with the level and direction of disagreement as outcome, including all demographic, lifestyle, and health variables in the model.

Results: We observed a Spearman correlation of 0.30 between questionnaire- and accelerometer-derived PA in the total population. The level of disagreement (ie, absolute difference) was 941.9 (standard deviation [SD] 747.0) minutes/week, and the PA reported by questionnaire was on average 529.4 (SD 1,079.5) minutes/week lower than PA obtained by the accelerometer. The level of disagreement decreased with higher educational levels. Additionally, participants with obesity, higher disability scores, and more depressive symptoms underestimated their self-reported PA more than their healthier counterparts.

Conclusion: We observed large differences in PA time derived from the LASA Physical Activity Questionnaire and the wrist-worn accelerometer. Differences between the methods were related to body-mass index, level of disability, and presence of depressive symptoms. Future studies using questionnaires and/or accelerometers should account for these differences.

Keywords: physical activity, questionnaire, accelerometer, disagreement, elderly

Introduction

Physical activity (PA) is an important modifiable risk factor in the prevention of diseases, including cardiovascular disease, diabetes mellitus, and several types of cancer.¹ For practical reasons, PA is often measured subjectively by questionnaire.² However, self-reported PA can suffer from reporting bias, partially attributable to the cognitive challenge of estimating the frequency, intensity, and duration of PA.³ Moreover, activities of light intensity are hard to recall and might not be reported.^{4,5} Objective methods, including accelerometers, offer a solution to these problems and can give objective estimates

Correspondence: Chantal Koolhaas
Department of Epidemiology, Erasmus
Medical Center, PO Box 2040,
Rotterdam 3000 CA, The Netherlands
Tel +31 10 704 3484
Fax +31 10 704 4657
Email c.koolhaas@erasmusmc.nl

on duration and intensity of PA. Therefore, accelerometers are increasingly being used in current research.^{2,6} However, when comparing questionnaire data with objective methods applied in large populations, major discrepancies emerge.^{7,8} The inconsistency might be related to recall bias, which can be influenced by population characteristics, such as age, sex, and health status.^{2,9} In addition, the inconsistency might stem from the fact that accelerometers cannot measure all PA accurately.^{10,11} For example, weightlifting and cycling are generally underestimated by accelerometers worn on the upper body.¹² Considering the increased use of accelerometers in current research,^{2,6} it is important to understand and quantify how PA assessed with questionnaires and accelerometers differs and how these differences relate to population characteristics.¹³

Previous studies have shown that the correlation between self-reported and objectively measured PA differs by age, sex, ethnicity, socioeconomic status, and level of PA.^{8,9} However, results presented referred to correlation coefficients, which are a measure of the extent to which two variables are linearly related, but do not take into account their measurement scales. As such, these studies did not provide information regarding the level of agreement and direction of disagreement of the two methods.¹⁴ A study that accounted for the direction of disagreement between objective and subjective measured PA in young adults suggested that overweight adults had a tendency to overestimate the time spent in vigorous PA more than normal-weight individuals.¹⁵ Information from older adults, a population with different PA patterns than younger adults,¹⁶ is currently lacking. Moreover, the effect of socioeconomic status and mental and physical health on the agreement between accelerometer and questionnaire remains unclear. Therefore, we aimed to quantify the level and direction of disagreement between questionnaire-assessed PA and accelerometer-assessed PA and to investigate if differences in agreement are explained by sociodemographic, lifestyle, and health factors.

Subjects and methods

Study population

This paper utilizes data from the Rotterdam Study, a population-based cohort designed to examine the onset and risk factors of diseases in older adults.¹⁷ The Rotterdam Study was approved by the medical ethics committee (The Ministry of Health, Welfare and Sport of the Netherlands) according to the Wet Bevolkingsonderzoek ERGO (Population Study Act, Rotterdam Study), executed by the Ministry of Health, Welfare and Sport of the Netherlands. Trained research assistants interviewed the participants at home to collect the baseline information and individuals visited the research center twice, where clinical measurements were obtained.

Participants were invited to wear an accelerometer and fill out a PA questionnaire between June 2011 and June 2014 (wave 1) and between July 2014 and May 2016 (wave 2). Participants were included in the analysis if they had valid accelerometer data (>1,200 min/day) for at least 4 days, including one weekend day. A total of 24 cases were excluded because of unreliably high (ie, higher than the mean + three times the standard deviation) levels of data on self-reported PA (Figure 1). Since the PA questionnaire used in the Rotterdam Study did not cover occupational activity, we also excluded participants with paid occupation (n=389) or without information on occupational status (n=27). Additionally, because 72 participants participated in both waves, we excluded observations from the second wave of these participants to avoid clustering of data. In the current analyses, we included 1,410 adults aged 50+ years with data on both objective and subjective PA. All subjects gave written informed consent.

Questionnaire-assessed physical activity

After the home interview, participants received the PA questionnaire and were requested to hand in their questionnaire during their first research-center visit. Questionnaire PA was assessed with the self-administered LASA Physical Activity Questionnaire (LAPAQ), a validated questionnaire with reasonably good test-retest reliability (0.65–0.75).¹⁸ Correlations with a pedometer and 7-day diary were 0.56 and 0.68, respectively.¹⁸ The LAPAQ includes questions on the frequency and duration of walking, cycling, sports (including two open-ended questions in which participants could report “other sports” they participated in), gardening, and housework. Participants reported how many hours/week they had spent in each activity in the previous 2 weeks. Detailed information on the assessment of questionnaire-derived PA can be found elsewhere.¹⁹ We used metabolic equivalent of task (MET) to quantify activity intensity. MET values were assigned to all activities in the questionnaire, using a compendium of activity energy costs.²⁰ Of all other activities, 18 (3.2%) were not sports and 28 (5.1%) were not in the compendium (eg, physiotherapy training, indoor sports, revalidation). No MET values were assigned to these activities, and they were not included in the analyses. Duration in total PA was computed by summing time spent in all activities mentioned in the questionnaire.

To be able to evaluate whether differences in total PA were driven by a particular intensity category (ie, light, moderate, or vigorous PA), we categorized activities into three groups. Activities with MET values lower than 3 were coded as light-intensity PA, 3–6 MET was coded as moderate-intensity PA, and ≥6 MET was coded as vigorous-intensity PA.²¹

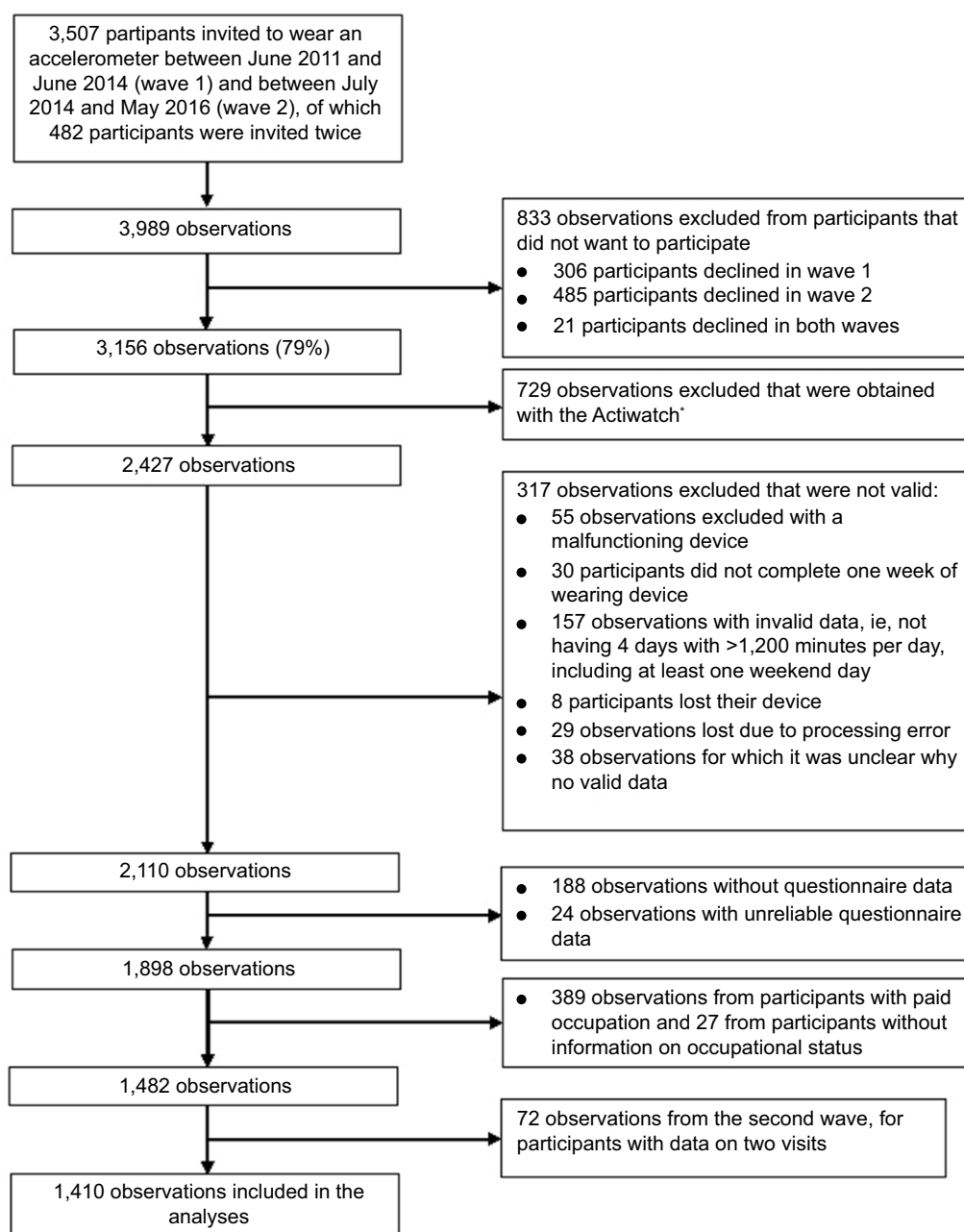


Figure 1 Flowchart of participant inclusion in the Rotterdam Study.

Notes: *The Actiwatch (Actiwatch model AW4; Cambridge Technology, Cambridge, UK) is a one-dimensional device and cannot be used to measure physical activity.

Accelerometer-assessed physical activity

At the first center visit, participants were requested to wear a triaxial accelerometer (GeneActiv; ActivinsightsLtd, Kimbolton, UK) on the nondominant wrist for 7 consecutive days and nights. Since the GeneActiv device is waterproof, the device can also be worn while bathing and swimming. The accelerometer was sampled at 50 Hz and acceleration was expressed relative to gravity ($1\text{ g} = 9.81\text{ m/second}^2$).^{9,22,23} Data were extracted from the first wearing day up to 7 days later, and all days with $\geq 1,200$ minutes of data were included.

To quantify the acceleration related to registered movement, we calculated the high-pass filtered vector magnitude. This approach applies a high-pass filter with a cutoff frequency of 0.2 Hz to the acceleration signal, treating gravity as a low-frequency component that is filtered out.²⁴ Accelerometer data were processed in Python (2.6.6) with the open-access Pampro software, a program for the systematic analysis of PA data collected in epidemiological studies.²⁵ Nonwear time was defined as all time periods where the standard deviation of acceleration in each of the three axes

fell below 13 mg for over 1 hour.²⁴ Any nonwear period was excluded from analyses. The pattern of nonwear time was accounted for by balancing the weighting of the data according to the diurnal profile.²⁶

Activity was categorized into sedentary (<48 mg), light (48–154 mg), moderate (154–389 mg), and vigorous activity (>389 mg), based on a recent validation study in 1,695 middle-aged adults (mean age 50 years).²⁴ Because not all participants wore the watch for 7 days, data were recoded such that our measures reflected PA over a 1-week course to match the questionnaire data. Average duration in light, moderate, and vigorous PA was calculated for weekdays and weekend days. Consequently, the weekly accelerometer-assessed total PA was calculated as $([5 \times \text{mean daily duration in total PA on weekdays}] + [2 \times \text{mean daily duration in total PA on weekend days}])$.

Assessment of factors

Self-reported alcohol use was obtained with a food-frequency questionnaire, expressed in grams/day, and categorized in tertiles. Socioeconomic status was evaluated by education, assessed in consonance with the international standard classification of education and categorized as primary, lower, intermediate, and higher education.²⁷ Smoking was categorized in three categories: current, former, and never. Body-mass index (BMI) was calculated using height and weight and defined as normal weight (<25 kg/m²), overweight (25–30 kg/m²), and obese (≥30 kg/m²). Marital status was defined as living with a partner or not. Disability was assessed by activities of daily living, from the Stanford Health Assessment Questionnaire Disability Index.²⁸ For descriptive purposes, disability was defined as a disability score value >0.5 and severe disability as a disability score value >1.²⁹ The presence of coronary heart disease, diabetes, stroke, and cancer were determined using medical records up to the year 2012 to define the number of comorbidities. Depressive symptoms were assessed with the Center for Epidemiologic Studies Depression (CESD) scale,³⁰ a self-report scale with 20 items and a maximum score of 60. A score of 16 or greater is traditionally accepted as the cutoff to define clinical depression,³¹ and was used for descriptive purposes. Cognitive function was assessed with the Mini-Mental State Examination (MMSE).³² MMSE scores range from 0 to 30, with a higher score indicating better cognitive performance. We used a cutoff of 26 to categorize participants as having a cognitive impairment in our descriptive analysis.³³

Statistical analyses

Descriptive statistics were used to characterize the sample. The difference in time spent in PA according to questionnaire

and accelerometer was examined using a paired-samples *t*-test, and the Bland–Altman method was used to visualize the level of (dis)agreement between questionnaire- and accelerometer-derived PA levels. We assessed the Spearman correlation between questionnaire- and accelerometer-derived PA in the total population and separately for the following categories: age-group (ie, 50–60 years, 60–70 years, 70–80 years, >80 years), sex, education, marital status, smoking status, BMI category, alcohol consumption, disability, prevalent chronic disease, cognitive function, and prevalent depression. Since the Spearman correlation coefficient is equal to the slope of the regression between the ranked values of the questionnaire and accelerometer, we used linear regression models to obtain correlation coefficients and corresponding 95% confidence intervals using accelerometer-derived PA as the dependent variable.

For our main analyses, we devised two measures of disagreement, in resemblance with previous research.³⁴ First, we computed the level of disagreement, expressed as the absolute time difference between questionnaire- and accelerometer-assessed PA. Second, we computed the direction of disagreement, indicating whether an individual had the tendency to overestimate or underestimate PA, by subtracting accelerometer-assessed PA from questionnaire-derived PA levels. To assess the association between assessed factors and the level and direction of disagreement, we used linear regression analyses with the level and direction of disagreement as outcome variables, including all demographic, lifestyle, and health variables, total PA measured by the accelerometer, time between interview date and the first accelerometer-wear date, cohort, and wave. Age, CESD score, disability index, MMSE score, and number of comorbidities were used as continuous variables in these analyses. To quantify if the association was dependent on the level of PA, β -values (95% CI) for the association of accelerometer-derived PA, expressed in hours/week, with the level and direction of disagreement were also obtained from these analyses.

We performed several sensitivity analyses. We compared baseline characteristics for those agreeing to wear an accelerometer versus those who rejected, and in those with paid occupation versus those without paid occupation. Next, we examined the level and direction of disagreement in each intensity category of PA (light, moderate, and vigorous). Additionally, we repeated the analyses in a population also including participants with paid occupation. In these analyses, we additionally included occupational status as a covariate. Moreover, because accelerometers do not measure cycling accurately, we repeated the analyses in those that did not report any cycling.

Our figures contained 49.1% missing data on alcohol use. Other covariates had <2% missing data. We imputed missing data using Markov chain Monte Carlo multiple imputation (n=20 imputations). All analyses were performed using IBM SPSS Statistics for Windows (IBM, Armonk, NY, USA) and R version 3.2.1 (R Foundation for Statistical Computing, Vienna, Austria). Statistical significance was set at $P<0.05$.

Results

Participants who agreed to wear an accelerometer were slightly younger, more often men, and showed a better health profile with regard to MMSE score, CESD score, and disability score than those who did not agree to wear an accelerometer (Table S1). The participants included in our study had a mean age of 73.8 (SD 7.6) years, and 742 (52.6%) were female (Table 1). The mean (SD) daily wearing time of the accelerometer was 23.8 (0.3 hours), and the median time difference between the interview and the first accelerometer-wear date was 8.7 weeks (interquartile range: 5.1–13.7). The mean (SD) weekly time spent in PA was 1,579.3 (386.1) minutes according to the accelerometer and 1,049.9 (1,108.3) minutes according to the questionnaire ($P<0.001$), indicating that overall, total PA was underestimated in the questionnaire. This is also visualized in the Bland–Altman plot in Figure S1. The Spearman correlation was 0.30 (95% CI 0.25–0.34) in the total sample, and differed as a function of education, disability, smoking status, depressive symptoms, and cognitive function (Table 2).

For every 1 hour/week more accelerometer-derived PA, the level of disagreement was 30.2 minutes/week more (95% CI 23.6–36.7, $P<0.001$). Independently of PA volume, those with intermediate and higher education showed smaller differences between the questionnaire PA and accelerometer PA than those with primary education (Figure 2). No other determinants were associated with the level of disagreement.

For every 1 hour/week more accelerometer-derived PA, questionnaire PA was underestimated more by 29.1 minutes (95% CI –38.6 to –19.6, $P<0.001$). Results of regression analyses with direction of disagreement are presented in Figure 3. Independently of PA volume, obese participants underestimated their self-reported PA levels more than normal-weight participants ($\beta=-234.3$ minutes/week, 95% CI –398.2 to –70.5; $P=0.002$). Additionally, the underestimation of self-reported PA in the questionnaire was larger for participants with higher disability and CESD scores (Figure 3).

Sensitivity analyses showed that light and vigorous PA were underestimated by the questionnaire (mean [SD]

Table 1 Characteristics of the study population (n=1,410)

| Characteristics | |
|--|-------------------|
| Demographic factors | |
| Age (years) | 73.8 (7.6) |
| Female, n (%) | 742 (52.6) |
| Education | |
| Elementary | 86 (6.1) |
| Lower secondary | 571 (40.5) |
| Higher secondary | 434 (30.8) |
| Tertiary | 319 (22.6) |
| Living with partner, n (%) | 989 (70.1) |
| Lifestyle factors | |
| Smoking, n (%) | |
| Never-smoker | 471 (33.4) |
| Former smoker | 833 (59.1) |
| Current smoker | 106 (7.5) |
| BMI, n (%) | |
| Normal weight | 418 (29.6) |
| Overweight | 668 (47.4) |
| Obese | 324 (23.0) |
| Alcohol consumption, n (%) | |
| Low | 502 (35.6) |
| Medium | 464 (32.9) |
| High | 444 (31.5) |
| Health factors | |
| Disability score | 0.64 (0.68) |
| Prevalent cancer, CVD or diabetes, n (%) | 635 (45) |
| MMSE score | 28.1 (1.7) |
| CESD score | 5.0 (6.3) |
| Physical activity variables | |
| Total PA according to questionnaire (minutes/week) | 1,049.9 (1,108.3) |
| Total PA according to accelerometer (minutes/week) | 1,579.3 (386.1) |
| Level of disagreement between questionnaire and accelerometer | 941.9 (747.0) |
| Direction of disagreement between questionnaire and accelerometer | –529.4 (1,079.5) |
| Time between interview and first accelerometer-wear date (weeks), median (IQR) | 8.7 (5.1–13.7) |

Note: Data are presented as mean (SD) unless otherwise stated.

Abbreviations: BMI, body-mass index; CESD, Center for Epidemiologic Studies Depression; CVD, cardiovascular disease; IQR, interquartile range; MMSE, Mini-Mental State Examination; PA, physical activity; SD, standard deviation.

708.4 [484.3] minutes/week and 25.6 [338.1] minutes/week, respectively), whereas the level of moderate PA was on average overestimated by 204.6 minutes/week (SD 780.8) (Table S2). In analyses with light, moderate, and vigorous PA, factors associated with the level and direction of disagreement were similar to the main analyses (Tables S3 and S4). When comparing those with and without paid occupation, we observed that those with paid occupation were younger, more highly educated, and showed a better health profile regarding CESD score, MMSE score, and disability score

Table 2 Spearman correlation between questionnaire-assessed total physical activity and accelerometer-assessed total physical activity, according to characteristics of the study population

| Total population, factors, and physical activity | n | Spearman's <i>r</i> | 95% CI |
|--|-------|---------------------|---------------|
| Total population | 1,410 | 0.30 | 0.25–0.34 |
| Demographic factors | | | |
| Age-group | | | |
| 50–60 years | 62 | 0.24 | 0.01–0.48 |
| 60–70 years | 334 | 0.26 | 0.16–0.37 |
| 70–80 years | 756 | 0.28 | 0.21–0.34 |
| >80 years | 258 | 0.24 | 0.13–0.36 |
| Sex | | | |
| Male | 668 | 0.32 | 0.25–0.39 |
| Female | 742 | 0.27 | 0.20–0.33 |
| Education | | | |
| Elementary | 86 | 0.46 | 0.28–0.64 |
| Lower secondary | 571 | 0.27 | 0.19–0.35 |
| Higher secondary | 434 | 0.28 | 0.19–0.37 |
| Tertiary | 319 | 0.31 | 0.21–0.42 |
| Marital status | | | |
| Living alone | 421 | 0.31 | 0.27–0.36 |
| Living with someone | 989 | 0.28 | 0.26–0.31 |
| Lifestyle factors | | | |
| Smoking | | | |
| Never-smoker | 472 | 0.28 | 0.22–0.35 |
| Former smoker | 833 | 0.32 | 0.29–0.35 |
| Current smoker | 106 | 0.17 | 0.08–0.26 |
| BMI | | | |
| Normal weight | 418 | 0.27 | 0.18–0.36 |
| Overweight | 668 | 0.29 | 0.21–0.36 |
| Obese | 324 | 0.26 | 0.16–0.36 |
| Alcohol consumption | | | |
| Low | 502 | 0.29 | 0.20–0.37 |
| Medium | 646 | 0.29 | 0.19–0.40 |
| High | 444 | 0.31 | 0.21–0.42 |
| Health factors | | | |
| Disability | | | |
| Not disabled | 743 | 0.27 | 0.21–0.34 |
| Disabled | 300 | 0.31 | 0.26–0.36 |
| Severely disabled | 367 | 0.23 | 0.14–0.33 |
| Prevalent chronic disease | | | |
| No prevalent chronic disease | 775 | 0.28 | 0.21–0.35 |
| Prevalent cancer, CVD, or diabetes | 635 | 0.29 | 0.22–0.36 |
| Cognitive function | | | |
| MMSE score <26 | 97 | 0.43 | 0.25–0.61 |
| MMSE score ≥26 | 1,313 | 0.29 | 0.23–0.34 |
| Depression | | | |
| No depression | 1,304 | 0.29 | 0.27–0.32 |
| Depression present | 106 | 0.33 | 0.15–0.51 |
| Physical activity | | | |
| Self-reported PA ^a | | | |
| Low | 485 | 0.16 | 0.07–0.25 |
| Medium | 445 | 0.16 | 0.07–0.25 |
| High | 480 | 0.10 | 0.01–0.19 |
| Accelerometer-derived PA ^b | | | |
| Low | 514 | 0.23 | 0.14–0.31 |
| Medium | 485 | 0.08 | –0.01 to 0.16 |
| High | 411 | 0.11 | 0.02–0.20 |

(Continued)

Table 2 (Continued)

Notes: ^aMedian reported levels of physical activity across low, medium, and high categories were 210, 765 and 1,807 minutes/week, respectively. ^bMedian accelerometer-derived level of physical activity across low, medium, and high categories were 1,229, 1,601 and 2,003 minutes/week, respectively.

Abbreviations: BMI, body-mass index; CI, confidence interval; CVD, cardiovascular disease; MMSE, Mini-Mental State Examination; PA, physical activity.

(Table S5). In sensitivity analyses on those with paid occupation ($n=1,823$), the level and direction of disagreement were larger compared to the main analyses (mean [SD] 971.8 [788] minutes/week and –589 [1,104] minutes/week, respectively). Factors associated with the disagreement were similar to the sample included in the main analyses (Table S6). In those not reporting cycling ($n=692$), factors associated with the direction of disagreement were equal compared to the sample included in the study (Table S7), although some associations were no longer significant.

Discussion

The aim of the current study was to quantify the level and direction of disagreement between questionnaire- and accelerometer-assessed PA and to investigate if sociodemographic, lifestyle, and health factors could explain differences in agreement. We found that the overall agreement between the LAPAQ and a wrist-worn accelerometer was low and that it differed across lifestyle and health variables. The level of disagreement was higher in less educated individuals. In addition, obese individuals, those with a higher disability score, and those with more depressive symptoms underreported PA more in the questionnaire compared to the reference group. Moreover, with higher levels of accelerometer-derived PA, the underestimation of PA in the questionnaire was more obvious.

The low correlation between accelerometer- and questionnaire-derived PA observed in the current study ($r=0.30$) was similar to results reported in previous studies in older adults,^{2,9} but lower than another study using the LAPAQ in 439 adults aged 69–92 years ($r=0.56$).¹⁸ This difference might be related to study design, as participants in Stel et al wore a pedometer directly after filling out the LAPAQ,¹⁸ whereas in the current study the time between questionnaire- and accelerometer-assessment had a median value of 8.7 weeks. Moreover, in Stel et al,¹⁸ participants additionally completed a 7-day diary while wearing the pedometer, which might have altered their PA patterns, due to increased awareness. Furthermore, in contrast to most studies, but similar to the study of Sabia et al,⁹

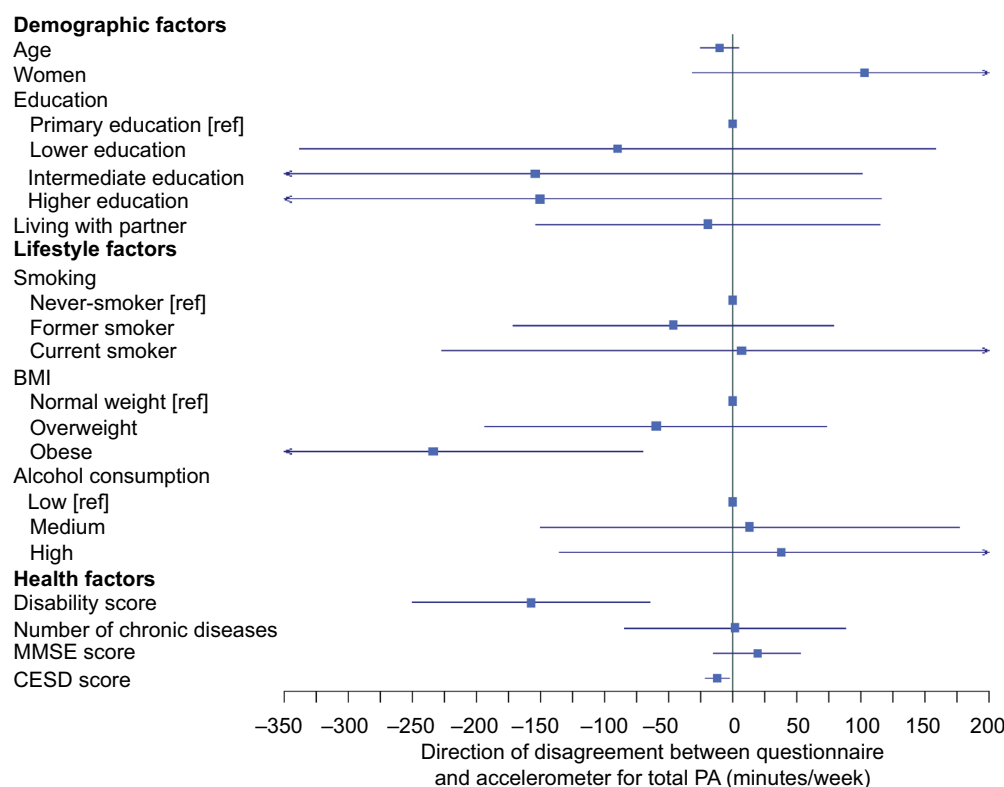


Figure 2 Factors associated with level of disagreement (ie, absolute difference) between accelerometer and questionnaire physical activity.

Notes: A positive difference indicates a larger absolute difference between the questionnaire and the accelerometer compared to the reference, whereas a negative difference indicates a smaller difference. Analyses included all demographics: lifestyle and health variables, total accelerometer-derived physical activity, time between interview date and first accelerometer-wear date, cohort, and wave. The coefficient for women was relative to men, and the coefficient for paid occupation relative to those without. For continuous variables (age, disability score, number of chronic diseases, MMSE score, and CESD score), values correspond to a 1-unit increase in the determinant.

Abbreviations: BMI, body-mass index; CESD, Center for Epidemiologic Studies Depression; MMSE, Mini-Mental State Examination; PA: physical activity; ref, reference.

our participants wore the accelerometer around the wrist. There is general consensus that a wrist-worn accelerometer is worse in estimating activity during large-muscle, dynamic activities like cycling than a hip-worn device.³⁵ This could lead to an underestimation of PA performed solely by the legs.^{10,11} Moreover, when the wrist is constrained during PA, eg, when carrying a briefcase or groceries, the PA level will likely be underestimated.³⁵ Considering the fact that these activities are usually of light–moderate activity, the agreement between questionnaire and accelerometer might be different for hip-worn accelerometers for these intensity domains.

However, an advantage of the wrist-worn accelerometer is that it allows for 24 hours of data collection per day, including during water-based activities, leading to a comprehensive overview of daily PA. This is in contrast to hip-worn devices that have to be removed during the night and are usually not waterproof. However, since the correlation between PA derived from questionnaires and accelerometers observed in the current study was similar to correlations in studies using hip-worn devices,^{2,9} the high compliance of the wrist-worn

accelerometers is unlikely to enhance the comparability between questionnaire and accelerometer.

In the total population, we found that compared to the accelerometer, the questionnaire on average underestimated PA levels by 529 minutes/week, equivalent to 76 minutes/day. This difference could be explained by missing items on the questionnaire that are measured by the accelerometer, including climbing stairs, walking within shops, playing with domestic animals, and dressing and cleaning oneself. These kinds of activities could add up to over 1 hour per day. An additional explanation for the underestimation of PA in the questionnaire could be that some activities are of very short duration (<1 minute), and thus might not be recalled by the participants when filling in the questionnaire, whereas the accelerometer accumulates all these small bouts of activity. These two explanations might be especially important for light activity, for which we found that in the total population, weekly light PA was 708 minutes/week (corresponding to 1.75 hours per day) lower according to questionnaires compared to accelerometers (Table S3), which is in line with previous evidence.^{36,37}

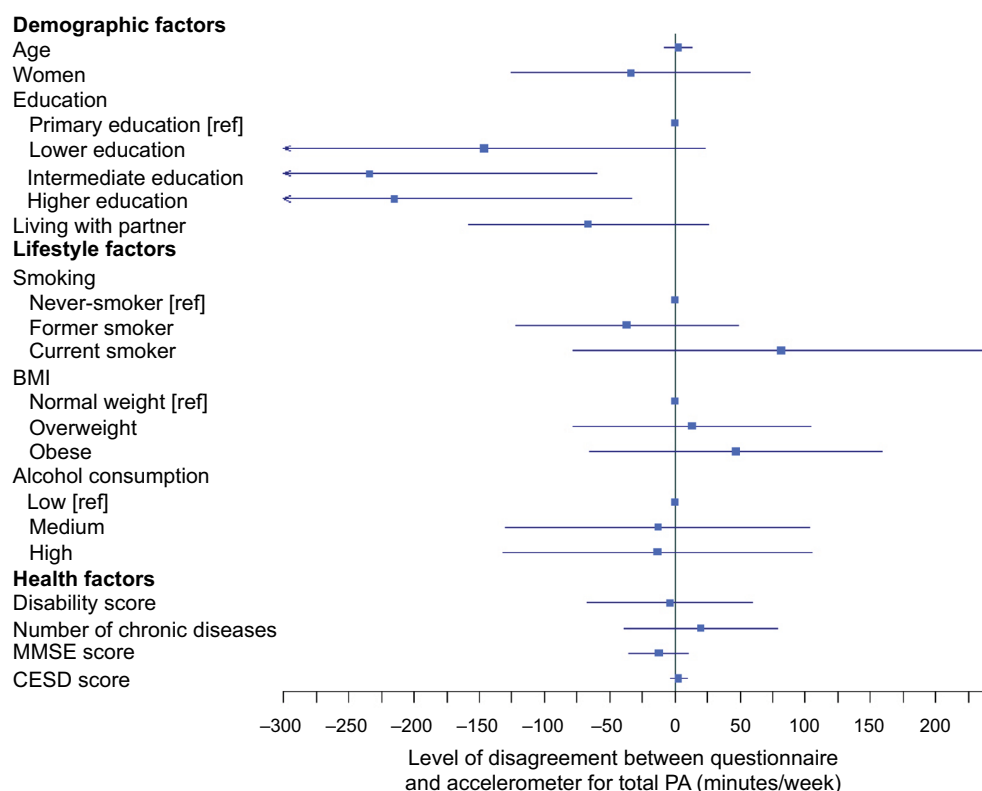


Figure 3 Factors associated with the direction of disagreement between accelerometer and questionnaire physical activity.

Notes: Considering the average direction of disagreement equaled -529.4 minutes/week, positive disagreement indicates less underestimation in the questionnaire, relative to the accelerometer. Negative disagreement indicates more underestimation in the questionnaire compared to the accelerometer. Analyses included all demographic, lifestyle, and health variables, total accelerometer-derived physical activity, time between interview date and first accelerometer-wear date, cohort, and wave. The coefficient for women was relative to men and the coefficient for paid occupation relative to those without paid occupation. For continuous variables (age, disability score, number of chronic diseases, MMSE-score, and CESD-score), values correspond with a 1-unit increase in the determinant.

Abbreviations: BMI, body-mass index; CESD, Center for Epidemiologic Studies Depression; MMSE, Mini-Mental State Examination; PA, physical activity.

To our knowledge, we are the first to report that more depressive symptoms and disability are associated with larger underestimation in the questionnaire compared to the accelerometer. An explanation could be that those with depressive symptoms or disability have a more pessimistic estimate of their PA levels than their counterparts. In other research domains, depressive symptoms have also been associated with reporting bias,³⁸ and it has been suggested that individuals suffering from depression process information about the self in a maladaptive fashion compared to their healthier counterparts.³⁹ For example, in a study on the agreement between self-reported and actigraphy-assessed sleep, participants with more depressive symptoms reported lower sleep duration than was measured using actigraphy.³⁴ In the same study, higher levels of functional disability were associated with larger absolute differences between self-reported and actigraphic sleep, indicating that being disabled might affect self-reported levels in domains other than PA as well.

Furthermore, our findings showed that the absolute difference between questionnaire and accelerometer was highest in

those with primary education, although there was no difference between educational groups regarding the direction of this disagreement. This indicates that the degree of underreporting differs for participants with a low education, which is canceled out in the absolute difference between questionnaire and accelerometer. Finally, obese individuals underreported their total PA more than normal-weight participants, which is contradictory to previous research showing overreporting of PA in questionnaires.^{15,40} However, in line with our observation, a study among 365 young and middle-aged adults found that participants who underestimated their PA levels were more likely to have a higher BMI.⁴¹ In this regard, it is possible that those with obesity perceive their PA levels as inadequate, because of the belief that they could not have been obese with adequate PA levels.^{41,42}

The results of our study may have implications for current and future practice on PA measurements. Considering the increase in the use of accelerometers in current research, researchers need to be aware that results between studies using either questionnaires or accelerometers should be compared

with caution. Importantly, the disagreement increased with higher levels of accelerometer-derived PA. More detailed instructions in the questionnaire may be needed,⁴³ and questionnaires might need to be improved to capture the overall PA pattern of older adults better. To this aim, future studies should focus on enhancing the agreement between questionnaire- and accelerometer-derived PA levels. Moreover, individuals working in clinical practice should be aware that especially those with worse perceived health might not have a realistic view of their PA levels. Asking more thoroughly about their activities could provide more information than a simple question on whether they perceive their PA as sufficient or not. Furthermore, the large discrepancy between accelerometer- and questionnaire-derived PA in those with paid occupations found in our sensitivity analysis implies that studies using questionnaires in working populations should always include occupational PA.

The use of accelerometers could also be improved in future studies to measure activities performed solely with the lower extremities more accurately. The bias between questionnaire and accelerometer might be reduced by future efforts to develop specified algorithms that can identify these activities, taking into account their specific rhythm when wearing an accelerometer. Moreover, our findings that (perceived) health and educational status can influence reported PA stresses the importance of taking into account these factors in future studies and analyses. Finally, regardless of the (dis)agreement between questionnaires and accelerometers, both methods should be used simultaneously in research to gain the most information. Whereas accelerometers provide an accurate distribution of intensity and duration, questionnaires might more accurately reflect perceived PA. In order to be able to influence the PA levels of individuals, both sources of information are important. We would thus recommend that for all individuals, including those with higher BMI, more depressive symptoms, and more disability, PA be measured using both methods whenever possible.

Our study contains several strengths. We are one of the first to examine both the correlation and disagreement between PA as derived from questionnaire and a wrist-worn accelerometer in a large community-based cohort of older adults, using a wide variety of determinants. Moreover, the use of a waterproof wrist-worn accelerometer ensured high compliance. Consequently, this ensured low nonwear periods in which assumptions would have to be made on whether this time was spent actively or sedentarily.^{44,45} As a result, PA is generally assessed more precisely.⁴⁵

However, some limitations also have to be acknowledged. First, the PA estimates derived from the accelerometer and

questionnaire were not obtained in the same week. This might have contributed to the low agreement in the current study. However, addressing questionnaires and accelerometers simultaneously might make participants more aware of their behavior and result in larger agreement than otherwise obtained.⁴³ Second, we did not have up-to-date information on cancer, cardiovascular disease, or diabetes. Consequently, some residual confounding might have been present, which can result in bias towards the null. Third, our results are based on the difference between the LAPAQ and the wrist-worn GeneActiv accelerometer. As such, our findings might not be generalizable to other questionnaires or hip-worn accelerometers. Fourth, using accelerometers providing raw output in milligravity is relatively new, and thus the cutoff points to define activity intensity have not been firmly established. Therefore, the use of different cutoffs might shift the distribution of time spent in light, moderate, and vigorous PA according to the accelerometer. Furthermore, we measured PA during 1 week, which might not represent overall engagement in PA. Additionally, participants that agreed to wear an accelerometer were more often men and showed a better health profile regarding MMSE score, CESD score, and disability score. This might affect generalizability toward the total population of older adults. Finally, the current participants were from a Dutch population, with relatively high levels of PA compared to other Western countries.⁴⁶ Considering the fact that the disagreement in the current study was higher for higher levels of accelerometer-derived PA, the disagreement between questionnaire- and accelerometer-derived PA might be lower in countries with generally lower PA levels.

In conclusion, we observed that correspondence between questionnaire- and accelerometer-derived PA is low. In the total population, the number of minutes per week spent in PA was in general underestimated in the questionnaire. This was a consequence of underreporting light PA, whereas moderate PA was overreported in the questionnaire. These differences were partly explained by BMI, disability, and depressive symptoms, as participants with higher BMI and more disability and depressive symptoms underestimated their self-reported PA more than their healthier counterparts. Moreover, the underestimation was larger with higher levels of accelerometer-derived PA. Future studies using questionnaires and/or accelerometers to measure PA should be aware of these differences.

Data sharing statement

Data can be obtained upon request. Requests should be directed to the management team of the Rotterdam Study

(secretariat.epi@erasmusmc.nl), which has a protocol for approving data requests. Because of restrictions based on privacy regulations and informed consent of the participants, data cannot be made freely available in a public repository.

Acknowledgments

This work was supported by a Netherlands Organization for Scientific Research grant (017.106.370) awarded to HT. The Rotterdam Study is funded by the Erasmus Medical Center and Erasmus University Rotterdam, the Netherlands Organization for the Health Research and Development, the Research Institute for Diseases in the Elderly, the Ministry of Education, Culture, and Science, the Ministry for Health, Welfare, and Sports, and the European Commission. We thank the staff of the Rotterdam Study and the participating general practitioners and pharmacists.

Author contributions

CMK, JDS, and OHF had the original idea for the study. CMK participated in analyses and interpretation of the data and drafted the article. All authors contributed toward data analysis, drafting and critically revising the manuscript for important intellectual content, gave final approval of the version to be published, and agree to be accountable for all aspects of the work.

Disclosure

CMK, JDS, OHF and MC work in ErasmusAGE, a center within the Erasmus Medical Center for aging research across the life course funded by Nestlé Nutrition (Nestec Ltd.), Metagenics Inc., and AXA. The authors report no conflicts of interest in this work.

References

- World Health Organization. *Global Status Report on Noncommunicable Diseases 2010*. Geneva: WHO; 2010.
- Skender S, Ose J, Chang-Claude J, et al. Accelerometry and physical activity questionnaires: a systematic review. *BMC Public Health*. 2016;16:515.
- Sallis JF, Saelens BE. Assessment of physical activity by self-report: status, limitations, and future directions. *Res Q Exerc Sport*. 2000;71 Suppl 2:1–14.
- Schmid D, Ricci C, Leitzmann MF. Associations of objectively assessed physical activity and sedentary time with all-cause mortality in US adults: the NHANES study. *PLoS One*. 2015;10(3):e0119591.
- Ensrud KE, Blackwell TL, Cauley JA, et al. Objective measures of activity level and mortality in older men. *J Am Geriatr Soc*. 2014;62(11):2079–2087.
- Westertorp KR. Physical activity assessment with accelerometers. *Int J Obes Relat Metab Disord*. 1999;23 Suppl 3:S45–S49.
- Scheers T, Philippaerts R, Lefevre J. Assessment of physical activity and inactivity in multiple domains of daily life: a comparison between a computerized questionnaire and the SenseWear armband complemented with an electronic diary. *Int J Behav Nutr Phys Act*. 2012;9:71.
- Prince SA, Adamo KB, Hamel ME, Hardt J, Gorber SC, Tremblay M. A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. *Int J Behav Nutr Phys Act*. 2008;5:56.
- Sabia S, van Hees VT, Shipley MJ, et al. Association between questionnaire- and accelerometer-assessed physical activity: the role of sociodemographic factors. *Am J Epidemiol*. 2014;179(6):781–790.
- Rosenberger ME, Haskell WL, Albinali F, Mota S, Nawyn J, Intille S. Estimating activity and sedentary behavior from an accelerometer on the hip or wrist. *Med Sci Sports Exerc*. 2013;45(5):964–975.
- Swartz AM, Strath SJ, Bassett DR Jr, O'Brien WL, King GA, Ainsworth BE. Estimation of energy expenditure using CSA accelerometers at hip and wrist sites. *Med Sci Sports Exerc*. 2000;32(9 Suppl):S450–S456.
- Chen KY, Bassett DR Jr. The technology of accelerometry-based activity monitors: current and future. *Med Sci Sports Exerc*. 2005;37(11 Suppl):S490–500.
- Shiroma EJ, Cook NR, Manson JE, Buring JE, Rimm EB, Lee IM. Comparison of self-reported and accelerometer-assessed physical activity in older women. *PLoS One*. 2015;10(12):e0145950.
- Miles J, Shevlin M. *Applying Regression and Correlation. A Guide for Students and Researchers*. London: Sage; 2001.
- Slootmaker SM, Schuit AJ, Chinapaw MJ, Seidell JC, van Mechelen W. Disagreement in physical activity assessed by accelerometer and self-report in subgroups of age, gender, education and weight status. *Int J Behav Nutr Phys Act*. 2009;6:17.
- Dong L, Block G, Mandel S. Activities contributing to total energy expenditure in the United States: results from the NHAPS study. *Int J Behav Nutr Phys Act*. 2004;1(1):4.
- Ikram MA, Brusselle GGO, Murad SD, et al. The Rotterdam Study: 2018 update on objectives, design and main results. *Eur J Epidemiol*. 2017;32(9):807–850.
- Stel VS, Smit JH, Pluijm SM, Visser M, Deeg DJ, Lips P. Comparison of the LASA physical activity questionnaire with a 7-day diary and pedometer. *J Clin Epidemiol*. 2004;57(3):252–258.
- Koolhaas CM, Dhana K, van Rooij FJ, Schoufour JD, Hofman A, Franco OH. Physical activity types and health-related quality of life among middle-aged and elderly adults: the Rotterdam study. *J Nutr Health Aging*. Epub 2017 Mar 9.
- Ainsworth BE, Haskell WL, Herrmann SD, et al. 2011 Compendium of Physical Activities: a second update of codes and MET values. *Med Sci Sports Exerc*. 2011;43(8):1575–1581.
- Haskell WL, Lee IM, Pate RR, et al. Physical activity and public health: updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. *Circulation*. 2007;116(9):1081–1093.
- Hildebrand M, van Hees VT, Hansen BH, Ekelund U. Age group comparability of raw accelerometer output from wrist- and hip-worn monitors. *Med Sci Sports Exerc*. 2014;46(9):1816–1824.
- da Silva IC, van Hees VT, Ramires VV, et al. Physical activity levels in three Brazilian birth cohorts as assessed with raw triaxial wrist accelerometry. *Int J Epidemiol*. 2014;43(6):1959–1968.
- White T, Westgate K, Wareham NJ, Brage S. Estimation of physical activity energy expenditure during free-living from wrist accelerometry in UK adults. *PLoS One*. 2016;11(12):e0167472.
- White T. Pampro: physical activity monitor processing. 2016. Available from: <https://github.com/thomite/pampro>. Accessed December 1, 2016.
- Brage S, Westgate K, Wijndaele K, Godinho J, Griffin S, Wareham N. Evaluation of a method for minimizing diurnal information bias in objective sensor data. Poster presented at: 3rd International Conference on Ambulatory Monitoring of Physical Activity and Movement; June 17–19, 2013; Amherst, MA.
- International Standard Classification of Education (ISCED). United Nations Educational, S.a.C.O.U. Division of Statistics and Education; Office of Statistics; UNESCO: 1976. Available from: <https://unstats.un.org/unsd/iiss/International-Standard-Classification-of-Education-ISCED.ashx>.
- Fries JF, Spitz P, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. *Arthritis Rheum*. 1980;23(2):137–145.

29. Odding E, Valkenburg HA, Stam HJ, Hofman A. Determinants of locomotor disability in people aged 55 years and over: the Rotterdam Study. *Eur J Epidemiol*. 2001;17(11):1033–1041.
30. Radloff LS. The CES-D scale a self-report depression scale for research in the general population. *Appl Psychol Meas*. 1977;1(3):385–401.
31. McDowell I. *Measuring Health: A Guide to Rating Scales and Questionnaires*. Oxford: Oxford University Press; 2006.
32. Folstein MF, Folstein SE, McHugh PR. “Mini-mental state”: a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*. 1975;12(3):189–198.
33. Kukull WA, Larson EB, Teri L, Bowen J, McCormick W, Pfanschmidt ML. The Mini-Mental State Examination score and the clinical diagnosis of dementia. *J Clin Epidemiol*. 1994;47(9):1061–1067.
34. Van Den Berg JF, Van Rooij FJ, Vos H, et al. Disagreement between subjective and actigraphic measures of sleep duration in a population-based study of elderly persons. *J Sleep Res*. 2008;17(3):295–302.
35. Rosenberger ME, Haskell WL, Albinali F, Mota S, Nawyn J, Intille S. Estimating activity and sedentary behavior from an accelerometer on the hip or wrist. *Med Sci Sports Exerc*. 2013;45(5):964–975.
36. España-Romero V, Golubic R, Martin KR, et al. Comparison of the EPIC physical activity questionnaire with combined heart rate and movement sensing in a nationally representative sample of older British adults. *PLoS One*. 2014;9(2):e87085.
37. Peters TM, Shu XO, Moore SC, et al. Validity of a physical activity questionnaire in Shanghai. *Med Sci Sports Exerc*. 2010;42(12):2222–2230.
38. Korn CW, Sharot T, Walter H, Heekeren HR, Dolan RJ. Depression is related to an absence of optimistically biased belief updating about future life events. *Psychol Med*. 2014;44(3):579–592.
39. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Arlington (VA): APA; 2013.
40. Warner ET, Wolin KY, Duncan DT, Heil DP, Askew S, Bennett GG. Differential accuracy of physical activity self-report by weight status. *Am J Health Behav*. 2012;36(2):168–178.
41. Watkinson C, van Sluijs EM, Sutton S, Hardeman W, Corder K, Griffin SJ. Overestimation of physical activity level is associated with lower BMI: a cross-sectional analysis. *Int J Behav Nutr Phys Act*. 2010;7:68.
42. van Sluijs EM, Griffin SJ, van Poppel MN. A cross-sectional study of awareness of physical activity: associations with personal, behavioral and psychosocial factors. *Int J Behav Nutr Phys Act*. 2007;4:53.
43. Lee PH, Yu YY, McDowell I, Leung GM, Lam TH, Stewart SM. Performance of the international physical activity questionnaire (short form) in subgroups of the Hong Kong Chinese population. *Int J Behav Nutr Phys Act*. 2011;8:81.
44. Zhang S, Rowlands AV, Murray P, Hurst TL. Physical activity classification using the Geneva wrist-worn accelerometer. *Med Sci Sports Exerc*. 2012;44(4):742–748.
45. Dieu O, Mikulovic J, Fardy PS, Bui-Xuan G, Beghin L, Vanhelst J. Physical activity using wrist-worn accelerometers: comparison of dominant and non-dominant wrist. *Clin Physiol Funct Imaging*. 2017;37(5):525–529.
46. Sjöström M, Oja P, Hagströmer M, Smith BJ, Bauman A. Health-enhancing physical activity across European Union countries: the Eurobarometer study. *J Public Health*. 2006;14(5):291–300.

Supplementary materials

Table S1 Characteristics of those who agreed to participate (N=3,156) and those who did not agree to participate (N=833)

| Characteristics | Did not agree to participate | Agreed to participate* | P |
|---|------------------------------|------------------------|--------|
| Age (years) | 72.7 (22.7) | 70.8 (9.0) | <0.001 |
| Women, n (%) | 528 (63.6) | 1,669 (52.9) | <0.001 |
| Education | | | <0.001 |
| Elementary | 66 (8.1) | 197 (6.3) | |
| Lower secondary | 374 (46.0) | 1,202 (38.5) | |
| Higher secondary | 239 (29.4) | 966 (30.9) | |
| Tertiary | 134 (16.5) | 757 (24.2) | |
| BMI, n (%) | | | 0.06 |
| Normal weight | 258 (31.3) | 880 (27.9) | |
| Overweight | 358 (43.4) | 1,507 (47.8) | |
| Obese | 209 (25.3) | 768 (24.3) | |
| Smoking, n (%) | | | 0.002 |
| Nonsmoker | 296 (35.8) | 1,048 (33.3) | |
| Former smoker | 424 (51.3) | 1,800 (57.2) | |
| Current smoker | 107 (12.9) | 301 (9.6) | |
| Disability score | 0.75 (0.87) | 0.55 (0.65) | <0.001 |
| Living with partner, n (%) | 508 (61.4) | 2,261 (71.7) | <0.001 |
| Currently with job, n (%) | 120 (16.7) | 599 (20.6) | 0.02 |
| Prevalent cancer, CVD, or diabetes, n (%) | 352 (42.3) | 1,276 (40.4) | 0.36 |
| MMSE score | 27.51 (2.40) | 28.12 (1.74) | <0.001 |
| CESD score | 6.44 (7.30) | 5.35 (6.79) | <0.001 |
| Total PA according to questionnaire (minutes/week) | 970.2 (1,176.5) | 1,110.1 (1,486.5) | 0.02 |

Notes: *The study sample of those agreeing to participate also includes participants that wore an Actiwatch device (Actiwatch model AW4; Cambridge Technology, Cambridge, UK). Data presented as mean (SD) unless otherwise stated.

Abbreviations: BMI, body-mass index; CESD, Center for Epidemiologic Studies Depression; CVD, cardiovascular disease; MMSE, Mini-Mental State Examination; PA, physical activity; SD, standard deviation.

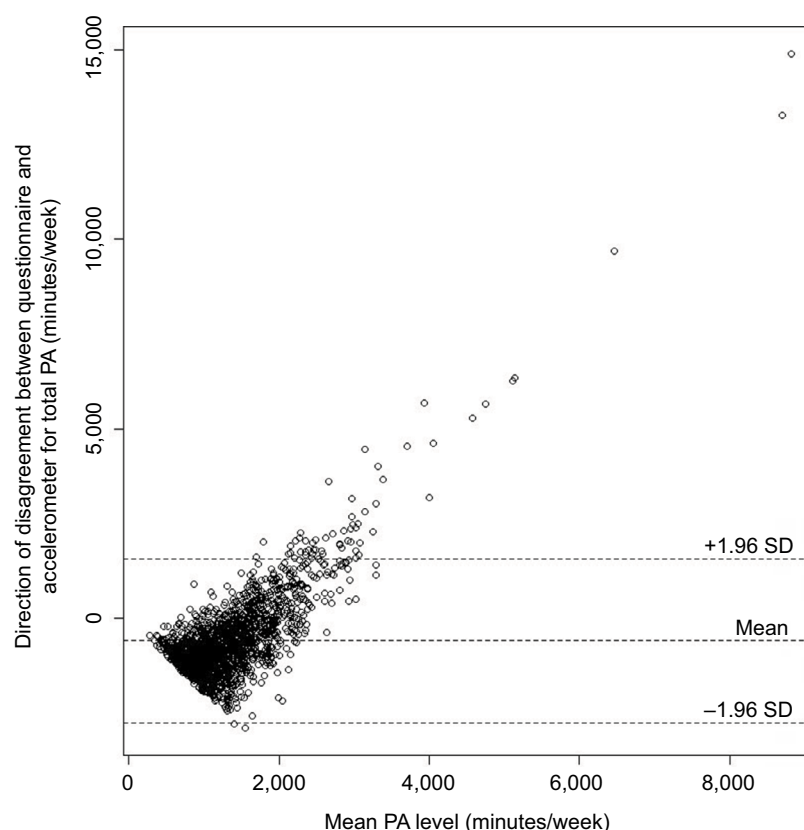


Figure S1 Bland–Altman plot for total physical activity.

Abbreviations: PA, physical activity; SD, standard deviation.

Table S2 Characteristics of light, moderate, and vigorous PA in the total population

| Characteristics | Mean (SD) |
|--|-----------------|
| PA according to questionnaire | |
| Light PA (minutes) | 316.1 (473.1) |
| Moderate PA (minutes) | 651.8 (800.8) |
| Vigorous PA (minutes) | 82.0 (338.7) |
| PA according to accelerometer | |
| Light PA (minutes) | 1,024.4 (218.1) |
| Moderate PA (minutes) | 447.3 (139.2) |
| Vigorous PA (minutes) | 107.6 (53.7) |
| Level of disagreement between questionnaire and accelerometer | |
| Light PA (minutes) | 781.6 (354.1) |
| Moderate PA (minutes) | 405.8 (697.6) |
| Vigorous PA (minutes) | 131.0 (312.7) |
| Direction of disagreement between questionnaire and accelerometer | |
| Light PA (minutes) | -708.4 (484.3) |
| Moderate PA (minutes) | 204.6 (780.8) |
| Vigorous PA (minutes) | -25.6 (338.1) |

Note: Paired t-tests indicated significant differences between light, moderate, and vigorous physical activity as measured by questionnaire and accelerometer.

Abbreviations: PA, physical activity; SD, standard deviation

Table S3 Factors associated with the level of disagreement between accelerometer and questionnaire PA, according to intensity of PA

| Factors | Light PA | | | Moderate PA | | | Vigorous PA | | |
|----------------------------|----------|-----------------|--------|-------------|----------------|------|-------------|-----------------|-------|
| | β | 95% CI | P | β | 95% CI | P | β | 95% CI | P |
| Demographic factors | | | | | | | | | |
| Age | 3.9 | -0.82 to 8.5 | 0.11 | -5.3 | -15.2 to 4.5 | 0.29 | -3.9 | -8.34 to 0.47 | 0.08 |
| Women | -115.0 | -156.3 to -73.7 | <0.001 | -20.2 | -107.4 to 7.0 | 0.65 | -43.4 | -82.3 to -4.6 | 0.03 |
| Education | | | | | | | | | |
| Primary | Ref | | | Ref | | | Ref | | |
| Lower | -18.1 | -95.1 to 59.0 | 0.65 | -66.6 | -228.2 to 95.1 | 0.42 | -111.3 | -183.5 to -39.0 | 0.003 |
| Intermediate | -64.4 | -143.5 to 14.6 | 0.11 | -116.6 | -282.7 to 49.6 | 0.17 | -121.7 | -195.9 to -47.5 | 0.001 |
| Higher | -50.9 | -133.4 to 31.6 | 0.23 | -113.6 | -286.6 to 59.5 | 0.20 | -114.3 | -191.7 to -37.0 | 0.004 |
| Living with partner | -4.1 | -45.7 to 37.5 | 0.85 | -64.4 | -152.0 to 23.3 | 0.15 | -34.0 | -73.2 to 5.2 | 0.09 |
| Lifestyle factors | | | | | | | | | |
| Smoking | | | | | | | | | |
| Never-smokers | Ref | | | Ref | | | Ref | | |
| Former smokers | -10.6 | -49.1 to 28.0 | 0.59 | 4.2 | -77.1 to 85.4 | 0.92 | -25.3 | -61.6 to 11.03 | 0.17 |
| Current smokers | 17.3 | -54.8 to 89.4 | 0.64 | 103.3 | -48.6 to 255.3 | 0.18 | -51.8 | -119.8 to 16.2 | 0.14 |
| BMI | | | | | | | | | |
| Normal weight | Ref | | | Ref | | | Ref | | |
| Overweight | 0.59 | -40.6 to 41.8 | 0.98 | -42.1 | -128.9 to 44.6 | 0.34 | 15.4 | -23.3 to 54.2 | 0.44 |
| Obese | 50.0 | -0.6 to 100.7 | 0.053 | -78.5 | -185.2 to 28.3 | 0.15 | -16.3 | -64.0 to 31.5 | 0.50 |
| Alcohol consumption | | | | | | | | | |
| Low | Ref | | | Ref | | | Ref | | |
| Medium | 0.9 | -47.4 to 49.2 | 0.97 | 48.4 | -61.2 to 158.0 | 0.39 | -14.9 | -58.8 to 29.0 | 0.51 |
| High | 7.0 | -44.9 to 58.9 | 0.79 | 27.7 | -83.0 to 138.4 | 0.62 | -10.0 | -55.6 to 35.6 | 0.67 |
| Health factors | | | | | | | | | |
| Disability score | 29.2 | 0.60 to 57.8 | 0.05 | -41.3 | -101.6 to 19.0 | 0.18 | -26.9 | -53.8 to -0.01 | 0.05 |
| Number of chronic diseases | 0.4 | -26.3 to 27.09 | 0.98 | 1.5 | -54.7 to 57.8 | 0.96 | 0.2 | -25.0 to 25.4 | 0.99 |
| MMSE score | -7.4 | -17.9 to 2.99 | 0.16 | 11.7 | -10.4 to 33.8 | 0.30 | -2.0 | -11.79 to 7.8 | 0.69 |
| CESD score | 2.9 | -0.05 to 5.76 | 0.05 | -5.3 | -11.4 to 0.82 | 0.09 | 0.7 | -2.02 to 3.46 | 0.61 |

Note: Analyses were adjusted for all included demographic, lifestyle, and health variables, total accelerometer-derived physical activity, time between interview date and first accelerometer-wear date, cohort, and wave.

Abbreviations: BMI, body-mass index; CESD, Center for Epidemiologic Studies Depression; CI, confidence interval; MMSE, Mini-Mental State Examination; PA, physical activity; ref, reference.

Table S4 Factors associated with the direction of disagreement between accelerometer and questionnaire PA according to intensity of PA

| Factors | Light PA | | | Moderate PA | | | Vigorous PA | | |
|----------------------------|----------|----------------|--------|-------------|-----------------|-------|-------------|-----------------|------|
| | β | 95% CI | P | β | 95% CI | P | β | 95% CI | P |
| Demographic factors | | | | | | | | | |
| Age | -3.2 | -9.8 to 3.3 | 0.33 | -5.7 | -16.7 to 5.3 | 0.31 | -1.5 | -6.3 to 3.3 | 0.54 |
| Women | 177.1 | 119.3 to 234.9 | <0.001 | -22.0 | -119.8 to 75.8 | 0.66 | -52.7 | -94.8 to -10.6 | 0.01 |
| Education | | | | | | | | | |
| Primary | Ref | | | Ref | | | Ref | | |
| Lower | 65.9 | -41.3 to 173.0 | 0.23 | -64.9 | -246.1 to 116.3 | 0.48 | -90.9 | -169.3 to -12.5 | 0.02 |
| Intermediate | 71.1 | -38.9 to 181.0 | 0.21 | -133.7 | -319.9 to 52.4 | 0.16 | -91.7 | -172.2 to -11.1 | 0.03 |
| Higher | 81.6 | -33.2 to 196.3 | 0.16 | -138.6 | -332.6 to 55.5 | 0.16 | -93.6 | -177.6 to -9.7 | 0.03 |
| Living with partner | 68.5 | 10.4-126.5 | 0.02 | -61.5 | -159.8 to 36.8 | 0.22 | -26.6 | -69.1 to 16.0 | 0.22 |
| Lifestyle factors | | | | | | | | | |
| Smoking | | | | | | | | | |
| Never-smokers | Ref | | | Ref | | | Ref | | |
| Former smokers | -18.0 | -71.8 to 35.8 | 0.51 | -0.3 | -91.3 to 90.8 | 1.00 | -28.3 | -67.7 to 11.14 | 0.16 |
| Current smokers | -26.4 | -127 to 74.2 | 0.61 | 102.5 | -67.9 to 272.9 | 0.24 | -69.5 | -143.2 to 4.2 | 0.06 |
| BMI | | | | | | | | | |
| Normal weight | Ref | | | Ref | | | Ref | | |
| Overweight | -8.2 | -65.7 to 49.2 | 0.78 | -76.2 | -173.5 to 21.0 | 0.12 | 24.4 | -17.7 to 66.4 | 0.26 |
| Obese | -76.0 | -146.6 to -5.4 | 0.04 | -147.4 | -267.0 to -27.8 | 0.02 | -11.0 | -62.8 to 40.9 | 0.68 |
| Alcohol consumption | | | | | | | | | |
| Low | Ref | | | Ref | | | Ref | | |
| Medium | -20.2 | -91.6 to 51.2 | 0.58 | 47.9 | -70.6 to 166.5 | 0.43 | -14.8 | -62.9 to 33.3 | 0.55 |
| High | 5.4 | -68.4 to 79.2 | 0.89 | 42.8 | -81.7 to 167.3 | 0.50 | -10.7 | -59.3 to 37.9 | 0.67 |
| Health factors | | | | | | | | | |
| Disability score | -43.4 | -83.3 to -3.4 | 0.03 | -76.8 | -144.4 to -9.2 | 0.03 | -37.7 | -67.0 to -8.5 | 0.01 |
| Number of chronic diseases | 8.2 | -29.1 to 45.4 | 0.67 | -7.5 | -70.6 to 55.5 | 0.82 | 0.9 | -26.4 to 28.2 | 0.95 |
| MMSE score | 8.3 | -6.25 to 22.8 | 0.26 | 12.5 | -12.3 to 37.3 | 0.32 | -1.7 | -12.4 to 8.9 | 0.75 |
| CESD score | -2.8 | -6.8 to 1.3 | 0.18 | -9.6 | -16.5 to -2.73 | 0.006 | 0.3 | -2.71 to 3.23 | 0.86 |

Note: Analyses were adjusted for all included demographic, lifestyle, and health variables, total accelerometer-derived physical activity, time between interview date and the first accelerometer-wear date, cohort, and wave.

Abbreviations: BMI, body-mass index; CESD, Center for Epidemiologic Studies Depression; CI, confidence interval; MMSE, Mini-Mental State Examination; PA, physical activity; ref, reference.

Table S5 Characteristics of the study population, stratified by occupational status

| Characteristics | Without paid occupation | With paid occupation* | P |
|---|-------------------------|-----------------------|--------|
| Participants | 1,410 | 386 | |
| Age (years) | 73.81 (7.61) | 60.31 (6.17) | <0.001 |
| Female, n (%) | 742 (52.6) | 174 (45.1) | 0.01 |
| Education | | | <0.001 |
| Elementary | 85 (6.1) | 15 (3.9) | |
| Lower secondary | 562 (40.5) | 94 (24.5) | |
| Higher secondary | 428 (30.8) | 120 (31.2) | |
| Tertiary | 314 (22.6) | 155 (40.4) | |
| BMI, n (%) | | | 0.70 |
| Normal weight | 418 (29.6) | 111 (28.8) | |
| Overweight | 668 (47.4) | 192 (49.7) | |
| Obese | 324 (23.0) | 83 (21.5) | |
| Smoking, n (%) | | | 0.005 |
| Never-smoker | 471 (33.4) | 115 (29.8) | |
| Former smoker | 832 (59.0) | 222 (57.5) | |
| Current smoker | 106 (7.5) | 49 (12.7) | |
| Disability score | 0.64 (0.68) | 0.25 (0.38) | <0.001 |
| Living with partner, n (%) | 988 (70.1) | 323 (83.7) | <0.001 |
| Prevalent cancer, CVD or diabetes, n (%) | 635 (45.0) | 79 (20.5) | <0.001 |
| MMSE score | 28.10 (1.74) | 28.61 (1.28) | <0.001 |
| CESD score | 5.04 (6.30) | 4.24 (5.74) | 0.03 |
| Total PA according to questionnaire (minutes/week) | 1,049.9 (1,108.3) | 970.9 (1,149.9) | 0.22 |
| Total PA according to accelerometer (minutes/week) | 1,579.3 (386.1) | 1,779.2 (396.7) | <0.001 |

Notes: *The reason that the number of participants with paid occupation in the table is lower than the 389 participants mentioned in Figure 1, is because a second observation was excluded for three additional participants. Data are presented as mean (SD) unless otherwise stated.

Abbreviations: BMI, body-mass index; CESD, Center for Epidemiologic Studies Depression; CVD, cardiovascular disease; IQR, interquartile range; MMSE, Mini-Mental State Examination; PA, physical activity; SD, standard deviation.

Table S6 Factors associated with level and direction of disagreement between accelerometer and questionnaire PA, including participants with paid occupation (n=1,823)

| Factors | Level of disagreement | | | Direction of disagreement | | |
|----------------------------|-----------------------|-----------------|-------|---------------------------|------------------|--------|
| | β | 95% CI | P | β | 95% CI | P |
| Demographic factors | | | | | | |
| Age | 0.8 | -8.8 to 10.5 | 0.86 | -8.8 | -22.4 to 4.8 | 0.21 |
| Women | -66.4 | -149.3 to 16.5 | 0.12 | 120.0 | 2.6 to 237.4 | 0.05 |
| Education | | | | | | |
| Primary | Ref | | | Ref | | |
| Lower | -176.1 | -340.2 to -12.1 | 0.04 | -59.9 | -291.8 to 171.9 | 0.61 |
| Intermediate | -268.6 | -436.3 to -101 | 0.002 | -115.1 | -352.0 to 121.8 | 0.34 |
| Higher | -241.5 | -413.9 to -69.0 | 0.01 | -101.9 | -345.4 to 141.6 | 0.41 |
| Paid occupation | 185 | 72.7 to 297.3 | 0.001 | -312.7 | -471.5 to -154 | <0.001 |
| Living with partner | -44.4 | -131.9 to 43.1 | 0.32 | 1.0 | -122.7 to 124.8 | 0.99 |
| Lifestyle factors | | | | | | |
| Smoking | | | | | | |
| Never-smoker | Ref | | | Ref | | |
| Former smoker | -2.5 | -82.2 to 77.3 | 0.95 | -18.3 | -131.0 to 94.5 | 0.75 |
| Current smoker | 88.1 | -53.1 to 229.4 | 0.22 | -12.6 | -212.5 to 187.3 | 0.90 |
| BMI | | | | | | |
| Normal weight | Ref | | | Ref | | |
| Overweight | 19.4 | -65.6 to 104.5 | 0.65 | -109 | -229.2 to 11.1 | 0.08 |
| Obese | 49.2 | -55.9 to 154.3 | 0.36 | -260.2 | -408.9 to -111.6 | 0.00 |
| Alcohol consumption | | | | | | |
| Low | Ref | | | Ref | | |
| Medium | -16.3 | -121.5 to 89.0 | 0.76 | -46.3 | -191.7 to 99.2 | 0.53 |
| High | -14.6 | -116.7 to 87.6 | 0.78 | -28.4 | -176.6 to 119.8 | 0.71 |
| Health factors | | | | | | |
| Disability score | -15.6 | -79.3 to 48.1 | 0.63 | -134.7 | -224.7 to -44.6 | 0.00 |
| Number of chronic diseases | 19.9 | -38.2 to 78.0 | 0.50 | -3.4 | -85.44 to 78.7 | 0.94 |
| MMSE score | -4.6 | -27.0 to 17.9 | 0.69 | 21.6 | -10.3 to 53.5 | 0.18 |
| CESD score | 2.4 | -3.7 to 8.5 | 0.44 | -12.5 | -21.1 to -3.95 | 0.004 |

Note: Analyses were adjusted for all included demographic, lifestyle, and health variables, total accelerometer-derived physical activity, time between interview date and the first accelerometer-wear date, cohort, and wave.

Abbreviations: BMI, body-mass index; CESD, Center for Epidemiologic Studies Depression Scale; CI, confidence interval; MMSE, Mini-Mental State Examination; PA, physical activity; ref, reference.

Table S7 Factors associated with the level and direction of disagreement between accelerometer and questionnaire PA in those not reporting cycling (n=692)

| Factors | Level of disagreement | | | Direction of disagreement | | |
|----------------------------|-----------------------|-----------------|------|---------------------------|-----------------|------|
| | β | 95% CI | P | β | 95% CI | P |
| Demographic factors | | | | | | |
| Age | -0.4 | -14.1 to 13.3 | 0.96 | -21.2 | -40.7 to -1.7 | 0.03 |
| Women | -31.6 | -165.1 to 101.9 | 0.64 | 94.5 | -95.1 to 284.1 | 0.33 |
| Education | | | | | | |
| Primary | Ref | | | Ref | | |
| Lower | 3.0 | -216.2 to 222.3 | 0.98 | 186.0 | -124.8 to 496.8 | 0.24 |
| Intermediate | -89.7 | -316.3 to 136.9 | 0.44 | 10.0 | -311.0 to 331.1 | 0.95 |
| Higher | -80.5 | -323.1 to 162.1 | 0.52 | 129.3 | -214.2 to 472.9 | 0.46 |
| Living with partner | -77.8 | -203.9 to 48.3 | 0.23 | -132.5 | -311.2 to 46.2 | 0.15 |
| Lifestyle factors | | | | | | |
| Smoking | | | | | | |
| Never-smoker | Ref | | | Ref | | |
| Former smoker | 8.4 | -113.8 to 130.6 | 0.89 | -15.2 | -188.3 to 158.0 | 0.86 |
| Current smoker | 162.4 | -47.9 to 372.7 | 0.13 | 83.6 | -214.5 to 381.7 | 0.58 |

(Continued)

Table S7 (Continued)

| Factors | Level of disagreement | | | Direction of disagreement | | |
|----------------------------|-----------------------|-----------------|------|---------------------------|-----------------|------|
| | β | 95% CI | P | β | 95% CI | P |
| BMI | | | | | | |
| Normal weight | Ref | | | Ref | | |
| Overweight | -81.4 | -214.0 to 51.2 | 0.23 | -149.3 | -337.1 to 38.6 | 0.12 |
| Obese | -14.9 | -170.0 to 140.3 | 0.85 | -283.7 | -503.4 to -63.9 | 0.01 |
| Alcohol consumption | | | | | | |
| Low | Ref | | | Ref | | |
| Medium | 10.4 | -154.7 to 175.5 | 0.90 | -27.6 | -282.9 to 227.7 | 0.83 |
| High | -42.5 | -224.1 to 139.1 | 0.64 | -26.9 | -287.8 to 234.0 | 0.84 |
| Health factors | | | | | | |
| Disability score | 5.0 | -74.5 to 84.4 | 0.90 | -91.9 | -204.6 to 20.8 | 0.11 |
| Number of chronic diseases | 37.6 | -46.6 to 121.8 | 0.38 | 49.9 | -69.5 to 169.2 | 0.41 |
| MMSE score | 8.3 | -23.3 to 39.8 | 0.61 | 26.0 | -18.7 to 70.7 | 0.25 |
| CESD score | -1.0 | -9.5 to 7.6 | 0.83 | -15.3 | -27.4 to -3.19 | 0.01 |

Note: Analyses were adjusted for all included demographic, lifestyle, and health variables, total accelerometer-derived PA, time between interview date and the first accelerometer-wear date, cohort, and wave.

Abbreviations: BMI, body-mass index; CESD, Center for Epidemiologic Studies Depression Scale; CI, confidence interval; MMSE, Mini-Mental State Examination; PA, physical activity; ref, reference.

Clinical Epidemiology

Publish your work in this journal

Clinical Epidemiology is an international, peer-reviewed, open access, online journal focusing on disease and drug epidemiology, identification of risk factors and screening procedures to develop optimal preventative initiatives and programs. Specific topics include: diagnosis, prognosis, treatment, screening, prevention, risk factor modification,

Submit your manuscript here: <https://www.dovepress.com/clinical-epidemiology-journal>

systematic reviews, risk and safety of medical interventions, epidemiology and biostatistical methods, and evaluation of guidelines, translational medicine, health policies and economic evaluations. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use.

Dovepress