Reliability and Validity of a Domain-Specific Last 7-d Sedentary Time Questionnaire

KATRIEN WIJNDAELE[1], ILSE DE BOURDEAUDHUIJ[2], JOB G. GODINO[1], BRIGID M. LYNCH[3,4], SIMON J. GRIFFIN[1], KATE WESTGATE[1], and SØREN BRAGE[1]

[1]MRC Epidemiology Unit, University of Cambridge, Cambridge, England, UNITED KINGDOM; [2]Department of Movement and Sport Sciences, Ghent University, Ghent, BELGIUM; [3]Physical Activity Laboratory, Baker IDI Heart and Diabetes Institute, Melbourne, AUSTRALIA; [4]Melbourne School of Population and Global Health, Faculty of Medicine, Dentistry and Health Sciences, University of Melbourne, Melbourne, AUSTRALIA

ABSTRACT

WIJNDAELE, K., I. DE BOURDEAUDHUIJ, J. G. GODINO, B. M. LYNCH, S. J. GRIFFIN, K. WESTGATE, and S. BRAGE. Statistical methods for assessing agreement between two methods of clinical measurement. *Med. Sci. Sports Exerc.*, Vol. 46, No. 6, pp. 1248–1260, 2014. **Purpose**: The objective of this study is to examine test–retest reliability, criterion validity, and absolute agreement of a self-report, last 7-d sedentary behavior questionnaire (SIT-Q-7d), which assesses total daily sedentary time as an aggregate of sitting/lying down in five domains (meals, transportation, occupation, nonoccupational screen time, and other sedentary time). Dutch (DQ) and English (EQ) versions of the questionnaire were examined. **Methods**: Fifty-one Flemish adults (ages 39.4 ± 11.1 yr) wore a thigh accelerometer (activPAL3™) and simultaneously kept a domain log for 7 d. The DQ was subsequently completed twice (median test–retest interval: 3.3 wk). Thigh-acceleration sedentary time was log annotated to create comparable domain-specific and total sedentary time variables. Four hundred two English adults (ages 49.6 ± 7.3 yr) wore a combined accelerometer and HR monitor (Actiheart®) for 6 d to objectively measure total sedentary time. The EQ was subsequently completed twice (median test–retest interval: 3.4 wk). In both samples, the questionnaire reference frame overlapped with the criterion measure administration period. All participants had five or more valid days of criterion data, including one or more weekend day. **Results**: Test–retest reliability (intraclass correlation coefficient (95% CI)) was fair to good for *total sedentary time* (DQ: 0.68 (0.50–0.81); EQ: 0.53 (0.44–0.62)) and poor to excellent for *domain-specific sedentary time* (DQ: from 0.36 (0.10–0.57) (meals) to 0.66 (0.46–0.79) (occupation); EQ: from 0.45 (0.35–0.54) (other sedentary time) to 0.76 (0.71–0.81) (meals)). For criterion validity (Spearman rho), significant correlations were found for *total sedentary time* (DQ: 0.52; EQ: 0.22; all $P < 0.001$). Compared with domain-specific criterion variables (DQ), modest-to-strong correlations were found for *domain-specific sedentary time* (from 0.21 (meals) to 0.76 ($P < 0.001$) (screen time)). The questionnaire generally overestimated sedentary time compared with criterion measures. **Conclusion**: The SIT-Q-7d appears to be a useful tool for ranking individuals in large-scale observational studies examining total and domain-specific sitting. **Key Words**: SITTING, ADULT, ACCELEROMETER, LOG, POSTURE, PSYCHOMETRICOPEN-ACCESSTRUE

Sedentary behavior includes all waking day behaviors in a sitting/lying posture expending ≤1.5 times the resting energy demand (35). It has been suggested that sedentary behavior is a highly prevalent, independent chronic disease risk factor (26,39). Most of the evidence for health effects has been based on leisure screen behavior (i.e., TV viewing) or total sitting/sedentary time derived from single self-report measures or objective monitoring (39). Although objective methods enhance precision of health risk estimates for total sedentary time, they are still expensive, difficult to administer to large free-living populations, and do not provide information on the domains in which sitting is accumulated (e.g., for transportation or at work). Determinants of adults' sitting time differ by domain (29), and some domain-specific sedentary behaviors (e.g., TV viewing) may have stronger associations with health outcomes (36). Large-scale observational studies therefore need reliable and valid sedentary behavior questionnaires to examine differential

health effects and to enable identification of determinants of domain-specific sedentary behaviors (29).

Few questionnaires characterize sedentary behavior across all common domains of daily life in the general adult population (19). Instead, they tend to focus on a specific domain (e.g., leisure time (9), workplace sitting (7,10,28)) or a specific population (e.g., working force (6) or elderly (15)). Two recent multidomain sedentary behavior questionnaires did not cover all common domains (e.g., sitting for meals), which may affect total sitting time estimates. More importantly, they require participants to recall sitting on a *usual* week/weekend day and exhibit relatively weak validity compared with accelerometry (25,31). Given their low intensity and habitual nature, sedentary behaviors are difficult to recall. Questionnaires therefore benefit from a short, recent recall frame, such as the last 7 d, which allows for recollection of *specific* rather than *usual* behaviors (27). A last 7-d recall frame also captures more intraindividual variability in sedentary behavior than recall frames that are even shorter, such as the past day (11). Only one questionnaire has so far incorporated sedentary time patterns (10), which have been shown to have independent health associations (20). Finally, self-report questionnaires have higher feasibility for implementation in large studies compared with equally valid interviewer-administered ones.

To date, sedentary behavior questionnaires have mostly been validated in specific populations (e.g., overweight adults (22,31), breast cancer patients (11), and middle-age women (25)), which limit generalizability of their psychometric properties to general adult populations. Furthermore, questionnaires have been compared predominantly with a waist- or hip-mounted accelerometer (6,7,10,12,15,25,31), which have important limitations. Removal of these monitors during sleep and water-based activities results in misclassification between sedentary and nonwear time (40), which may be overcome by waterproof monitors that allow for a $24\text{-}h\cdot d^{-1}$ wear protocol. Second, because these accelerometers do not discriminate between postures, they misclassify standing as sedentary time. Thigh-worn monitors that measure thigh acceleration and position with respect to gravity have been shown to accurately measure posture-based sitting time (17,21,23,32). This method has, however, rarely been used as criterion when validating sedentary behavior questionnaires (11,22). Finally, domain-specific criterion measures, derived from simultaneously collected objective monitor and log data, are also underused. No studies have so far annotated thigh accelerometry-based sedentary time with domain-specific log data, despite the merit of such a criterion when validating domain-specific questionnaires (9,19).

We examined the test–retest reliability, criterion validity, and absolute agreement of a self-report, domain-specific, last 7-d sedentary behavior questionnaire (SIT-Q-7d), which assesses volume and patterns of sedentary time, in a sample of the Flemish and English general adult population. Criterion measures allowing for a 24-h wear protocol were used, providing 1) total and domain-specific sedentary time

and patterns, derived from a combined criterion of thigh accelerometry and a domain log, and 2) total sedentary time derived from individually calibrated combined HR and movement sensing.

METHODS

Sedentary Behavior Questionnaire: SIT-Q-7d

The SIT-Q-7d is a self-administered questionnaire that quantifies time spent sedentary in the last 7 d as well as sleeping/napping time. Sedentary time ("sitting or lying down") is assessed across five different domains covering adults' daily life activities. These include, in this order, 1) *meals* (sum of breakfast, lunch, and dinner), 2) *transportation* (sum of to and from occupation, as part of occupation and getting about apart from occupation), 3) *occupation* (i.e., work, study, and volunteering, sum of two main occupations), 4) *leisure screen time* (sum of watching TV/DVDs/videos, using computer apart from work, and playing sedentary computer games), and 5) *time spent sedentary in other activities* (sum of reading, household tasks, caring for children, grandchildren, elderly, or disabled relatives, hobbies, socializing, listening to music, and other activities). The questionnaire enables calculation of domain-specific and total sedentary time. The first page of the SIT-Q-7d provides general instructions for correct completion of the questionnaire, which also emphasize the importance of entering each period of sitting only once, to avoid double counting of sitting time. Subsequently, these instructions on mutual exclusivity of reported sitting time are repeated throughout the questionnaire, tailored to each domain specifically (e.g., section 2—Meals: "DO NOT INCLUDE time spent eating while watching TV. This is part of section 5 (i.e., screen time and other activities)"). Questions on sleeping/napping, meals, nonoccupational transportation, screen time, and other sedentary activities are queried for weekdays and weekend days separately to account for week and weekend differences. The number of interruptions in sedentary time (standing up or walking somewhere) in occupational and TV viewing sedentary time is also assessed. The SIT-Q-7d was based on the SIT-Q, a domain-specific measure of habitual sedentary behaviors with a reference frame of the past year. The SIT-Q was developed through a three-stage process of expert review, cognitive interviewing, and pilot testing and demonstrates acceptable measurement properties for use in epidemiological studies (24).

The SIT-Q-7d was developed by changing the reference frame from the past year (SIT-Q) to the last 7 d (SIT-Q-7d). Given the shorter reference frame, an introductory question determines whether the amount of sitting during the last 7 d was similar to a typical week on a five-point Likert scale. Furthermore, due to shortening the reference frame, participants are restricted to reporting on two main occupations instead of three. For ease of completion, coding, and analysis, the response format was changed to closed categorical rather than open-ended questions. Sections on sitting during

meals, transportation, and leisure were broken down (e.g., transportation to and from occupation, as part of occupation, and getting about apart from occupation) to lower cognitive demands on participants in terms of presummarizing sitting durations (27). A food-frequency section on snacking while watching TV was added, because this may be an important mediator of the health effects of TV viewing (18). To maximize clarity of questions, appropriateness of response categories, question order, and presentation, both the Dutch and English versions were pilot tested before the reliability and validity study. The questionnaire and recommended processing codes are available in http://www.mrc-epid.cam.ac.uk/research/resources/.

Participants and Design

Flemish sample. Dutch-speaking Flemish adults, ages 20–60 yr, were recruited while visiting a variety of theoretical and practical adult education classes in an adult education school located in a suburb of Ghent, Belgium. A total of 62 adults (40% men) agreed to participate. They self-reported their sex, date of birth, height, weight, and education level on the day of recruitment (day 0). They were instructed how to continuously wear a thigh accelerometer (activPAL3™; PAL Technologies, Glasgow, UK) and simultaneously complete a simple domain log for 7 d while maintaining their normal activities, starting the same evening until their evening class the following week (day 7). On day 7, participants handed in both thigh accelerometer and domain log, before they received the SIT-Q-7d questionnaire for the first time (test), to avoid reference to the log when completing the test questionnaire. The "last 7-d" reference frame of the questionnaire coincided with the 7-d administration of the combined criterion measure (thigh accelerometer and domain log). The test questionnaire was completed during the recess break. During a final visit, 2 wk later (day 21), participants received the retest questionnaire, which was again completed during recess. Participants who were unable to complete their retest questionnaire during the break were instructed to complete it at home the same evening and send it back in a prestamped envelope. Those with a test–retest interval longer than 8 wk were excluded from the analysis. The study was approved by the Ghent University Ethics Committee (reference number EC/2011/1236), and all participants provided written informed consent.

English sample. Participants were recruited from the randomized controlled Diabetes Risk Communication Trial (DRCT) (16). In short, Cambridgeshire residents born between 1950 and 1975 and registered with participating general practices were eligible to participate if they were able to walk unaided and were not pregnant, lactating, or previously diagnosed with diabetes, a terminal illness with a prognosis <1 yr, or a psychotic illness at recruitment. The vast majority (≥90%) of potential participants was eligible for inclusion, and 569 were randomized in the DRCT (16). Participants were instructed to continuously wear a combined HR and movement sensor (Actiheart®, CamNtech Ltd., Papworth, UK) for six days and nights while maintaining their normal activities. They completed a log that indicated the date and time they a) started wearing the monitor, b) removed it (along with the reason) and replaced it again, and c) completed measurement. Participants were asked to complete the SIT-Q-7d immediately after removing the combined sensor and to return all of the materials on the same day by post. After the materials were received, participants were sent a response form that asked them to indicate if they would be willing to complete the questionnaire for a second time (constituting the reliability subsample). Participants in the reliability subsample were sent the SIT-Q-7d again (retest) as soon as their response form was received by the study coordinator (approximately 5-d turnaround time). They were instructed to complete the questionnaire as soon as they received it and to immediately return it in a freepost envelope that was provided. Only those with a test–retest interval between 2 and 8 wk (as determined by the date of completion provided by participants on the questionnaire) were included in the analysis. Ethical approval was obtained from the Cambridgeshire 1 Research Ethics Committee (reference number 10/H0304/78). All participants provided written informed consent.

Criterion Measurement of Total and Domain-Specific Sedentary Time

Flemish sample. The activPAL3™ monitor is a small (35 mm × 53 mm × 7 mm), light weight (15 g) electronic accelerometer worn on the anterior midline of the thigh, one-third in distance from the hip to the knee. Waterproof attachment with medical dressing allows for continuous 24-h multiple day wearing. The monitor uses a micro-electromechanical sensor-based accelerometer to measure gravity and activity-related acceleration of the thigh at a sampling frequency of 20 Hz. With this information, participants' daily activities are first classified as static or dynamic. Sitting/lying and standing postures (i.e., static activity) are discriminated on the basis the angle of the thigh relative to gravity (where approximate thresholds of 32° from horizontal for transition from sedentary to upright posture and 22° for the opposing transition are applied, in combination with a time buffer). Dynamic activity is classified as stepping (activPAL software version 6.0.8). The method is highly valid for measuring sitting time and breaks in sitting time (i.e., transitions from sitting/lying to a nonsitting/lying posture) in adults by comparison with direct observation, both in the laboratory (17,32) and in the free-living setting (21,23). Participants were instructed to only remove the monitor when strictly necessary and were provided with additional dressing to reattach the monitor when needed. They recorded every nonwear episode in the domain log, accompanied by the reason.

For domain-specific annotation of the thigh accelerometry-based sedentary time data, participants simultaneously kept a domain log, adapted from Bouchard's log (2). In this log, every day is broken down into 15-min blocks, using 24 rows

(day hours) and four columns (hour quarters). From a list of 10 categories (1: *sleep*; 2: *meal* ("eat breakfast, lunch, or dinner"); 3: *transportation* ("in whichever mode (by foot, bike, car, bus, train, etc.) and to whichever destination (work, sports club, supermarket, etc.")); 4: *work* ("all tasks done to earn money"); 5: *study* ("educational activities"); 6: volunteering ("work that you do for no pay"); 7: *TV/DVD/ video viewing*; 8: *non–work-related computer use/computer games*; 9: *other* ("sports, reading, household activities (cooking, cleaning,…), caring for children, hobbies, listening to music, etc."); and X: *monitor nonwear*), participants were asked to write the corresponding number in each block or draw a horizontal line in the next blocks if a domain was carried over until there was a change. They were instructed that the log captured information on the domains they resided in, irrespective of whether the activities covered in those domains were performed in an active or sitting manner. They were asked to complete the domain log during the day to minimize recall bias. By annotating the thigh accelerometry-based sitting time with the domain log data, the same domain-specific and overall sedentary time summary variables were calculated from the combined criterion (thigh accelerometry and domain log) and the SIT-Q-7d questionnaire, allowing detailed validity assessment in terms of domains and week/ weekend days. Bouts of ≤15-min nonwear time due to replacing the medical dressing of the monitor, as reported in the domain log, were included in analysis. Days on which participants were provided with or handed in the monitor and log, as well as days with ≥16 min of nonwear time, were considered as nonvalid, given the limited time spent sitting in some domains. Participants with ≥5 valid days, including ≥1 weekend day, were included for the validity analyses.

English sample. The Actiheart® is a small (33-mm diameter × 7-mm thick), light weight (<8 g), waterproof monitor worn on the chest with standard ECG electrodes. It simultaneously measures HR (128 Hz) and uniaxial acceleration (32 Hz) by a piezoelectric element. Data were integrated into 60-s epochs (3). Participants' HR response was individually calibrated with a graded submaximal exercise test (5). Free-living HR data were preprocessed before further analysis (37). Branched equation modeling (4) integrated the HR and acceleration signal into estimates of physical activity energy expenditure (PAEE), which are valid across all intensities from sedentary behavior to vigorous physical activities (13). Participants were asked to change the electrodes every 2 d to ensure high-quality HR signals. Every nonwear episode during the day was recorded in a log sheet accompanied by the reason. Discrimination between awake and sleep time was accomplished by annotating the combined sensing output with self-reported sleep time, and agreement was visually inspected. Potential bias due to patterns in nonwear time, which were imbalanced across the 24-h clock, was minimized by regression against two phase-shifted sine functions. Total sedentary time was defined as awake time spent at an intensity <1.5 times the resting metabolic rate (EE-based criterion), in agreement with sedentary behavior definitions (35). Given

the low weighting of the HR signal at the sedentary end of the intensity spectrum in branched equation modeling (4), an additional total awake sedentary time estimate was derived using only the trunk acceleration signal (ACC-based criterion), implemented as a $<0.075 \text{ m·s}^{-2}$ threshold (30). Participants with a wear time corresponding to five full days or more (i.e., ≥120 total wear hours), including one ore more full weekend day (i.e., ≥24 weekend wear hours), were included in analysis.

Statistical Analysis

For self-reported total sedentary behavior and each of the domains, average time per day was calculated to enable comparison between the 6-d criterion measures and the 7-d questionnaire time frame. Midpoint values were assigned for all categories, and weekdays and weekend days were weighted five- and twofold, respectively, before dividing the sum by 7. Separate weekday and weekend day sedentary variables were also calculated for meals, screen time, TV viewing, nonoccupational computer use (using computer apart from work + playing sedentary computer games), and other sedentary time. Descriptive characteristics (medians and interquartile range) were calculated for all sedentary behavior variables.

Test–retest reliability was examined using single-measure intraclass correlation coefficients (ICC) and 95% confidence intervals (CI), with ICC <0.40 indicating poor agreement, 0.40–0.74 fair to good agreement, and ≥0.75 excellent agreement (8). We examined the absolute difference between the test and retest administration of the questionnaires via intraindividual differences and Wilcoxon signed-rank test. As the test–retest interval between the two administrations of the questionnaire (7-d reference frame) varied between participants (between 2 and 8 wk), true behavior change could have affected the reliability results. We therefore conducted a sensitivity analysis in a subgroup of the larger English sample, including only those who reported a similar test and retest reference week (i.e., who provided the same Likert scale rating for the introductory question of the questionnaire at the test and retest administration).

Criterion validity of the SIT-Q-7d (test administration) compared with the respective criteria was examined using Spearman rank-order correlations (relative validity). We assessed the absolute agreement between the questionnaire and criterion-derived variables using Wilcoxon signed-rank test and by Bland–Altman plots (1). All analyses were conducted using IBM SPSS Statistics 19 (SPSS, Inc., Chicago, IL) and STATA 12.0 (StataCorp LP, College Station, TX). Statistical significance was set at $P < 0.05$.

RESULTS

Test–Retest Reliability

Flemish sample. Fifty-three participants (38% men) with a test and retest interval of 2–8 wk were included in the

reliability analysis (median interval, 3.3 wk; exactly 2 wk in 47% of participants). They were 38.4 ± 11.3 yr old (men, 39.8 ± 10.0; women, 37.6 ± 12.1) and had an average body mass index (BMI) of 24.2 ± 3.4 $kg \cdot m^{-2}$ (men, 25.3 ± 3.4 $kg \cdot m^{-2}$; women, 23.5 ± 3.2 $kg \cdot m^{-2}$; 32% overweight, 6% obese). Twenty-three percent had a secondary school or lower education level, and 77% participants had a university or college degree. As shown in Table 1, average day total sedentary time showed good test–retest reliability (ICC, 0.68). Almost all summary indicators of domain-specific sedentary time (transportation, occupation, screen time, and other sedentary time) showed fair to good reliability (ICC = 0.50–0.66), except for meal sedentary time, which showed poor reliability (ICC = 0.36). Reliability for most specific sedentary behaviors (TV viewing, nonoccupational computer use, reading, caring, and hobbies) was fair to excellent (ICC = 0.48–0.72), except for sitting during household tasks and while socializing and listening to music, which were poor (ICC = 0.26–0.38). Poor reliability was also found for the number of breaks in occupational sitting and in sitting while watching TV (ICC = 0.26–0.31). ICC was similar between weekday and weekend day variables and comparable with those for an average day. Intraindividual differences between the test and retest administration were small and only statistically significant for screen time, which was mainly driven by TV viewing during the week, and breaks in occupational sitting, as indicated by the Wilcoxon signed-rank test (Table 1).

English sample. The reliability subsample consisted of 281 participants (46% men) with a test–retest interval of 2–8 wk (median interval, 3.4 wk). They were 50.3 ± 7.4 yr old (men, 50.0 ± 7.4 yr of age; women, 50.5 ± 7.4 yr of age). Average BMI was 25.6 ± 4.1 $kg \cdot m^{-2}$ (men, 26.2 ± 3.5 $kg \cdot m^{-2}$;

TABLE 1. Descriptive characteristics (median (interquartile range)), ICC (95% CI) examining test–retest reliability, and intraindividual differences (median (interquartile range)) and Wilcoxon signed-rank test for sedentary behavior summary variables from the Sit-Q-7d between the test and retest measurement, Dutch version.

Summary Variable ($h \cdot d^{-1}$, Unless Stated Otherwise)	Test	Retest	ICC (95% CI)	Intraindividual Difference
Average day				
Total sedentary time (n = 48)	10.70 (3.96)	9.38 (4.16)	0.68 (0.50–0.81)***	0.84 (4.10)
Meals (n = 52)	0.97 (0.40)	1.08 (0.48)	0.36 (0.10–0.57)**	−0.07 (0.50)
Transportation (n = 53)	0.73 (1.14)	0.86 (1.02)	0.58 (0.37–0.74)***	−0.11 (0.70)
Occupation (n = 46)[a]	3.93 (2.77)	3.21 (3.12)	0.66 (0.46–0.79)***	0.08 (1.53)
Screen time (n = 50)	2.77 (1.51)	2.25 (1.74)	0.50 (0.26–0.68)***	0.55 (1.35)*
TV viewing (n = 48)	1.38 (1.57)	0.96 (1.23)	0.54 (0.31–0.71)***	0.00 (0.98)
Nonoccupational computer use (n = 49)	0.96 (1.55)	0.96 (1.37)	0.51 (0.27–0.69)***	0.18 (1.07)
Other sedentary time (n = 51)	2.05 (3.09)	2.25 (1.91)	0.52 (0.29–0.70)***	−0.07 (2.23)
Sitting while reading (n = 50)	0.38 (0.63)	0.34 (0.29)	0.48 (0.24–0.67)***	0.00 (0.41)
Sitting while doing household tasks (n = 48)	0.00 (0.11)	0.00 (0.28)	0.34 (0.07–0.57)**	0.00 (0.00)
Sitting while caring (n = 52)	0.00 (0.09)	0.00 (0.21)	0.72 (0.56–0.83)***	0.00 (0.00)
Sitting for hobbies (n = 52)	0.00 (0.33)	0.00 (0.38)	0.51 (0.27–0.68)***	0.00 (0.25)
Sitting for socializing (n = 51)	0.80 (1.29)	0.71 (0.61)	0.38 (0.12–0.59)**	0.00 (0.91)
Sitting while listening to music (n = 52)	0.00 (0.09)	0.00 (0.18)	0.26 (−0.01 to 0.50)*	0.00 (0.00)
Average breaks occupation (breaks per hour, n = 35)[b]	1.54 (2.87)	1.00 (2.29)	0.26 (−0.07 to 0.54)	0.57 (1.59)**
Average breaks TV viewing (breaks per hour, n = 39)[c]	2.00 (2.95)	2.67 (3.97)	0.31 (−0.01 to 0.57)*	−0.13 (2.99)
Sleeping (n = 50)	8.23 (0.87)	8.13 (1.16)	0.40 (0.14–0.61)**	0.07 (0.64)
Weekday				
Meals (n = 52)	0.92 (0.45)	1.06 (0.54)	0.40 (0.15–0.61)*	0.00 (0.42)
Screen time (n = 52)	2.50 (1.66)	2.25 (2.06)	0.52 (0.29–0.69)***	0.69 (1.09)*
TV viewing (n = 51)	1.50 (2.13)	0.75 (1.13)	0.52 (0.28–0.69)***	0.00 (0.75)*
Nonoccupational computer use (n = 52)	0.75 (1.34)	0.75 (1.13)	0.47 (0.23–0.66)***	0.00 (0.75)
Other sedentary time (n = 52)	1.31 (3.16)	1.38 (2.00)	0.51 (0.28–0.69)***	0.00 (1.69)
Sitting while reading (n = 50)	0.38 (0.63)	0.38 (0.25)	0.51 (0.28–0.69)***	0.00 (0.41)
Sitting while doing household tasks (n = 48)	0.00 (0.00)	0.00 (0.31)	0.36 (0.09–0.58)**	0.00 (0.00)
Sitting while caring (n = 51)	0.00 (0.00)	0.00 (0.13)	0.77 (0.62–0.86)***	0.00 (0.00)
Sitting for hobbies (n = 52)	0.00 (0.31)	0.00 (0.38)	0.56 (0.34–0.72)***	0.00 (0.25)
Sitting for socializing (n = 50)	0.38 (1.50)	0.38 (0.75)	0.29 (0.01–0.52)*	0.00 (0.84)
Sitting while listening to music (n = 52)	0.00 (0.00)	0.00 (0.13)	0.08 (−0.19 to 0.35)	0.00 (0.00)
Sleeping (n = 50)	7.90 (1.04)	8.00 (1.31)	0.33 (0.07–0.56)**	0.00 (0.75)
Weekend day				
Meals (n = 52)	1.25 (0.54)	1.29 (0.72)	0.28 (0.01–0.51)*	−0.10 (0.72)
Screen time (n = 50)	3.00 (1.84)	2.94 (2.09)	0.40 (0.14–0.61)**	0.31 (1.53)*
TV viewing (n = 48)	1.50 (2.13)	1.50 (1.75)	0.56 (0.33–0.73)***	0.00 (1.13)
Nonoccupational computer use (n = 49)	1.50 (1.31)	0.75 (1.81)	0.56 (0.34–0.73)***	0.00 (0.94)
Other sedentary time (n = 51)	3.50 (3.50)	3.25 (3.00)	0.43 (0.18–0.63)***	0.00 (3.50)
Sitting while reading (n = 47)	0.38 (0.63)	0.38 (0.63)	0.34 (0.06–0.57)**	0.00 (0.75)
Sitting while doing household tasks (n = 46)	0.00 (0.19)	0.00 (0.19)	0.37 (0.10–0.60)**	0.00 (0.00)
Sitting while caring (n = 51)	0.00 (0.13)	0.00 (0.38)	0.63 (0.43–0.77)***	0.00 (0.00)
Sitting for hobbies (n = 51)	0.00 (0.38)	0.00 (0.38)	0.36 (0.09–0.57)**	0.00 (0.38)
Sitting for socializing (n = 50)	1.50 (2.19)	1.50 (2.13)	0.43 (0.17–0.63)***	0.00 (1.75)
Sitting while listening to music (n = 51)	0.00 (0.00)	0.00 (0.38)	0.46 (0.22–0.65)***	0.00 (0.00)
Sleeping (n = 49)	8.75 (1.44)	8.38 (1.13)	0.40 (0.14–0.61)**	0.25 (1.25)

*P < 0.05.
**P < 0.01.
***P < 0.001.
[a] Three participants did not have an occupation.
[b] Three participants did not have an occupation and two reported no sedentary time at work.
[c] Five participants reported no sedentary time during TV viewing.

women, 25.1 ± 4.6 kg·m^{-2}; 40% overweight, 11% obese). Thirty-eight percent finished full time education at ≥ 20 yr of age, 27% between 16 and 20 yr of age, and 35% ≤ 16 yr of age. Test–retest reliability results are shown in Table 2. Average total sedentary time had fair reliability (ICC = 0.53). All domain-specific sedentary time indicators showed fair to excellent reliability (ICC = 0.45–0.76). Specific sedentary behavior indicators showed fair to good reliability (ICC = 0.48–0.63), except for sitting during household tasks, hobbies, and socializing (ICC = 0.06–0.39). Similar to the Dutch version, the number of breaks in occupational sitting or during TV viewing showed poor reliability (ICC = 0.12–0.28). No significant intraindividual differences between both administrations were found, except for total sedentary time (median difference, 26 min·d^{-1}), breaks in occupational sitting (0.79 breaks per hour), breaks during TV viewing (0.21 breaks per hour), weekday sitting while listening to music (0 min·d^{-1}), weekend TV viewing (0 min·d^{-1}), and weekend sitting for hobbies (0 min·d^{-1}), which were higher in the first administration (Table 2).

Comparing both sexes, in general, ICC was higher among men compared with women [see Table, Supplemental Digital Content 1, http://links.lww.com/MSS/A321, descriptive characteristics (median (interquartile range)), ICC (95% CI) examining test–retest reliability, and intraindividual differences (median (interquartile range)) and Wilcoxon signed-rank test for sedentary behavior summary variables from the Sit-Q-7d between the test and retest measurement, English version, men only; and Table, Supplemental Digital Content 2, http://links.lww.com/MSS/A322, descriptive characteristics (median (interquartile range)), ICC (95% CI) examining test–retest reliability, and intraindividual differences (median

TABLE 2. Descriptive characteristics (median (interquartile range)), ICC (95% CI) examining test–retest reliability, and intraindividual differences (median (interquartile range)) and Wilcoxon signed-rank test for sedentary behavior summary variables from the Sit-Q-7d between the test and retest measurement, English version.

Summary Variable (h·d^{-1}, Unless Stated Otherwise)	Test	Retest	ICC (95% CI)	Intraindividual Difference
Average day				
Total sedentary time (n = 237)	9.29 (4.21)	8.87 (4.34)	0.53 (0.44–0.62)***	0.44 (3.23)**
Meals (n = 271)	0.86 (0.59)	0.87 (0.62)	0.76 (0.71–0.81)***	0.00 (0.32)
Transportation (n = 273)	0.79 (1.32)	0.71 (1.13)	0.50 (0.40–0.58)***	0.00 (0.91)
Occupation (n = 197)^a	2.50 (4.11)	2.50 (3.93)	0.74 (0.67–0.80)***	0.00 (1.12)
Screen time (n = 278)	2.79 (2.10)	2.79 (2.13)	0.61 (0.53–0.67)***	0.00 (1.37)
TV viewing (n = 269)	1.79 (1.54)	1.79 (1.50)	0.69 (0.62–0.75)***	0.00 (0.88)
Nonoccupational computer use (n = 276)	0.75 (1.26)	0.75 (1.23)	0.57 (0.48–0.64)***	0.00 (0.56)
Other sedentary time (n = 275)	1.71 (2.05)	1.75 (1.80)	0.45 (0.35–0.54)***	0.00 (1.61)
Sitting while reading (n = 278)	0.40 (0.55)	0.38 (0.55)	0.59 (0.51–0.66)***	0.00 (0.39)
Sitting while doing household tasks (n = 272)	0.00 (0.00)	0.00 (0.00)	0.06 (−0.06 to 0.17)	0.00 (0.00)
Sitting while caring (n = 274)	0.00 (0.00)	0.00 (0.05)	0.63 (0.55–0.70)***	0.00 (0.00)
Sitting for hobbies (n = 269)	0.00 (0.38)	0.00 (0.27)	0.28 (0.16–0.38)***	0.00 (0.05)
Sitting for socializing (n = 276)	0.53 (1.00)	0.53 (1.13)	0.39 (0.29–0.49)***	0.00 (0.80)
Sitting while listening to music (n = 274)	0.00 (0.21)	0.00 (0.13)	0.48 (0.38–0.57)***	0.00 (0.00)
Average breaks occupation (breaks per hour; n = 144)^b	2.31 (3.51)	1.33 (2.25)	0.12 (−0.04 to 0.28)	0.79 (1.44)***
Average breaks TV viewing (breaks per hour; n = 227)^c	2.00 (1.54)	2.24 (2.51)	0.28 (0.15–0.39)***	−0.21 (1.72)*
Sleeping (n = 275)	8.04 (1.11)	8.11 (1.11)	0.70 (0.64–0.76)***	0.00 (0.64)
Weekday				
Meals (n = 271)	0.75 (0.54)	0.83 (0.54)	0.74 (0.68–0.79)***	0.00 (0.33)
Screen time (n = 280)	2.63 (2.38)	2.63 (2.38)	0.53 (0.44–0.61)***	0.00 (1.50)
TV viewing (n = 277)	1.50 (1.75)	1.50 (1.00)	0.63 (0.55–0.69)***	0.00 (1.19)
Nonoccupational computer use (n = 276)	0.75 (1.38)	0.75 (1.38)	0.46 (0.37–0.55)***	0.00 (0.72)
Other sedentary time (n = 276)	1.50 (2.00)	1.50 (1.72)	0.45 (0.35–0.54)***	0.00 (1.34)
Sitting while reading (n = 276)	0.38 (0.63)	0.38 (0.63)	0.50 (0.41–0.59)***	0.00 (0.50)
Sitting while doing household tasks (n = 271)	0.00 (0.00)	0.00 (0.00)	0.06 (−0.06 to 0.17)	0.00 (0.00)
Sitting while caring (n = 272)	0.00 (0.00)	0.00 (0.00)	0.58 (0.49–0.65)***	0.00 (0.00)
Sitting for hobbies (n = 267)	0.00 (0.38)	0.00 (0.38)	0.27 (0.15–0.37)***	0.00 (0.00)
Sitting for socializing (n = 271)	0.38 (0.75)	0.38 (0.75)	0.42 (0.31–0.51)***	0.00 (0.75)
Sitting while listening to music (n = 272)	0.00 (0.13)	0.00 (0.13)	0.40 (0.30–0.50)***	0.00 (0.00)*
Sleeping (n = 275)	8.00 (1.25)	8.00 (1.13)	0.70 (0.64–0.76)***	0.00 (0.63)
Weekend day				
Meals (n = 274)	0.96 (0.71)	0.92 (0.71)	0.74 (0.68–0.79)***	0.00 (0.38)
Screen time (n = 278)	3.00 (2.53)	3.13 (2.38)	0.65 (0.57–0.71)***	0.00 (2.00)
TV viewing (n = 270)	2.50 (2.00)	2.50 (2.00)	0.64 (0.57–0.71)***	0.00 (1.75)*
Nonoccupational computer use (n = 276)	0.75 (1.38)	0.75 (1.38)	0.72 (0.65–0.77)***	0.00 (0.63)
Other sedentary time (n = 275)	2.38 (2.88)	2.25 (3.00)	0.41 (0.31–0.51)***	0.00 (2.38)
Sitting while reading (n = 270)	0.75 (0.81)	0.38 (0.63)	0.67 (0.59–0.73)***	0.00 (0.63)
Sitting while doing household tasks (n = 271)	0.00 (0.00)	0.00 (0.00)	0.03 (−0.09 to 0.15)	0.00 (0.00)
Sitting while caring (n = 272)	0.00 (0.00)	0.00 (0.00)	0.57 (0.49–0.65)***	0.00 (0.00)
Sitting for hobbies (n = 263)	0.00 (0.38)	0.00 (0.38)	0.26 (0.15–0.37)***	0.00 (0.00)*
Sitting for socializing (n = 271)	0.75 (1.50)	0.75 (2.50)	0.33 (0.22–0.43)***	0.00 (1.50)
Sitting while listening to music (n = 272)	0.00 (0.13)	0.00 (0.13)	0.51 (0.41–0.59)***	0.00 (0.00)
Sleeping (n = 275)	8.50 (1.25)	8.50 (1.00)	0.61 (0.53–0.68)***	0.00 (1.00)

*$P < 0.05$.
**$P < 0.01$.
***$P < 0.001$.
^{a}Thirty-three participants did not have an occupation.
^{b}Thirty-three participants did not have an occupation and 22 reported no sedentary time at work.
^{c}Ten participants reported no sedentary time during TV viewing.

(interquartile range)) and Wilcoxon signed-rank test for sedentary behavior summary variables from the Sit-Q-7d between the test and retest measurement, English version, women only]. Men showed good to excellent reliability for average total and domain-specific sitting indicators (ICC = 0.61–0.81), whereas women showed fair to good reliability for total sedentary time, sitting during meals, occupation, and screen time (ICC = 0.42–0.72), but poor reliability for sitting during transportation, and other activities (ICC = 0.31–0.36). A similar pattern of higher reliability among men was found for the week- and weekend-specific variables. No substantial difference in ICC was found between weekdays and weekend days in both sexes.

A sensitivity analysis only including participants who reported similar reference weeks for both administrations (n = 97) is shown in Supplemental Digital Content 3 [see Table, Supplemental Digital Content 3, http://links.lww.com/MSS/A323, descriptive characteristics (median (interquartile range)), ICC (95% CI) examining test–retest reliability, and intraindividual differences (median (interquartile range)) and Wilcoxon signed-rank test for sedentary behavior summary variables from the Sit-Q-7d between the test and retest measurement, English version, participants with a comparable test and retest week only]. ICC was relatively comparable with those from the whole sample (Table 2). As expected, neither total sedentary time nor most other sitting indices differed significantly between both administrations, except for breaks during occupational sitting (median difference: 0.83 breaks per hour) and sitting for hobbies during the week (0 min·d^{-1}).

Criterion Validity

Flemish sample. Fifty-one Flemish participants (41% men) were included in the validity analyses, after excluding participants because of technical problems with the monitor (n = 3) or those who did not provide sufficient (i.e., ≥5 d, including ≥1 weekend day) valid criterion data because of insufficient monitor wear time and/or insufficient log time (n = 8). Of those included, 84% had six full days of monitoring time. The age, education, and BMI of the included sample showed very similar distributions compared with the reliability subsample. Table 3 and Figure 1 present the results for relative (Spearman rho correlation coefficients) and absolute validity (Wilcoxon signed-rank tests and Bland–Altman plots). The average total sitting time (rho = 0.52) and the domain-specific sitting indicators were all significantly correlated with their criterion (rho = 0.36–0.84), with the exception of sitting during meals (rho = 0.21) and breaks in occupational sitting and while watching TV (rho = 0.06–0.09). Correlations were relatively similar between weekdays and weekend days, except for meal sitting (nonsignificant for weekend day) and other sedentary time (nonsignificant for weekday).

TABLE 3. Descriptive characteristics (median (interquartile range)), Spearman correlation coefficients (rho) examining criterion validity, and intraindividual differences (median (interquartile range)) and Wilcoxon signed-rank test examining absolute validity between the sedentary behavior summary variables from the Sit-Q-7d and the criterion measure (thigh accelerometry + domain log), Dutch version.

Summary Variable (h·d^{-1}, Unless Stated Otherwise)	Questionnaire	Criterion	Rho	Intraindividual Difference
Average day				
Total sedentary time (n = 49)	10.93 (4.35)	9.35 (2.69)	0.52***	0.98 (4.25)**
Meals (n = 50)	1.01 (0.46)	0.86 (0.48)	0.21	0.07 (0.66)*
Transportation (n = 51)	0.82 (1.39)	0.87 (0.84)	0.46***	0.10 (0.65))
Occupation (n = 48)[a]	4.02 (3.34)	3.00 (2.39)	0.63***	0.28 (1.74)*
Screen time (n = 51)	2.75 (1.54)	1.81 (1.56)	0.76***	0.77 (1.07)***
TV viewing (n = 51)	1.50 (1.57)	1.22 (1.71)	0.84***	0.06 (0.66)
Nonoccupational computer use (n = 51)	0.96 (1.43)	0.28 (1.07)	0.72***	0.58 (0.66)***
Other sedentary time (n = 50)	1.96 (2.65)	2.51 (1.45)	0.36*	−0.30 (2.01)
Average breaks occupation (breaks per hour; n = 40)[b]	2.00 (2.28)	3.78 (1.77)	0.06	−2.23 (2.80)**
Average breaks TV viewing (breaks per hour; n = 43)[c]	1.84 (2.53)	2.46 (1.49)	0.09	−0.37 (2.93)
Sleeping (n = 49)	8.24 (0.87)	7.86 (0.69)	0.46***	0.13 (0.74)*
Weekday				
Meals (n = 50)	0.92 (0.51)	0.83 (0.60)	0.34*	0.05 (0.61)
Screen time (n = 51)	2.50 (1.75)	1.49 (1.43)	0.60***	0.79 (1.27)***
TV viewing (n = 51)	1.50 (2.13)	1.09 (1.51)	0.75***	0.00 (0.89)
Nonoccupational computer use (n = 51)	0.75 (1.25)	0.24 (0.74)	0.63***	0.63 (0.82)***
Other sedentary time (n = 50)	1.31 (2.38)	1.78 (1.49)	0.19	0.04 (2.57)
Sleeping (n = 49)	8.00 (1.00)	7.53 (0.78)	0.58***	0.31 (0.82)***
Weekend day				
Meals (n = 50)	1.25 (0.75)	0.88 (0.64)	0.19	0.50 (0.78)***
Screen time (n = 51)	3.00 (2.13)	2.06 (2.27)	0.74***	0.94 (1.33)***
TV viewing (n = 51)	1.50 (2.13)	1.05 (2.24)	0.86***	0.27 (0.84)*
Nonoccupational computer use (n = 51)	1.50 (0.75)	0.00 (1.21)	0.57***	0.75 (1.28)***
Other sedentary time (n = 51)	3.25 (3.50)	3.55 (1.88)	0.35*	−0.48 (3.07)
Sleeping (n = 47)	9.00 (1.00)	8.89 (1.44)	0.35**	0.17 (1.33)

*P < 0.05.
**P < 0.01.
***P < 0.001.
[a] Three participants did not have an occupation.
[b] Three participants did not have an occupation and two reported no sedentary time at work.
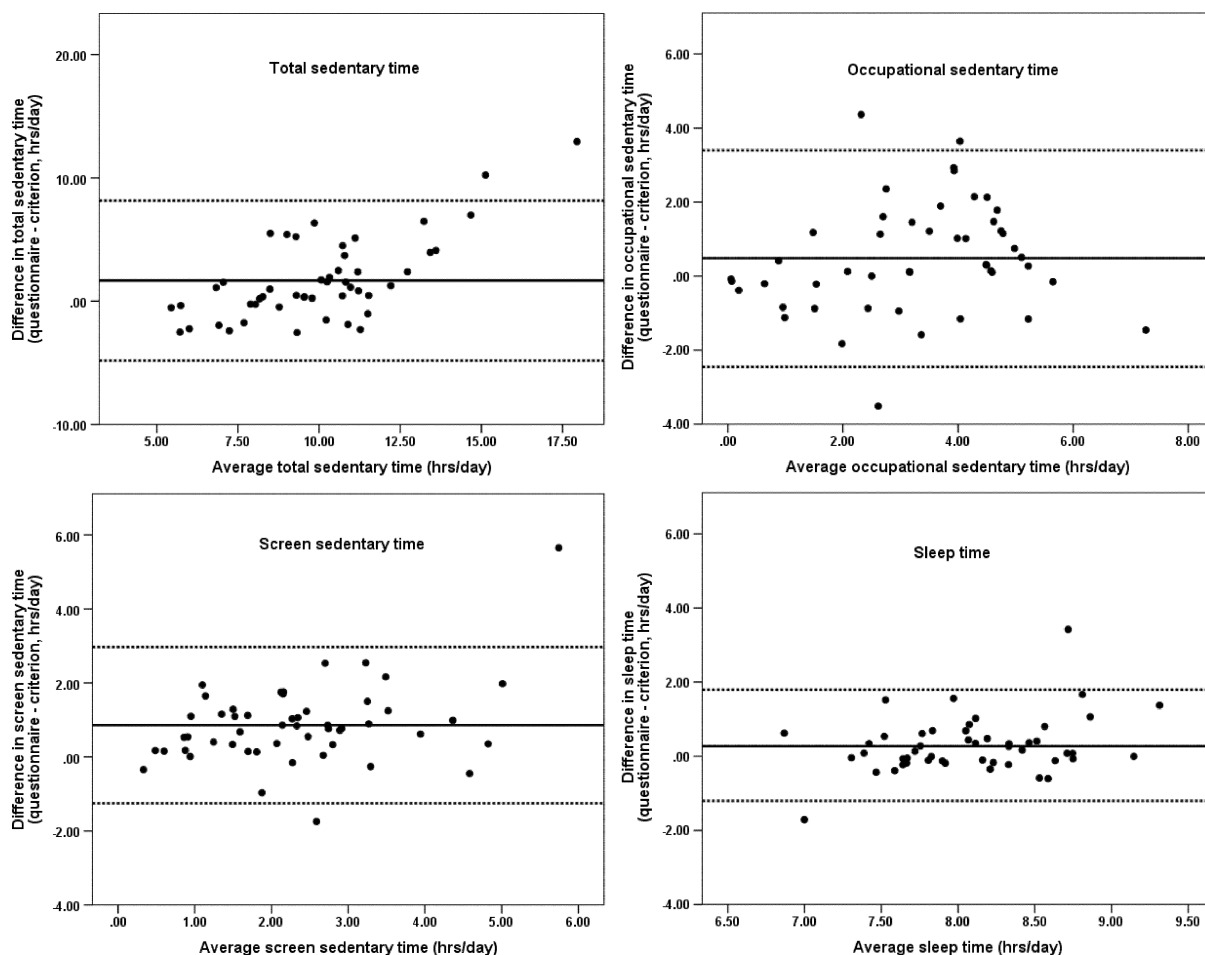[c] Five participants reported no sedentary time during TV viewing.

FIGURE 1—Bland–Altman plot of absolute agreement of total sedentary time, occupational sedentary time, screen sedentary time, and sleep time derived from the questionnaire, with the equivalent variable derived from the criterion measure (combination of the thigh accelerometry and domain log).

Participants significantly overreported their total sedentary time (median difference, 59 min·d^{-1}). This bias was predominantly due to significant overreporting of their nonoccupational computer (35 min·d^{-1}) and occupational (17 min·d^{-1}) sitting time, as well as sitting during meals (30 min·d^{-1}) and while watching TV (16 min·d^{-1}) in the weekend. They significantly underreported their number of breaks in occupational sitting time. Bland–Altman plots (Fig. 1) showed relatively wide 95% limits of agreement for total sedentary time (−4.81, 8.17 h·d^{-1}), occupational sedentary time (−2.44, 3.40 h·d^{-1}), and screen sedentary time (−1.25, 2.97 h·d^{-1}). The difference between the questionnaire and the criterion was not significantly correlated with the average sedentary time for any of the variables, except for total sedentary time (rho = 0.55, $P < 0.001$); participants with lower sedentary time tended to underreport their sitting time, whereas the reverse was seen for those with high levels of sitting time.

English sample. A total of 402 participants (48% men) were included in the validity analysis, after excluding those who had technical problems with the monitor ($n = 6$), insufficient valid monitor wear time ($n = 85$), self-report sleep data

deemed implausible by visual inspection ($n = 38$), and missing self-reported sedentary time data ($n = 38$). Forty-two percent of the 402 included participants had six full days of monitoring time. The validity subsample showed similar distributions for age, BMI, and education level as the reliability subsample. Relative and absolute validity for self-reported total sedentary behavior (test questionnaire) compared with total sedentary time estimated from combined sensing is shown in Table 4 and Figure 2, for the total sample and both sexes separately. In the total sample, significant correlations were found with both the EE-based (rho = 0.22) and ACC-based (rho = 0.37) criteria (Pearson correlation between both objective criterion measures: $r = 0.66$, $P < 0.001$). The correlation with the ACC-based criterion was the strongest, both in the total and the sex-specific analysis. Relative validity among women was somewhat stronger than among men, especially for the EE-based criterion.

On group level, participants overreported total sedentary time compared with the EE-based criterion (median difference, 25 min·d^{-1}) and underreported it compared with the ACC-based criterion (44 min·d^{-1}). Bland–Altman plots showed that the 95% limits of agreement were wide, both

TABLE 4. Descriptive characteristics (median (interquartile range)), Spearman correlation coefficients (rho) examining criterion validity, and intraindividual differences (median (interquartile range)) and Wilcoxon signed-rank test examining absolute validity between average day total sedentary time from the Sit-Q-7d and total sedentary time derived from combined HR and trunk acceleration sensing, English version.

Summary Variable (h·d^{-1})	Questionnaire	Criterion	Rho	Intraindividual Difference
Total sample (n = 402)				
Energy expenditure-based criterion	9.20 (4.37)	8.78 (3.16)	0.22***	0.41 (4.60)***
Accelerometry-based criterion	9.20 (4.37)	10.04 (2.29)	0.37***	−0.73 (3.71)***
Men (n = 193)				
Energy expenditure-based criterion	9.94 (4.43)	8.92 (3.11)	0.16*	1.00 (4.82)***
Accelerometry-based criterion	9.94 (4.43)	10.54 (1.97)	0.32***	−0.59 (3.66)*
Women (n = 209)				
Energy expenditure-based criterion	8.98 (4.29)	8.73 (3.28)	0.26***	0.17 (4.34)
Accelerometry-based criterion	8.98 (4.29)	9.62 (1.76)	0.35***	−0.84 (3.77)***

*$P < 0.05$.
**$P < 0.01$.
***$P < 0.001$.

compared with the EE-based (−6.85, 8.57 h·d^{-1}) and ACC-based criterion (−7.24, 6.52 h·d^{-1}). Similar to the Dutch version, differences between questionnaire and criterion were significantly correlated with sedentary time derived from both measures, and this was the case for both criterion variables (EE-based, rho = 0.29; ACC-based, rho = 0.60; both $P < 0.001$). Again, this indicates underreporting in those with low sedentary time and increasingly greater overreporting in those with higher sedentary time. In sex-specific analysis, compared with the EE-based criterion, men significantly overreported their sedentary time (median difference, 60 min·d^{-1}), whereas women did not (10.2 min·d^{-1}, $P = 0.46$). Compared with the ACC-based criterion, both men (35 min·d^{-1}) and women (50 min·d^{-1}) significantly underreported their sedentary time. Although somewhat tighter among women, wide 95% limits of agreement were found for both sexes (EE-based: men, −6.95,

10.02 h·d^{-1}; women, −6.48, 6.95 h·d^{-1}; ACC-based: men, −7.82, 7.71 h·d^{-1}; women, −6.55, 5.28 h·d^{-1}). Significant correlations were found between the differences between questionnaires and criterion, and sedentary time calculated from both measures (EE-based: men, rho = 0.33; women, rho = 0.24; ACC-based: men, rho = 0.64; women, rho = 0.59; all $P < 0.001$).

DISCUSSION

This study provides evidence for the reliability and validity in the general adult population of the SIT-Q-7d, a self-report instrument designed to assess total and domain-specific sedentary behavior with a reference frame of the last 7 d. A unique feature of our study is the use of domain- and week/weekend-specific criterion variables based on domain log-annotated sitting/lying time derived from thigh accelerometry
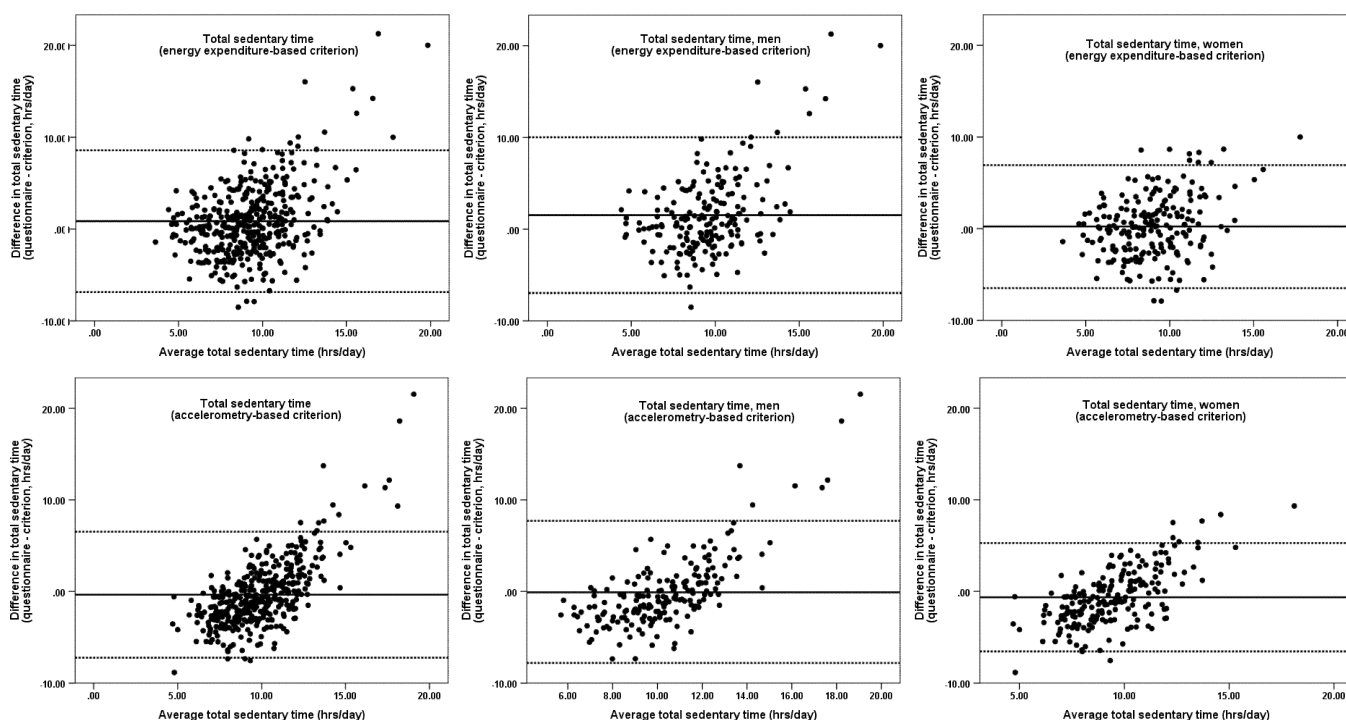

FIGURE 2—Bland–Altman plot of absolute agreement of total sedentary time derived from the questionnaire with the total sedentary time derived from the criterion measure (combined HR and trunk acceleration sensing).

(17,21,23,32). This combination provides an improved criterion measure compared with previous studies examining the validity of a sedentary behavior questionnaire (19).

The validity correlations for the total and domain-specific sitting indices with this combined criterion were relatively strong and at the high end of the spectrum of validity coefficients from previous sedentary behavior questionnaires (6,7,10–12,15,19,25,31). No other study has evaluated domain-specific measures of sitting against log-annotated thigh accelerometry. However, for total self-reported sitting, the SIT-Q-7d showed similar correlations with thigh accelerometry as the PAST questionnaire, which is an interviewer-administered, past-day recall of sitting time across domains (single administration, $r = 0.58$; two-day administration, $r = 0.53$) (11), and as the last-week International Physical Activity Questionnaire sitting question (weekdays, $r = 0.41$; weekends, $r = 0.55$), and it showed a better correlation than another composite sitting time measure (22). Lower relative validity of SIT-Q-7d total sedentary time was found when using EE-based and ACC-based sedentary time measures derived from the combined sensing method as criterion in our larger English sample. Although this method has high validity for PAEE of specific sedentary behaviors (13,38), it has only limited capacity to distinguish between postures, and both the EE-based and ACC-based sedentary measures depend on cut points, which are yet to be evaluated across a wider range of sedentary behaviors. The comparison between the SIT-Q-7d total sedentary time and the sedentary measures from combined sensing therefore needs to be interpreted in this context, acknowledging that discrepancies may be due to error in both methods.

Consistent with previous findings, estimates of sitting time during TV viewing, nonoccupational computer use, total screen time, and occupational sitting (i.e., behaviors which tend to happen in a more structured manner and therefore may be easier to recall accurately) were more valid (all >0.6) compared with estimates of sitting during transportation, meals, and other sedentary time (19,25). Scheers et al. (34) recently annotated sedentary time derived from the SenseWear Armband (≤1.8 MET) with electronic activity diary information to determine validity of domain-sitting indices derived from the Flemish Physical Activity Computerized Questionnaire (FPACQ). This questionnaire showed similar correlations for domain-specific sitting indices that were common between the FPACQ and SIT-Q-7d instruments (FPACQ: screen time, $r = 0.57$; motorized transport, $r = 0.58$; eating time, $r = 0.26$). Relatively high validity was found for the SIT-Q-7d for TV viewing, nonoccupational computer use, and total screen time for both weekdays and weekend days (all ≥0.57). Few studies have so far examined domain-specific sitting by weekdays and weekend days. Compared with a log, Marshall et al. (25) found relatively similar results for weekday TV viewing and computer use, but could not replicate validity for the weekend. Although weekend days generally tend to be less structured in terms of time use, potentially resulting in lower validity results (6,12), this argument may be less applicable to screen time compared with other behaviors such as meal time or other leisure sitting. Screen time may still mainly be an evening activity and was relatively similar in duration between the week and the weekend.

The SIT-Q-7d generally overestimated total and some domain-specific sedentary time indices compared with the log-annotated thigh accelerometry criterion. Previous studies using total thigh accelerometry-based sitting time as a criterion measure have shown both under- and overestimations for composite total sitting time (11,22). The SIT-Q-7d aimed to be comprehensive in terms of daily sedentary behaviors, also including domains (e.g., meals) that were not included in questionnaires showing underestimation of total sedentary time (15). Furthermore, sedentary behaviors have a much higher likelihood to cooccur (e.g., sitting while working and listening to music) compared with higher intensity activities. Double reporting of such composite behaviors may have caused overestimation, despite repeated questionnaire instructions on mutual exclusivity of reported time spent in different domains. Administration by interview instead of self-report could diminish this error and explain discrepancies with other measures such as the PAST questionnaire (11). In addition, brevity of wording in the domain-specific log compared with the SIT-Q-7d may have influenced comparison between the questionnaire and the criterion measure as well. Similar to the majority of other questionnaires (11,19,22), the 95% limits of agreement were wide for all measures examined, making the SIT-Q-7d less suitable for estimations of sedentary time at the individual level and for capturing change in sedentary time in intervention studies.

Test–retest reliability was fair to good for total and fair to excellent for domain-specific sedentary time in both samples, except for sitting during meals, which was poor in the Flemish but excellent in the English sample. The difference in reliability for sitting during meals between both samples may be explained by several factors, including differences in regularity of meal patterns (due to differences in sociocultural determinants), which would make associated behaviors easier to recall reliably, and differences in the cultural importance assigned to meal circumstances, which would differentially influence memory allocation and thus recall bias (14), although we have no firm evidence to support these hypotheses. Although variation in study populations, test–retest intervals, and reference time frames hamper direct comparison between studies, reliability coefficients derived in this study were broadly comparable with those found in previous studies (11,15,19,25). A consistent pattern in both samples and across the literature is that those indicators with poor reliability tended to be specific leisure sedentary behaviors, such as sitting for household tasks, hobbies, socializing, and listening to music, which are typically performed on a more sporadic and/or less prolonged basis (15,19,33). To examine whether behavior change between both administrations, due to the last 7-d reference frame and the relatively long test–retest interval in some participants, would have

negatively influenced the reliability results, a sensitivity analysis only including those who reported a similar test and retest week was performed. ICC was comparable, suggesting that the effect of behavior change on our results was minimal. In the English sample, reliability was generally somewhat stronger among men than women. Few studies have so far examined sex differences due to smaller study samples and results are mixed and difficult to interpret due to potential bias in the convenience samples used, including sex differences in education level and weight status (6,25). Women may have less structured lives due to higher prevalence of household and child care activities, which may compromise ability to report sedentary time across the day in a reliable manner. More women in our study did not work outside the home compared with men (15% vs 9%, respectively). Further large-scale studies should aim to examine this potential sex pattern.

Our study showed very poor psychometric properties, as well as substantial missing data, for the items determining the number of breaks in occupational sitting and while watching TV, indicating the difficulty of recalling this irregular behavior in a reliable and accurate manner. One previous study has determined the validity of a self-report measure of breaks during occupational sitting, showing modest validity (rho = 0.26) compared with breaks derived from hip-mounted accelerometry. Comparability between studies is hampered due to differences in administration method (self-report vs interviewer administered), criterion, and study population. Clark et al. (10) evaluated this item in full-time employees with predominantly sedentary jobs, who likely have less varied patterns of occupational sitting and breaks in occupational sitting. Further research in heterogeneous populations is needed to establish whether breaks in sedentary time could be assessed reliably and validly by questionnaire. In addition, studies implementing comprehensive sedentary behavior questionnaires such as the SIT-Q-7d should aim to take into account the different measurement properties of the individual sedentary behavior items when interpreting results, for example, relating to associations with health outcomes.

This study confirms reliability and relative validity of a questionnaire specifically designed to estimate total and domain-specific sedentary time, in the general adult population ages 20–60 yr. Unlike most previous validation studies, which used homogenous study populations, we were able to use population-based samples that were rather balanced in terms of sex, age, education level (which was less balanced in the Flemish sample compared with the English), and BMI, supporting the use of the SIT-Q-7d in general adult population studies in this age range. Furthermore, the questionnaire is self-administered rather than by interview, increasing its feasibility for implementation in large-scale observational studies. Self-report, noncomputerized questionnaires may however be associated with relatively more missing data, as was also the case in our study. Another unique strength of the study is the use of a combined thigh-accelerometer and domain log criterion, which allowed for detailed relative and absolute validity analysis by domain

and weekdays and weekend days with a strong criterion (17,21,23,32). The overlap between the questionnaire recall frame and criterion administration time frame in both substudies eliminates the possibility that lack of agreement is caused by fluctuations in behavior. Sensitivity analysis in those who provided two valid weekend days of criterion data (94% of the Flemish sample and 63% of the English sample) instead of only one showed very similar validity results (data not shown). In addition, both criterion monitors allowed for a 24-h wearing protocol, reducing misclassification of nonwear and sedentary time (40), which may be particularly relevant for sedentary behaviors in the evening (25). Finally, the large population-based sample of English adults enabled stratified reliability and validity analysis by sex, and sensitivity analysis in those with a comparable test and retest week.

The following limitations should also be considered. The validity of branched equation model estimates of the 1.5 MET threshold, or the 0.075 m·s^{-2} threshold for trunk acceleration, is yet to be fully evaluated. For example, Scheers et al. (34) used a threshold of 1.8 MET on the basis of estimates from another combined sensor. In addition, the keeping of a daily domain log, as was employed in the relatively small Flemish study sample, may have resulted in priming of memory, hence better recall for the test administration of the SIT-Q-7d. This effect is however likely to be minimal as the log predominantly aimed to time-stamp domains, irrespective of whether this time was spent sedentary or active, and the log design did not prompt participants to estimate durations. Other sources of error associated with the domain log may include its 15-min time resolution, potential delayed completion of the log (e.g., at the end of the day instead of contemporaneously), and potential nonsynchronization between participant's watch or other time-keeping device and the thigh accelerometer. The test–retest interval was not standard for all participants, which might have increased the difference in actual behavior between both reference weeks, thus potentially reducing reliability estimates. However, sensitivity analysis indicated this effect to be minimal. Finally, our study did not assess responsiveness to behavior change of the SIT-Q-7d, which would need to be established before using this questionnaire to track longitudinal shifts of sitting at a population level.

In conclusion, the SIT-Q-7d is a comprehensive questionnaire providing total and domain-specific sedentary behavior estimates with acceptable reliability and relative validity in the general adult population ages 20–60 yr. It is less reliable and valid for assessing breaks in sedentary time. These psychometric properties, together with the self-administration method, support the usefulness of this questionnaire for population surveillance purposes, as well as observational studies aiming to examine determinants of overall and domain-specific sitting, and associations with health outcomes. Future studies are warranted to examine psychometric properties of the SIT-Q-7d in other populations, such as the elderly, as well as the instrument's ability to detect change in sedentary behavior.

REFERENCES

1. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1(8476):307–10.

2. Bouchard C, Tremblay A, Leblanc C, Lortie G, Savard R, Theriault G. A method to assess energy expenditure in children and adults. *Am J Clin Nutr*. 1983;37(3):461–7.

3. Brage S, Brage N, Franks PW, Ekelund U, Wareham NJ. Reliability and validity of the combined heart rate and movement sensor Actiheart. *Eur J Clin Nutr*. 2005;59(4):561–70.

4. Brage S, Brage N, Franks PW, et al. Branched equation modeling of simultaneous accelerometry and heart rate monitoring improves estimate of directly measured physical activity energy expenditure. *J Appl Physiol*. 2004;96(1):343–51.

5. Brage S, Ekelund U, Brage N, et al. Hierarchy of individual calibration levels for heart rate and accelerometry to measure physical activity. *J Appl Physiol*. 2007;103(2):682–92.

6. Chau JY, van der Ploeg HP, Dunn S, Kurko J, Bauman AE. A tool for measuring workers' sitting time by domain: the Workforce Sitting Questionnaire. *Br J Sports Med*. 2011;45(15):1216–22.

7. Chau JY, van der Ploeg HP, Dunn S, Kurko J, Bauman AE. Validity of the occupational sitting and physical activity questionnaire. *Med Sci Sports Exerc*. 2012;44(1):118–25.

8. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychol Asses*. 1994;6(4):284–90.

9. Clark BK, Sugiyama T, Healy GN, Salmon J, Dunstan DW, Owen N. Validity and reliability of measures of television viewing time and other non-occupational sedentary behaviour of adults: a review. *Obes. Rev*. 2009;10(1):7–16.

10. Clark BK, Thorp AA, Winkler EA, et al. Validity of self-reported measures of workplace sitting time and breaks in sitting time. *Med Sci Sports Exerc*. 2011;43(10):1907–12.

11. Clark BK, Winkler E, Healy GN, et al. Adults' past-day recall of sedentary time: reliability, validity, and responsiveness. *Med Sci Sports Exerc*. 2013;45(6):1198–207.

12. Clemes SA, David BM, Zhao Y, Han X, Brown W. Validity of two self-report measures of sitting time. *J Phys Act Health*. 2012;9(4):533–9.

13. Crouter SE, Churilla JR, Bassett DR Jr. Accuracy of the Actiheart for the assessment of energy expenditure in adults. *Eur J Clin Nutr*. 2008;62(6):704–11.

14. De Castro JM. Socio-cultural determinants of meal size and frequency. *Br J Nutr*. 1997;77(1 Suppl):S39–54.

15. Gardiner PA, Clark BK, Healy GN, Eakin EG, Winkler EA, Owen N. Measuring older adults' sedentary time: reliability, validity, and responsiveness. *Med Sci Sports Exerc*. 2011;43(11):2127–33.

16. Godino JG, van Sluijs EM, Marteau TM, Sutton S, Sharp SJ, Griffin SJ. Effect of communicating genetic and phenotypic risk for type 2 diabetes in combination with lifestyle advice on objectively measured physical activity: protocol of a randomised controlled trial. *BMC Public Health*. 2012;12:444. doi: 10.1186/1471-2458-12-444.

17. Grant PM, Ryan CG, Tigbe WW, Granat MH. The validation of a novel activity monitor in the measurement of posture and motion during everyday activities. *Br J Sports Med*. 2006;40(12):992–7.

18. Harris JL, Bargh JA, Brownell KD. Priming effects of television food advertising on eating behavior. *Health Psychol*. 2009;28(4):404–13.

19. Healy GN, Clark BK, Winkler EA, Gardiner PA, Brown WJ, Matthews CE. Measurement of adults' sedentary time in population-based studies. *Am J Prev Med*. 2011;41(2):216–27.

20. Healy GN, Matthews CE, Dunstan DW, Winkler EA, Owen N. Sedentary time and cardio-metabolic biomarkers in US adults: NHANES 2003–06. *Eur Heart J*. 2011;32(5):590–7.

21. Kozey-Keadle S, Libertine A, Lyden K, Staudenmayer J, Freedson PS. Validation of wearable monitors for assessing sedentary behavior. *Med Sci Sports Exerc*. 2011;43(8):1561–7.

22. Kozey-Keadle S, Libertine A, Staudenmayer J, Freedson P. The feasibility of reducing and measuring sedentary time among overweight, non-exercising office workers. *J Obes*. 2012;2012:282303. doi: 10.1155/2012/282303.

23. Lyden K, Kozey Keadle SL, Staudenmayer JW, Freedson PS. Validity of two wearable monitors to estimate breaks from sedentary time. *Med Sci Sports Exerc*. 2012;44(11):2243–52.

24. Lynch BM, Friedenreich CM, Csizmadi I. Adult perception of sedentary behavior self-report assessment: Cognitive interviewing the SIT-Q. *J Sci Med Sport*. 2012;15(6):295.

25. Marshall AL, Miller YD, Burton NW, Brown WJ. Measuring total and domain-specific sitting: a study of reliability and validity. *Med Sci Sports Exerc*. 2010;42(6):1094–102.

26. Matthews CE, Chen KY, Freedson PS, et al. Amount of time spent in sedentary behaviors in the United States, 2003–2004. *Am J Epidemiol*. 2008;167(7):875–81.

27. Matthews CE, Moore SC, George SM, Sampson J, Bowles HR. Improving self-reports of active and sedentary behaviors in large epidemiologic studies. *Exerc Sport Sci Rev*. 2012;40(3):118–26.

28. Miller R, Brown W. Steps and sitting in a working population. *Int J Behav Med*. 2004;11(4):219–24.

29. Owen N, Sugiyama T, Eakin EE, Gardiner PA, Tremblay MS, Sallis JF. Adults' sedentary behavior determinants and interventions. *Am J Prev Med*. 2011;41(2):189–96.

30. Ridgway CL, Brage S, Sharp SJ, et al. Does birth weight influence physical activity in youth? A combined analysis of four studies using objectively measured physical activity. *PLoS One*. 2011;6(1):e16125. doi: 10.1371/journal.pone.0016125.

31. Rosenberg DE, Norman GJ, Wagner N, Patrick K, Calfas KJ, Sallis JF. Reliability and validity of the Sedentary Behavior Questionnaire (SBQ) for adults. *J Phys Act Health*. 2010;7(6):697–705.

32. Ryan CG, Grant PM, Tigbe WW, Granat MH. The validity and reliability of a novel activity monitor as a measure of walking. *Br J Sports Med*. 2006;40(9):779–84.

33. Salmon J, Owen N, Crawford D, Bauman A, Sallis JF. Physical activity and sedentary behavior: a population-based study of barriers, enjoyment, and preference. *Health Psychol*. 2003;22(2):178–88.

34. 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. doi: 10.1186/1479-5868-9-71.:71-9.

35. Sedentary Behaviour Research Network. Standardized use of the terms "sedentary" and "sedentary behaviours". *Appl Physiol Nutr Metab*. 2012;37:540–2.

36. Stamatakis E, Davis M, Stathi A, Hamer M. Associations between multiple indicators of objectively-measured and self-reported sedentary behaviour and cardiometabolic risk in older adults. *Prev Med*. 2012;54(1):82–7.

37. Stegle O, Fallert SV, MacKay DJ, Brage S. Gaussian process robust regression for noisy heart rate data. *IEEE Trans Biomed Eng*. 2008;55(9):2143–51.

38. Thompson D, Batterham AM, Bock S, Robson C, Stokes K. Assessment of low-to-moderate intensity physical activity thermogenesis in young adults using synchronized heart rate and accelerometry with branched-equation modeling. *J Nutr*. 2006;136(4):1037–42.

39. Thorp AA, Owen N, Neuhaus M, Dunstan DW. Sedentary behaviors and subsequent health outcomes in adults a systematic review of longitudinal studies, 1996–2011. *Am J Prev Med*. 2011;41(2):207–15.

40. Winkler EA, Gardiner PA, Clark BK, Matthews CE, Owen N, Healy GN. Identifying sedentary time using automated estimates of accelerometer wear time. *Br J Sports Med*. 2012;46(6):436–42.