MEASURING OLDER ADULTS' SEDENTARY TIME: RELIABILITY, VALIDITY AND RESPONSIVENESS

Paul A Gardiner[1], Bronwyn K Clark[1], Genevieve N Healy[1,2], Elizabeth G Eakin[1,2], Elisabeth AH Winkler[1] & Neville Owen[1,2]

[1] The University of Queensland, School of Population Health, Cancer Prevention Research Centre, Brisbane, Australia

[2] Baker IDI Heart and Diabetes Institute, Melbourne, Australia

Corresponding Author

Paul Gardiner

The University of Queensland

Cancer Prevention Research Centre, School of Population Health

Level 3 Public Health Building

Herston Rd, Herston, Qld 4006 Australia

Telephone: +61 7 3365 5163

Fax: +61 7 3365 5540

Email: p.gardiner@uq.edu.au

Funding:

This study was funded by a National Health and Medical Research Council Program Grant (#569663) and Supplementary Research Funding from the School of Population Health at the University of Queensland.

Inclusions:

Abstract (word count: 275)

Word count (excluding title page, abstract, references, figures and tables): 4,160

Number of tables for text: 2

Number of figures: 1

Running title: Measuring older adults' sedentary time

ABSTRACT

Purpose: With evidence that prolonged sitting has deleterious health consequences, decreasing sedentary time is a potentially important preventive health target. High-quality measures, particularly for use with older adults, who are the most sedentary population group, are needed to evaluate the impact of sedentary behavior interventions. We examined the reliability, validity and responsiveness to change of a self-report sedentary behavior questionnaire that assessed time spent in behaviors common among older adults: watching television (TV), computer use, reading, socialising, transport and hobbies, and a summary measure (total sedentary time).

Methods: In the context of a sedentary behavior intervention, non-working older adults (n = 48; mean age 73 [SD=8] years) completed the questionnaire on three occasions over a two week period (seven days between administrations), and wore an accelerometer (Actigraph model GT1M) for two periods of six days. Test-retest reliability (for the individual items and the summary measure) and validity (self-reported total sedentary time compared to accelerometer-derived sedentary time) were assessed over the one-week pre-intervention period, using Spearman's (ρ) correlations [95% CI]. Responsiveness to change following the intervention was assessed using the responsiveness statistic (RS).

Results: Test-retest reliability was excellent for TV viewing time (ρ=0.78 [0.63-0.89]), computer use (ρ=0.90 [0.83-0.94]) and reading (ρ=0.77 [0.62-0.86]), acceptable for hobbies (ρ=0.61 [0.39-0.76]) and poor for socialising and transport (ρ<0.45). Total sedentary time had acceptable test-retest reliability (ρ =0.52 [0.27 to 0.70]) and validity (ρ=0.30 [0.02-0.54]). Self-report total sedentary time was similarly responsive to change (RS=0.47) as accelerometer-derived sedentary time (RS=0.39).

Conclusions: The summary measure of total sedentary time has good repeatability, modest validity and is sufficiently responsive to change suggesting that it is suitable for use in interventions with older adults.

Keywords: sitting; responsiveness to change; accelerometer; questionnaire; sedentary behavior; intervention

INTRODUCTION

Paragraph 1 Time spent sedentary (too much sitting as distinct from lack of exercise) such as daily sitting time, watching television, driving in cars, or sitting during leisure-time is independently associated with an increased risk of premature all-cause and cardiovascular disease mortality (10, 17, 28, 37). While it has been shown via accelerometer-derived assessment that older adults have the highest levels of sedentary time (22), these objective measures do not provide the contextual information necessary for the identification of intervention targets and public health messages on how to reduce sedentary time (26).

Paragraph 2 As the use of objective measures is not always feasible for use in epidemiological and health-behavior intervention studies, high-quality self-report measures are needed, yet there is limited evidence on the measurement properties of existing questionnaires, particularly with older adults. In a review of 60 articles reporting non-occupational sedentary behavior in adults, nine studies reported on reliability and three studies reported on validity (7); only one reported on the measurement properties of a measure administered with older adults (23). Two previous studies have reported good test-retest reliability with older adults using a measure of global sedentary time designed for use in the general adult population (9, 18). However, in these studies validity was assessed against weekly pedometer step counts (9) or accelerometer-assessed moderate- to vigorous-intensity physical activity and total accelerometer counts (18), and not directly against an objective measure of sedentary time.

Paragraph 3 Decreasing sedentary time is a potentially important preventive health target. To evaluate interventions, measures that are responsive to change are required. By comparing the responsiveness to change in sedentary time of several different measures, we can evaluate which instruments can detect significant changes with the smallest number of participants, i.e. provide the most power for a given sample size (34). In contrast with the evidence for responsiveness to change of several physical activity measurement tools (12, 29, 36), no studies to date have examined the responsiveness of either self-report or accelerometer measures of sedentary time. Thus, it is not clear which type of sedentary time measure would be more responsive: the self-report measures may be better able to capture the specific behaviors targeted by intervention; on the other hand, the accelerometer measures are likely to have lower background variability due to their good reliability (5).

Paragraph 4 We examined the measurement properties of a self-reported total sedentary time measure and specific sedentary time items in non-working older adult participants of a sedentary behavior intervention trial. The specific measurement properties examined were test-retest reliability, responsiveness to change, and validity compared to accelerometer-derived sedentary time. We also assessed the responsiveness to change of accelerometer-derived sedentary time.

METHODS

Participants

Paragraph 5 Non-working older adults were recruited for a study examining the feasibility of an intervention (Stand Up For Your Health) designed to reduce sedentary time (27). Sixty participants were recruited through community-based organisations for older adults and retirement villages within urban areas of Brisbane, Australia. Eligibility criteria included: age 60 years or older, self-reported TV viewing time of two or more hours/day, not in paid employment, able to ambulate without assistance from another person, and residence within 50 km of the research center. Participants were enrolled in the study between May and

December 2009 and were not paid for their involvement. The study was approved by The University of Queensland's Behavioural Social Sciences Ethics Review Committee.

Study design and protocols

Paragraph 6 This pre-post study involved multiple baseline assessments over a pre-intervention period, followed immediately by an intervention and a post-intervention assessment period. Self-reported sedentary time (described below) was assessed using an interviewer-administered questionnaire (face-to-face) during three home visits conducted approximately one week apart. During the first visit participants provided written informed consent, had their height and weight measured, and completed a questionnaire to assess demographic characteristics, and sedentary time (T1 assessment). Participants were instructed to wear a uniaxial accelerometer (Actigraph model GT1M; http://www.theactigraph.com/) fitted firmly around their waist over the right hipbone, during all waking hours (except when bathing or participating in water-based activities) and to complete a log to record wear times. During the second visit, the questionnaire was re-administered (T2 assessment), data from the accelerometer were downloaded, and the battery recharged. A short intervention session was conducted and participants were asked to adhere to the accelerometer wearing protocol for a further week. During visit three, the accelerometer and log were collected and the questionnaire was re-administered (T3 assessment). Data from T1 and T2 and the first accelerometer assessment (pre-intervention period) were used to assess test-retest reliability and validity. Data from T3 and the second accelerometer assessment (post-intervention) were used to assess responsiveness to change from T2 and the first accelerometer assessment.

Paragraph 7 The inclusion criteria for this measurement study were: at least four valid days of accelerometer data during each assessment period (valid day = at least 10 hours of wear established from both accelerometer and log data; data from days on which the researcher conducted the home visits were not included); and complete responses to the interviewer-administered questionnaire at the three assessment time points (T1, T2, and T3). All participants provided valid accelerometer data for at least half of the assessment period covered by self-report.

Measures

Demographic measures and anthropometry

Paragraph 8 Participants were asked to report their date of birth (from which age was calculated), highest level of education completed, the number of people in their household, and how much time they usually spent sleeping each night over the past week. Participants' height (via stadiometer: nearest 0.5cm) and weight (via electronic scales: nearest 0.1kg) were measured to derive body mass index (BMI; kg/m^2).

Self-report measure of sedentary time

Paragraph 9 Self-reported sedentary time, in the specific domains of leisure-time and transport, was assessed using a seven-item, one-week recall questionnaire adapted from a previous measure of leisure-time sitting developed for the general adult population (32). The original summary measure was shown to have good test-retest reliability (ICC (95%CI) = 0.79 (0.71-0.85); 145 participants with mean (SD) age of 50.8 (13.5) years) and modest validity assessed against a behavior log (Spearman's rho = 0.30; 130 participants with mean (SD) age of 38.8 (15) years), however some individual items showed poor reliability and validity (32). Our adapted questionnaire asked participants to report on activities they did over the last week while they were sitting or lying down (not including time spent in bed) and report the total time spent in each activity. The seven individual sedentary items were [a] television (TV) or video/DVD watching; [b] computer use; [c] reading; [d] socialising with friends or family; [e] time travelling in a motor vehicle or on public transport; [f] doing hobbies; [g] and any other activities they did while they were sitting or lying down (see Appendix, Supplemental Digital Content 1, Questionnaire to measure sedentary time in older adults).

Paragraph 10 We adapted Salmon's original measure (32) to meet the needs of our intervention trial and to increase their relevance to our target population, by: [i] ensuring that time in the individual behaviors was mutually exclusive as participants were instructed to only count time when it was their main activity; for example they would count time spent knitting while watching TV as either TV time or time doing hobbies, but not both; [ii] addressing problems of low test-retest reliability in the original measure by removing two activities (listening to music/CD/radio; and relaxing, thinking and resting); [iii] revising two items to reflect the activities of older adults (talking on the telephone was included as part of socialising, and going for a recreational drive was expanded to include all transport 2)]; [iv] administering the questionnaire in a face-to-face interview (which allowed for prompting of participants, e.g. to identify activities undertaken on specific days), in contrast to studies with the original measure which were completed without any contact with an interviewer; [v] reducing participant burden by only asking about total time spent in activities during the last week, instead of asking separately about weekdays and weekend days as per the original measure; and, [vi] ensuring that the individual behaviors were suitable targets for behavior change. As per the original measure responses for the items were reported as continuous time (hours and minutes per week). Weekly time in each activity was converted to hours per day. Total sedentary time was calculated as the sum of daily time in each activity and reported as hours/day.

Accelerometer-derived sedentary time

Paragraph 11 Data were collected in one minute epochs. The commonly used cut-point of <100 counts per minute (cpm) (8, 22) was used to derive sedentary time. Days on which the researcher conducted home visits and when the accelerometer was worn for <10 hours were excluded. Wear time was determined from a combination of data from the accelerometer and wearing logs completed by participants. Average sedentary time was calculated as [total sedentary time / number of valid days] and expressed as hours per day. When examining responsiveness to change, sedentary time was standardised to 16 hours of waking time (participants reported a median sleeping time of 8 hours/night) to account for differences in accelerometer wear time in each assessment period. Data were summarized using SAS 9.1 (SAS Institute Inc., Cary, NC, US) via a modified version of the program available from the National Cancer Institute [http://riskfactor.cancer.gov/tools/nhanes_pam/; (20)].

Statistical Analyses

Paragraph 12 Analyses were conducted using STATA Statistical Software Release 11.0 (StataCorp LP, College Station, TX, USA) and SPSS version 17.0 (SPSS, Inc. Chicago IL). All sedentary variables had non-normal distributions, with the exception of total self-report sedentary time. Statistical significance was set at $p<0.05$.

Reliability

Paragraph 13 Reliability was assessed using a one week (pre-intervention) test-retest protocol. Correlation between T1 and T2 measures of time in each sedentary behavior and total sedentary time was assessed with Spearman's rank correlation coefficient (ρ); with 95% confidence intervals (95% CI) calculated using Fisher's transformation. To allow comparison to the original measure (32), we also assessed test-retest using single-measures intraclass correlation coefficients (ICC) with 95% CI, with an absolute agreement definition, which were calculated using two-way mixed effects models (33).

Validity

Paragraph 14 The relative validity of self-report total sedentary time was assessed with Spearman's rank correlation coefficient (ρ) with 95% CI using data from T2 and accelerometer-derived sedentary time over the pre-intervention period. Bland-Altman (3, 4) plots were used to assess absolute agreement between the two measures. Regressing average self-report sedentary time/accelerometer-derived sedentary time on the differences between the two measures revealed that the mean difference increased significantly as average values increased. The variability, however, remained constant across average values. Therefore, the Bland-Altman plot presents the trend line for mean difference obtained from the regression and limits of agreement (± 2 standard deviations).

Responsiveness to change

Paragraph 15 Responsiveness to change was assessed using the responsiveness statistic (RS), (13) which quantifies the minimum clinically important difference, or if this is unknown, the difference observed in an intervention, in relation to variability over a stable period. We calculated RS as mean change (Δ) within participants over the intervention period (T2 to T3) divided by the square root of two times mean squared error (MSE), our measure of background variability over a stable period (pre-intervention, T1-T2). Repeated measures ANOVAs were used to determine MSE. Thus RS was calculated as [$\Delta/\sqrt{(2x\ MSE)}$] using the change and MSE for each individual behavior and total sedentary time (both self-report and accelerometer-derived), with the direction of the change removed to ease interpretation of the magnitude of the RS. We also examined the proportion of participants who made substantial changes (15 minutes/day for individual behaviors and 30 minutes/day for total sedentary time).

RESULTS

Paragraph 16 Forty-eight (80%) of the participants in the intervention study met the inclusion criteria and were included in the current investigation. The majority of participants in this measurement study were women (n=35, 72.9%); had completed post-secondary level (e.g. university) or professional level (e.g. teaching certificate) education (n=35, 72.9%); lived with others (n=27, 56.3%) and were community dwelling (n=39, 81.3%). The participants' mean (SD) age and BMI were 72.8 (8.1) years and 27.2 (4.8) kg/m^2.

Paragraph 17 The median (minimum, maximum) number of days from T1 to T2 and T2 to T3 was 7 (6, 11) and 7 (6, 14) respectively. The duration for each individual sedentary item [median (25th, 75th percentile)] and total self-reported sedentary time [mean (SD)] at each assessment (i.e. T1, T2, and T3) is shown in Table 1. TV viewing time comprised the largest component of total sedentary time (37.6%, 37.1%, and 36.8% at T1, T2, and T3 respectively). A large proportion of participants reported no time in computer use, hobbies, or "other" sedentary behavior.

Paragraph 18 The median (minimum, maximum) number of valid days the accelerometer was worn was 6 (4, 10) in both assessment periods. The mean (SD) accelerometer wear time was 14.6 (0.97) in the first assessment period and 14.4 (1.10) hours/day in the second assessment period. Total sedentary time median (25th, 75th percentile) was 10.2 (9.5, 10.9) and 9.8 (8.5, 10.5) hours/day during the pre- and post-intervention assessment periods respectively. Accelerometer-derived sedentary time significantly decreased from pre- to post-intervention ($p<0.001$), as assessed using Wilcoxon Signed-Rank test. Total self-reported sedentary time did not exceed accelerometer wear time for any participant.

INSERT TABLE 1 ABOUT HERE

Reliability

Paragraph 19 Table 1 presents the test-retest reliability of individual sedentary behavior items and total self-reported sedentary time. The Spearman's correlations were high for computer use, TV viewing time and reading ($\rho>0.75$), modest to acceptable for the other individual items (ρ=0.23-0.61), and acceptable for total sedentary time ($\rho=0.56$). ICCs were similar to the Spearman's correlations for total sedentary time and most individual items, but were considerably lower for 'other' sedentary time and hobbies.

Validity

Paragraph 20 The correlation between total self-report and accelerometer-derived sedentary time was statistically significant, but modest (ρ [95% CI] = 0.30 [0.02 to 0.54]). The Figure shows the Bland-Altman plot for total self-reported and accelerometer-derived sedentary time. Linear regression showed a significant negative association between the difference in the two measures (self-reported minus accelerometer-derived sedentary time) and the average of these two measures (B= -0.67, SE (0.21), $p=0.003$). Thus, the mean difference is estimated at -9.20 hours + 0.67 * average of the two measures. At mean levels of average self-reported/accelerometer-derived sedentary time (8.36 hours), the mean difference indicated self-reported sedentary time was 3.60 hrs/day lower than accelerometer-derived sedentary time with wide limits of agreement (mean difference ± 3.82 hours).

INSERT FIGURE ABOUT HERE

Responsiveness to change

Paragraph 21 Table 2 shows the variability in the pre-intervention period [√(2*MSE)], the change from pre- to post-intervention, and the responsiveness to change for the self-report and accelerometer-derived measures of sedentary time. The reduction detected was greater for self-report (0.85 hrs/day) than accelerometer-derived sedentary time (0.50 hrs/day). The variability for self-report was greater than for accelerometer-derived sedentary time, however, they were both similarly responsive to change (0.47 and 0.39, respectively) because the amount of change was assessed to be larger by self-report than by accelerometer. Of the specific sedentary behavior items, the greatest reductions were seen in TV viewing time and hobbies, and these also had better responsiveness (0.34 and 0.33, respectively) than the other items. Table 2 also shows the proportion of participants making substantial changes from pre- to post-intervention. A larger proportion of participants made a substantial reduction in sedentary time as assessed by self-report than derived by accelerometer, while more participants reported making a substantial reduction in TV viewing time than the other individual sedentary behaviors.

INSERT TABLE 2 ABOUT HERE

DISCUSSION

Paragraph 22 This study examined the measurement properties of a self-report sedentary time questionnaire adapted specifically for use in epidemiological and health behavior intervention studies with non-working older adults. Our questionnaire is unique among measures designed for older adults in that it assesses continuous time spent in specific sedentary behaviors as well as the total time spent in these behaviors. Other novel aspects of this measure are the ability to capture waking time spent lying down, not just sitting time, and also the mutually exclusive nature of the items, which may improve the face validity of the summed measure of total sedentary time. This study is also unique in that it is the first in older adults to report validity for self-report sedentary time compared with accelerometer-derived sedentary time. Importantly, we provided the novel evidence that accelerometer-derived and self-reported sedentary time measures are both responsive to change. Coupled with finding adequate, although not ideal, validity and reliability for our self-report measure of sedentary time, this suggests that both our self-report and accelerometer-derived measures of sedentary time are suitable for use in interventions with older adults.

Paragraph 23 The test-retest reliability of the individual sedentary time items was similar to that reported for measures previously used with a general adult population sample (32), and for similar sedentary measures in varied populations (7). The reliability of total self-report sedentary time (ICC = 0.52) was fair to good, but lower than that for the earlier study with the general adult population sample (ICC = 0.82; 32). This is broadly consistent with previous studies that have shown measures used with older adults have lower test-retest reliability than in general adult populations, including for screen time (23) and for physical activity (24). This may be reflective of differences in the amount of true variation in sedentary behaviors over one week, possibly due to the age and working status of the participants. Other possible explanation are differences in the amount of measurement error which may be related to factors specific to older adults such as deficits in concentration / cognition / memory (30), or the higher test-retest reliability in the study by Salmon and colleagues (29) could reflect the greater between-person variability in the general adult population, which increases the ICC. To explore this, we repeated analyses excluding participants who reported that their sedentary time had changed during the pre-intervention period. At T1 and T2 participants were asked whether their sitting was comparable to a typical week (5-point scale from much less than normal to a lot more than normal). Test-retest reliability was examined for participants (n=32) who responded that their sitting at T2 was at a similar level to that at T1. The reliability of all measures improved: self-report sedentary time had good reliability (ICC [95% CI] = 0.74 [0.44 to 0.88] and ρ [95% CI] = 0.71 [0.48 to 0.85]); the Spearman's correlations for individual items ranged from ρ=0.41 (socialising) to ρ=0.93 (computer use). The increase in test-retest reliability in these sub-analyses suggests there may be substantial true variability in sedentary behaviors in older adults. Thus the measure of total sedentary time has acceptable repeatability when considering both random and systematic error.

Paragraph 24 Validity of this self-report measure of total sedentary time was less than ideal. However, our findings were comparable to what has been seen for global and composite measures in adult populations (assessed against accelerometer-derived sedentary time) in terms of correlation (Spearman's ρ ranging from -0.01 to 0.61 6, 8, 11, 31), mean difference (ranging from -2.2 to 0.18 hours/day 14, 21), and limits of agreement (ranging from ±5.53 to ± 6.90 hours/day 14, 21). Participants tended to report less sedentary time than was recorded by the accelerometer, with the discrepancy decreasing at higher levels of sedentary time. Previous studies have shown that adults under-report time spent in individual sedentary behaviors such as watching TV (23, 25), which may exacerbate the differences between self-report and comparison measures when using a summary score of total sedentary time, as is the case for our study. Further, it is possible that other salient aspects of sedentary time were not captured by the items in this questionnaire such as time spent eating.

Paragraph 25 Alternatively, the discrepancy could relate to over-detection by the accelerometer, e.g. some standing still might have been classified as sedentary. The accelerometer is not considered a gold-standard criterion for sedentary time as it does not detect posture. Recent evidence suggests that devices that measure body position (such as the *activ*PAL™ which distinguishes between time spent sitting/lying, standing, or walking) may be a better comparison measure for validity studies and for assessing changes in sedentary time. One study reported a smaller discrepancy between self-report sedentary time and sedentary time as measured by the *activ*PAL™ than between self-report sedentary time and sedentary time derived from the Actigraph GT1M accelerometer (15). Another study reported that the *activ*PAL™ detected greater reductions in sedentary time than the Actigraph GT3X accelerometer (19). Furthermore, the ideal accelerometer cut-point for sedentary time for older adults is not known. The <100cpm cut-point has been shown to be detrimentally associated with biomarkers of cardio-metabolic health and inflammation in population-based studies (16), and was thus used as the primary comparison value. Conclusions regarding validity were only minimally affected by the choice of cut-point. In a sensitivity analysis where sedentary time was classed as <50cpm, results were similar to those reported in the paper in terms of correlations, limits of agreement, and responsiveness. Mean differences were smaller than when using the commonly cited <100cpm cut-point. It may be that cut-points lower than <100cpm might be more suitable in this older adult population for distinguishing sedentary time from activity.

Paragraph 26 A pertinent finding of our study was that both the self-report and accelerometer measures of sedentary time were able to detect reductions following the intervention and were similarly responsive to change. This would suggest that the individual items we chose to include in the self-report measure were among the behaviors that people changed during the intervention. Notably, the amount of change detected was greater via self-report than accelerometer (51 vs 30 minutes/day). The greater change via self-report could be due to biased reporting; alternatively, the accelerometer may have failed to detect some sitting reductions (e.g. if sitting or lying down was replaced with standing still). The individual sedentary items were less responsive to change than total self-report sedentary time, with time watching TV and doing hobbies being the most responsive compared to the other items. Notably, most items had lower background variability than total sedentary time (both self-report and accelerometer-derived), so the lower responsiveness could relate to participants not changing these behaviors, rather than poor measurement quality. The lack of a control group in our pre-post study limits our ability to assess whether changes in individual items and total sedentary time are as a result of the intervention or if the participants are more accurately reporting their behavior. In our study, a larger RS (in absolute terms) reflects a greater magnitude of change observed over the intervention period, not necessarily a better ability to detect a minimum clinically meaningful change. Caution with these findings is recommended as the minimum clinically meaningful change is as yet unknown for sedentary time, and could be smaller, or larger, than what was achieved in this intervention.

Paragraph 27 The strengths of this study are our ability to report on responsiveness to change in addition to reliability and validity, and the use of a self-report measure designed specifically for older adults that assessed time in individual sedentary behaviors and total sedentary time. The non-normal distribution of individual sedentary time items and accelerometer-derived sedentary time limits the usefulness of the ICCs and responsiveness statistics reported for these items (which rely on assumptions of normality); however this was not a limitation for total self-report sedentary time. We did not have an appropriate criterion to assess validity of the individual sedentary items, and recommend that future studies use a combination of log and objective measures for this purpose (35). Although participants wore the accelerometer for at least half of the period covered by self-report, there was not total overlap between assessment periods which may have contributed to the less than ideal correlation and agreement between the self-report and accelerometer-derived sedentary time measures. We are also limited in our ability to generalize the findings given the small, non-random sample who were more highly educated than the general population of older Australian adults (1), and possibly had other differences related to motivation to participate in the intervention (38).

Paragraph 28 Our findings suggest that the summary measure of total sedentary time has good reliability but may be assessing behaviors that are not stable from week to week. The responsiveness is sufficiently high that the numbers of participants required to detect behavior change are achievable within the typical sample sizes of many behavioral interventions. Hence, in terms of responsiveness, the measure is suitable for detecting change following interventions, however the interpretation of changes that would be detected are questionable, given that we did not identify strong validity. This may be improved by providing specific prompts in the 'other' sedentary time item, such as for time spent eating. We administered the questionnaire via a face-to-face interview with non-working older adults. The viability of alternate delivery methods such as via the telephone or self-completion, and the utility of this questionnaire to evaluate change following interventions in younger adult populations across the leisure and transport domains of sedentary time should also be investigated.

Acknowledgments: This study was funded by a National Health and Medical Research Council (NHMRC) Program Grant (#569940) and Supplementary Research Funding from the School of Population Health at the University of Queensland. All authors are supported by a NHMRC Program Grant (#569940) and a Queensland Health Core Research Infrastructure Grant. In addition, Paul Gardiner is supported by a National Heart Foundation Postgraduate Scholarship (PP 06B 2889); Bronwyn Clark is supported by an Australian Postgraduate Award; Elizabeth Eakin is supported by a NHMRC Senior Research Fellowship (#511001); Genevieve Healy is supported by a NHMRC (#569861) / National Heart Foundation (PH 08B 3905) Postdoctoral Fellowship; and Neville Owen is supported by a NHMRC Senior Principal Research Fellowship (#1003960). The authors would like to sincerely thank the participants involved in this research.

Conflict of interest: The authors report no conflict of interest. The results of the present study do not constitute endorsement by ACSM.

References:

1. Australian Bureau of Statistics. Australian social trends 1999. Cat. No. 4102.0 [Internet] Canberra: ABS; [Cited 2010 October 29]. Available from: http://www.abs.gov.au/AUSSTATS/abs@.nsf/2f762f95845417aeca25706c00834efa/52ffc28684910e47ca2570ec00112c3e!OpenDocument.

2. Australian Bureau of Statistics. Time use on recreation and leisure activities, 2006. Cat No. 4173.0 [Internet] Canberra, Australia: ABS; [Cited 2010 October 29]. Available from:

http://www.ausstats.abs.gov.au/ausstats/subscriber.nsf/0/91FB93C8E82F220CCA25771F0018AE29/$File/41730_2006.pdf.

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

4. Bland JM, and Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res*. 1999;8(2):135-60.

5. Buman MP, Hekler EB, Haskell WL, Pruitt L, Conway TL, Cain KL, Sallis JF, Saelens BE, Frank LD, and King AC. Objective light-intensity physical activity associations with rated health in older adults. *Am J Epidemiology*. 2010;172(10):1155-65.

6. Chinapaw MJ, Slootmaker SM, Schuit AJ, van Zuidam M, and van Mechelen W. Reliability and validity of the Activity Questionnaire for Adults and Adolescents (AQuAA). *BMC Med Res Methodol*. 2009;9:58.

7. Clark BK, Sugiyama T, Healy GN, Salmon J, Dunstan DW, and 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.

8. Craig CL, Marshall AL, Sjostrom M, Bauman AE, Booth ML, Ainsworth BE, Pratt M, Ekelund U, Yngve A, Sallis JF, and Oja P. International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc*. 2003;35(8):1381-95.

9. Deng HB, Macfarlane DJ, Thomas GN, Lao XQ, Jiang CQ, Cheng KK, and Lam TH. Reliability and validity of the IPAQ-Chinese: the Guangzhou Biobank Cohort study. *Med Sci Sports Exerc*. 2008;40(2):303-7.

10. Dunstan DW, Barr EL, Healy GN, Salmon J, Shaw JE, Balkau B, Magliano DJ, Cameron AJ, Zimmet PZ, and Owen N. Television viewing time and mortality: the Australian Diabetes, Obesity and Lifestyle Study (AusDiab). *Circulation*. 2010;121(3):384-91.

11. Fjeldsoe BS, Marshall AL, and Miller YD. Measurement properties of the Australian Women's Activity Survey. *Med Sci Sports Exerc*. 2009;41(5):1020-33.

12. Forsen L, Loland NW, Vuillemin A, Chinapaw MJ, van Poppel MN, Mokkink LB, van Mechelen W, and Terwee CB. Self-administered physical activity questionnaires for the elderly: a systematic review of measurement properties. *Sports Med*. 2010;40(7):601-23.

13. Guyatt G, Walter S, and Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. *J Chronic Dis*. 1987;40(2):171-8.

14. Hagströmer M, Ainsworth BE, Oja P, and Sjöström M. Comparison of a subjective and an objective measure of physical activity in a population sample. *J Phys Act Health*. 2010;7(4):541-50.

15. Hart TL, Ainsworth BE, and Tudor-Locke C. Objective and subjective measures of sedentary behavior and physical activity. *Med Sci Sports Exerc*. 2011;43(3):449-56.

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

17. Katzmarzyk PT, Church TS, Craig CL, and Bouchard C. Sitting time and mortality from all causes, cardiovascular disease, and cancer. *Med Sci Sports Exerc*. 2009;41(5):998-1005.

18. Kolbe-Alexander TL, Lambert EV, Harkins JB, and Ekelund U. Comparison of two methods of measuring physical activity in South African older adults. *J Aging Phys Act*. 2006;14(1):98-114.

19. Kozey-Keadle S, Libertine A, Lyden K, Staudenmayer J, and Freedson P. Validation of wearable monitors for assessing sedentary behavior. *Med Sci Sports Exerc*. 2011;Jan 12 (published ahead of print).

20. Loney T, Standage M, Thompson D, Sebire SJ, and Cumming S. Self-report vs. objectively assessed physical activity: Which is right for public health? *J Phys Act Health*. 2011;8(1):62-70.

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

22. Matthews CE, Chen KY, Freedson PS, Buchowski MS, Beech BM, Pate RR, and Troiano RP. Amount of time spent in sedentary behaviors in the United States, 2003-2004. *Am J Epidemiol*. 2008;167(7):875-81.

23. Matton L, Wijndaele K, Duvigneaud N, Duquet W, Philippaerts R, Thomis M, and Lefevre J. Reliability and validity of the Flemish Physical Activity Computerized Questionnaire in adults. *Res Q Exerc Sport*. 2007;78(4):293-306.

24. Norman A, Bellocco R, Bergstrom A, and Wolk A. Validity and reproducibility of self-reported total physical activity--differences by relative weight. *Int J Obes Relat Metab Disord*. 2001;25(5):682-8.

25. Otten JJ, Littenberg B, and Harvey-Berino JR. Relationship between self-report and an objective measure of television-viewing time in adults. *Obesity*. 2010;18(6):1273-5.

26. Owen N, Bauman A, and Brown W. Too much sitting: a novel and important predictor of chronic disease risk? *Br J Sports Med*. 2009;43(2):81-3.

27. Owen N, Ekelund U, Hamilton M, Gardiner P, and Dunstan DW. Sedentary behavior in adults: longitudinal, experimental, and intervention evidence. *J Phys Act Health*. 2010;7(Suppl 3):S334-6.

28. Patel AV, Bernstein L, Deka A, Feigelson HS, Campbell PT, Gapstur SM, Colditz GA, and Thun MJ. Leisure time spent sitting in relation to total mortality in a prospective cohort of US adults. *Am J Epidemiol*. 2010;172(4):419-29.

29. Reeves MM, Marshall AL, Owen N, Winkler EA, and Eakin EG. Measuring physical activity change in broad-reach intervention trials. *J Phys Act Health*. 2010;7(2):194-202.

30. Rikli RE. Reliability, validity, and methodological issues in assessing physical activity in older adults. *Res Q Exerc Sport*. 2000;71(2 Suppl):S89-96.

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

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

33. Shrout PE, and Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979;86(2):420-8.

34. Tuley MR, Mulrow CD, and McMahan CA. Estimating and testing an index of responsiveness and the relationship of the index to power. *J Clin Epidemiol*. 1991;44(4-5):417-21.

35. van der Ploeg HP, Merom D, Chau JY, Bittman M, Trost SG, and Bauman AE. Advances in population surveillance for physical activity and sedentary behavior: reliability and validity of time use surveys. *Am J Epidemiol*. 2010;172(10):119-1206.

36. van Poppel MN, Chinapaw MJ, Mokkink LB, van Mechelen W, and Terwee CB. Physical activity questionnaires for adults: a systematic review of measurement properties. *Sports Med*. 2010;40(7):565-600.

37. Warren TY, Barry V, Hooker SP, Sui X, Church TS, and Blair SN. Sedentary behaviors increase risk of cardiovascular disease mortality in men. *Med Sci Sports Exerc*. 2010;42(5):879-85.

38. Waters LA, Galichet B, Owen N, and Eakin EG. Who participates in physical activity intervention trials? *J Phys Act Health*. 2011;8(1):85-103.

Supplemental Digital Content 1.doc: Appendix that details the sedentary time questions. Questionnaire to measure sedentary time in older adults

Figure legend:

Figure: Bland-Altman plot of agreement of total self-report sedentary time with accelerometer-derived sedentary time in older adults (n=48).

The y axis is the difference between the two measures (self-report - accelerometer-derived sedentary time) and the x axis is the average of the two measures ([total self-report sedentary time + accelerometer-derived sedentary time]/2), both in hours per day. The solid line shows the mean difference between the two measures (-9.20 hours/day + 0.67 * average sedentary time), with the dashed lines representing the limits of agreement (mean difference ± 3.82 hours/day).

Appendix:

The questionnaire used in *Stand Up For Your Health* to assess sedentary time in older adults.

I am going to ask you about activities you did over the ***last week whilst sitting or lying down***. Don't count the time you spent in bed. So today is _____, I want you to think about the time from last _____ to yesterday.

For each of the activities only count the time where this was your main activity; for example if you are watching TV and doing a crossword, count it as TV time or crossword time but not as both.

During the last week, how much time in total did you spend ***sitting or lying down and****……?*

SEDENTARY ITEM	**TIME**
Watching television or videos/DVDs	___ HOURS ___MINUTES
Using the computer/Internet	___ HOURS ___MINUTES
Reading	___ HOURS ___MINUTES
Socializing with friends or family	___ HOURS ___MINUTES
Driving or riding in a car, or time on public transport	___ HOURS ___MINUTES
Doing hobbies, e.g. craft, crosswords	___ HOURS ___MINUTES
Doing any other activities	___ HOURS ___MINUTES

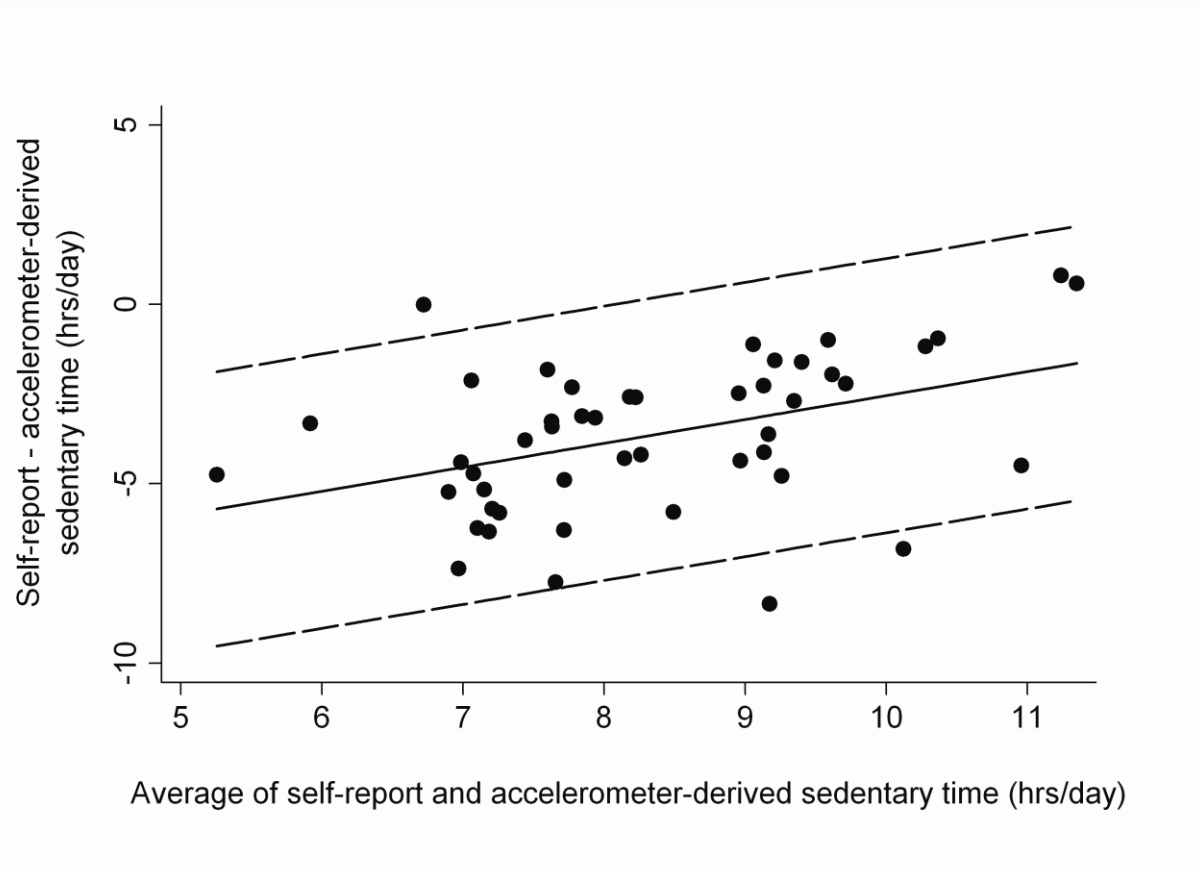
Self-report - accelerometer-derived sedentary time (hrs/day)
5
0
-5
-10
5
6
7
8
9
10
11
Average of self-report and accelerometer-derived sedentary time (hrs/day)

Table 1 Duration (hours/day) and test-retest reliability of self-reported sedentary behaviors in older adults (n=48).

Variable	Duration, n who report not doing the activity			Test –retest (T1 versus T2)	
	T1	**T2**	**T3**	**Spearman's ρ (95% CI)**	**ICC[a] (95% CI)**
TV viewing	2.9 (1.7, 3.5), n=0	2.1 (1.4, 3.3), n=0	1.8 (1.3, 2.5), n=0	0.78 (0.63, 0.89)	0.76 (0.62, 0.86)
Computer use	0.5 (0.0, 1.1), n=17	0.3 (0.0, 0.8), n=14	0.2 (0.0, 0.6), n=16	0.90 (0.83, 0.94)	0.79 (0.65, 0.88)
Reading	1.0 (0.6, 2.0), n=5	0.7 (0.5, 1.5), n=5	0.6 (0.3, 1.6), n=9	0.77 (0.62, 0.86)	0.74 (0.51, 0.86)
Socialising	1.1 (0.7, 1.5), n=3	1.1 (0.6, 1.6), n=5	1.0 (0.4, 1.7), n=3	0.38 (0.11, 0.60)	0.38 (0.11, 0.60)
Transport	0.4 (0.2, 0.7), n=1	0.4 (0.2, 0.7), n=2	0.4 (0.2, 0.8), n=0	0.45 (0.19, 0.65)	0.40 (0.14, 0.61)
Hobbies	0.5 (0.0, 0.9), n=15	0.4 (0.0, 0.8), n=15	0.2 (0.0, 0.5), n=17	0.61 (0.39, 0.76)	0.35 (0.07, 0.58)
'Other' sedentary	0.0 (0.0, 0.0), n=37	0.0 (0.0, 0.0), n=41	0.0 (0.0, 0.0), n=34	0.23 (0.38, 0.74)	0.04 (-0.25, 0.32)
Total self-report sedentary time	7.3 (1.8)	6.5 (2.1)	5.7 (2.0)	0.56 (0.33, 0.73)	0.52 (0.27, 0.70)

Data are reported as median (25^{th}, 75^{th} percentile) with the exception of total self-report sedentary time which is presented as mean (SD).
[a] Intraclass correlations (single-measures with an absolute agreement definition, calculated using two-way mixed effects models).

Table 2 Responsiveness, change of self-reported and accelerometer-derived sedentary time in older adults (n=48).

Variable (all hours/day)	**% substantial change[a]**	**Mean pre-, post - intervention change (95% CI)**	**√(2*MSE)[b]**	**RS[c]**
Accelerometer-derived sedentary time	45.8	-0.50 (-0.68, -0.32)	1.29	0.39
Total self-report sedentary time	60.4	-0.85 (-1.40, -0.30)	1.80	0.47
TV viewing	43.8	-0.35 (-0.68, -0.02)	1.03	0.34
Computer use	25.0	-0.03 (-0.21, 0.16)	0.49	0.06
Reading	35.4	-0.13 (-0.29, 0.02)	0.63	0.21
Socialising	37.5	-0.09 (-0.33, 0.14)	0.83	0.11
Transport	22.9	0.10 (-0.11, 0.30)	0.53	0.18
Hobbies	33.3	-0.25 (-0.45, -0.04)	0.75	0.33
‘Other’ sedentary time	8.3	-0.09 (-0.33, 0.15)	0.82	0.11

[a] percentage of participants who decreased by at least 15 minutes per day (each individual behavior) or 30 minutes per day (total and accelerometer-derived sedentary time)
[b] √(2*MSE) is a measure of variability in the pre-intervention period. Mean square error (MSE) is from repeated measures ANOVA of the pre-intervention data.
[c] Responsiveness statistic (RS) is calculated as (mean change / √[2 x MSE]), with direction of change removed.