Advances in Population Surveillance for Physical Activity and Sedentary Behavior: Reliability and Validity of Time Use Surveys

Hidde P. van der Ploeg*, Dafna Merom, Josephine Y. Chau, Michael Bittman, Stewart G. Trost, and Adrian E. Bauman

* Correspondence to Dr. Hidde P. van der Ploeg, Cluster for Physical Activity and Health, Level 2, Medical Foundation Building (K25), Sydney School of Public Health, University of Sydney, Sydney, New South Wales 2006, Australia (e-mail: hiddep@health.usyd.edu.au).

Initially submitted April 9, 2010; accepted for publication July 13, 2010.

Many countries conduct regular national time use surveys, some of which date back as far as the 1960s. Time use surveys potentially provide more detailed and accurate national estimates of the prevalence of sedentary and physical activity behavior than more traditional self-report surveillance systems. In this study, the authors determined the reliability and validity of time use surveys for assessing sedentary and physical activity behavior. In 2006 and 2007, participants (n = 134) were recruited from work sites in the Australian state of New South Wales. Participants completed a 2-day time use diary twice, 7 days apart, and wore an accelerometer. The 2 diaries were compared for test-retest reliability, and comparison with the accelerometer determined concurrent validity. Participants with similar activity patterns during the 2 diary periods showed reliability intraclass correlations of 0.74 and 0.73 for nonoccupational sedentary behavior and moderate/vigorous physical activity, respectively. Comparison of the diary with the accelerometer showed Spearman correlations of 0.57–0.59 and 0.45–0.69 for nonoccupational sedentary behavior and moderate/vigorous physical activity, respectively. Time use surveys appear to be more valid for population surveillance of nonoccupational sedentary behavior and health-enhancing physical activity than more traditional surveillance systems. National time use surveys could be used to retrospectively study nonoccupational sedentary and physical activity behavior over the past 5 decades.

epidemiologic measurements; exercise; population surveillance; public health; sedentary lifestyle

Abbreviations: ICC, intraclass correlation coefficient; MET, metabolic equivalent; NHANES, National Health and Nutrition Examination Survey.

Regular moderate- to vigorous-intensity physical activity is known to reduce the risk of many chronic diseases, including cardiovascular disease, colon and breast cancer, diabetes, and obesity (1). While physical activity is already a widely accepted public health priority, sedentary behavior is quickly emerging as an important risk factor for the development of a number of chronic diseases, independently of physical activity (2). Hence, monitoring of both population physical activity and sedentary behavior levels is essential for public health.

Traditional population physical activity surveillance systems, such as national health surveys, are mostly based on self-report physical activity recall questionnaires, which are prone to reporting error and social desirability bias (3). Furthermore, space constraints in national health surveys allow a limited level of detail for physical activity assessment, and sedentary behavior has been mostly overlooked. For these reasons, the National Center for Health Statistics implemented accelerometers in the 2003–2004 National Health and Nutrition Examination Survey (NHANES) in order to objectively assess physical activity and sedentary behavior (4). However, this was an expensive and logistically complex surveillance exercise, which will be harder to replicate in other countries and in future NHANES surveys. Additionally, the accelerometer data do not allow for study of physical activity behaviors across different domains.

Recently, the possibility of utilizing time use surveys for surveillance of population physical activity and sedentary...
behavior has been explored (5, 6). Time use surveys have their origin in sociology and are conducted regularly in population-representative samples in many countries (7, 8). Time use surveys generally ask participants to keep a diary in which they record what they are doing in (usually) 5-minute intervals over the course of 1 or several days; the information is then coded into an activity classification with more than 200 different activity codes. There are a number of important advantages of using time use surveys for physical activity surveillance in comparison with more traditional self-report surveillance questionnaires.

Firstly, the data obtained from time use surveys provide an enormous level of detail, allowing investigators to study both physical activity and sedentary behaviors over a range of different domains, as well as to study specific behaviors such as active transportation or television viewing. However, assigning intensity levels to the activity classification is sometimes difficult, since time use surveys were not created with physical activity surveillance in mind. For occupational activities, assigning an intensity level is especially difficult because these activities are generally just coded as “occupational,” without further clues as to what actually happened.

Secondly, time use diaries are less susceptible to measurement bias than the more traditional surveillance questionnaires (6, 9). Since the participant is asked for a comprehensive account of activities engaged in throughout the day and the survey does not prompt the participant specifically for information on physical activity and sedentary behaviors, responses will probably be less affected by social desirability bias. Furthermore, recall errors are limited, since the time use diary is completed throughout the day or at the end of the day.

Thirdly, time use surveys have been conducted in population-representative samples since the 1960s in many countries (including Australia, Canada, the United States, and many African, Asian, European, and Latin American countries). Hence, they appear to have the potential to aid in the study of national trends in physical activity and sedentary behavior over the past 5 decades. However, changes and differences in methods between and within countries might not always make this an easy task (6).

Given these advantages of time use surveys, there has recently been increased interest in utilizing time use surveys in research on physical activity and sedentary behavior (5, 6, 10–14). For example, a recent study commissioned by the National Cancer Institute and the National Institutes of Health (14) assigned intensity levels to the activity classification of the American Time Use Survey by linking it to the Compendium of Physical Activities (15). Despite this increasing body of work, little is known about the reliability and validity of time use surveys for classifying activities into intensity levels in order to study population levels of physical activity and sedentary behavior. Hence, our objective in the current study was to determine the test-retest reliability and concurrent validity of time use surveys in assessing sedentary, light-intensity, and moderate- and vigorous-intensity activities across different domains.

MATERIALS AND METHODS

Participants

Eligible study participants were adults in the Australian state of New South Wales. A convenience sample of participants was recruited between August 2006 and July 2007, through 6 New South Wales work sites and via word of mouth. For inclusion, a person had to be aged 18 years or older and had to speak English sufficiently well to successfully complete the study materials. All workers from participating work sites received an invitation by e-mail to join the study. People who expressed further interest in the study (n = 293) received the participant information sheet. Of those interested, 172 people volunteered to participate and gave written informed consent, and 148 (86%) completed the study. The study protocol was approved by the University of Sydney Human Research Ethics Committee.

Measurements

Participants wore an accelerometer for 10 days to assess their physical activity and sedentary behavior. During days 2 and 3, participants completed the first 2-day time use diary. The second time use diary was completed on days 9 and 10, matching the same days of the week as the first time use diary. At the end of the second time use diary, participants were asked how different their activity patterns in the past 2 days had been in comparison with the same 2 days during the previous week; the response was scored on a 10-point Likert-type scale (1 = not different at all, 10 = very different).

All study materials were mailed to the participant and returned via paid reply envelope after the 10-day measurement period. Reminder text messages were sent to the participant’s mobile phone on days when he or she had to start using the accelerometer or completing the first or second time use diary. Start days were allocated in such a way that all days of the week were equally represented for the diaries in the study sample.

At the start of the measurement period, participants completed a questionnaire on personal characteristics. Body mass index (weight (kg)/height (m)²) was calculated from self-reported height and weight. In order to determine general activity patterns during work time, participants were asked whether, at work, they were mostly sitting or standing, mostly walking, or mostly performing heavy labor or physically demanding tasks (16).

Time use diary. For the time use diary, we used the 2006 Australian Bureau of Statistics time use survey, which consists of a 2-day diary formatted into 5-minute time intervals and is completed during or at the end of each day (17). Subjects recorded in their own words their primary activity, any secondary activity they were engaged in at the same time, the location of the activity, other persons present during the activity, and who they did this activity for. Recorded activities were classified by a trained coder into a nesting 3-digit activity classification using the Australian Bureau of Statistics 2006 time use survey diary coding rules (17). The classification contained 221 activity codes and was derived

Am J Epidemiol 2010;172:1199–1206
from an internationally established classification that forms the basis of the majority of time use surveys in the world (7). Activities were then assigned an intensity classification based on the Compendium of Physical Activities (15) and the intensity coding system used in the American Time Use Survey (14). The intensity classification consisted of 4 categories based on metabolic equivalent (MET) energy expenditure (1 MET = 4.184 kJ per kilogram of body weight per hour): sedentary activities (≤1.5 METs), light activities (>1.5–<3 METs), moderate and vigorous activities (≥3 METs), and occupational activities. The occupational category was added because the occupational time use activity codes are too generic to allow for the assignment of an intensity category. All activities were also recoded into categories based on the intensity coding system used in the American Time Use Survey (14). The intensity classification consisted of 4 categories based on metabolic equivalent (MET) energy expenditure (1 MET = 4.184 kJ per kilogram of body weight per hour): sedentary activities (≤1.5 METs), light activities (>1.5–<3 METs), moderate and vigorous activities (≥3 METs), and occupational activities. The occupational category was added because the occupational time use activity codes are too generic to allow for the assignment of an intensity category. All activities were also recoded into domains commonly used in physical activity research (occupational, household, leisure, and transportation).

**Accelerometer.** Physical activity and sedentary behavior were assessed with an ActiGraph GT1M accelerometer (ActiGraph, Pensacola, Florida), which was worn on the right hip during waking hours for the 10-day measurement period. This small (38 mm × 37 mm × 18 mm), lightweight (27 g) accelerometer detected accelerations ranging from 0.05 × g to 2 × g, with a frequency response of 0.25–2.5 Hz to limit the measurement of nonhuman movements. The accelerations were converted into activity counts for every 1-minute epoch, which were then used to calculate time spent in sedentary (<100 counts/minute), light (100–759 counts/minute), and moderate and vigorous activities (≥760 counts/minute) on the basis of the most frequently used sedentary cutpoint (18) and the moderate-intensity cutpoint of Matthews (19).

**Statistical analysis**

All analyses were performed in SPSS, version 17.0 (SPSS, Inc., Chicago, Illinois). Intraclass correlation coefficients (ICCs) (2-way mixed consistency model) were used to compare the first and second 2-day time use diaries for test-retest reliability. ICCs were classified as low (<0.30), moderate (0.30–0.59), or high (≥0.60) (20). To correct for actual changes in behavior between the first and second diaries, the analysis was repeated separately for people who reported that their activity patterns were similar during the first and second diary periods and those who reported that their activity patterns during the 2 periods were quite different (scores of ≥5 on the question “How different were the last 2 days compared with the same 2 days last week?”).

For the validity analysis, the accelerometer was used as the criterion measure of physical activity because of the lack of a gold standard. The accelerometer data were temporally linked to each activity in the participant’s time use diary using the recorded starting and stopping times of the activity. For each activity, periods of consecutive zero activity counts lasting for at least 10 minutes were considered time in which the participant did not wear the accelerometer. Activities for which the accelerometer was not worn for more than 25% of the activity’s duration were excluded from the analysis. This exclusion criterion is stricter than the commonly used minimal wearing requirements for accelerometers (21). Activities that lasted only 5 minutes (the minimum duration in the diary) and had no activity counts were also excluded from the analysis. In addition, all occupational activities were excluded from the validity analyses, because intensity categories could not be assigned to them on the basis of the time use diary. The concurrent validity of the 2-day time use survey in estimating the relative amounts of time spent in sedentary, light, and moderate and vigorous activity was determined by comparison with the accelerometer using Spearman correlations. In order to check for consistency in the validity correlation coefficients, we assessed validity separately for the first and second 2-day time use diaries. We also performed subgroup analyses for gender, age (<45 years, ≥45 years), and body mass index (<25, ≥25).

**RESULTS**

Of the 148 people who completed the study, 134 had complete test-retest reliability data; of those, 129 also had complete validity data. Table 1 shows that the study sample had reasonably representative gender, age, and body mass index distributions. The majority of participants had a higher

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>No.</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>79</td>
<td>59.0</td>
</tr>
<tr>
<td>Female</td>
<td>55</td>
<td>41.0</td>
</tr>
<tr>
<td>Age, years</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18–30</td>
<td>26</td>
<td>19.4</td>
</tr>
<tr>
<td>31–40</td>
<td>30</td>
<td>22.4</td>
</tr>
<tr>
<td>41–50</td>
<td>49</td>
<td>36.6</td>
</tr>
<tr>
<td>51–63</td>
<td>29</td>
<td>21.6</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Some high school</td>
<td>5</td>
<td>3.7</td>
</tr>
<tr>
<td>Completion of high school</td>
<td>8</td>
<td>6.0</td>
</tr>
<tr>
<td>Trade/technical certificate or diploma</td>
<td>32</td>
<td>23.9</td>
</tr>
<tr>
<td>University</td>
<td>89</td>
<td>66.4</td>
</tr>
<tr>
<td>Self-reported body mass index&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal-weight (&lt;25)</td>
<td>67</td>
<td>50.0</td>
</tr>
<tr>
<td>Overweight (25–29)</td>
<td>46</td>
<td>34.3</td>
</tr>
<tr>
<td>Obese (≥30)</td>
<td>18</td>
<td>13.4</td>
</tr>
<tr>
<td>Missing data</td>
<td>3</td>
<td>2.2</td>
</tr>
<tr>
<td>Current occupational activity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mostly sitting or standing</td>
<td>126</td>
<td>94.0</td>
</tr>
<tr>
<td>Mostly walking</td>
<td>3</td>
<td>2.2</td>
</tr>
<tr>
<td>Mostly heavy labor or physically demanding work</td>
<td>2</td>
<td>1.5</td>
</tr>
<tr>
<td>Not employed</td>
<td>3</td>
<td>2.2</td>
</tr>
</tbody>
</table>

<sup>a</sup> Weight (kg)/height (m)<sup>2</sup>.
education, and almost all had an occupation that involved mostly sitting or standing.

Test-retest reliability

Table 2 shows moderate test-retest reliability ICCs for correlation between the first and second time use diaries in the whole study sample. However, for people who reported that their activity patterns were similar during the first and second assessments, mostly high correlations were found. For people with similar activity patterns, test-retest reliability ICCs were close to 0.75 for nonoccupational sedentary and moderate- and vigorous-intensity activities. The test-retest reliability of light-intensity activities was lower (around 0.50).

Concurrent validity

Table 3 shows the concurrent validity of the time use diary for estimating time spent in sedentary, light-intensity, and moderate- and vigorous-intensity activities when compared with the accelerometer. The time use diary showed moderate correlations with the accelerometer for time spent in sedentary, light-intensity, moderate- and vigorous-intensity activities when compared with the accelerometer. The time use diary showed similar validity Spearman correlation coefficients within the different domain and intensity categories.

Validity correlations for light-intensity activity were low to moderate and ranged from −0.02 to 0.39. Total nonoccupational light-intensity physical activity showed significant correlations for both time use diaries (0.39 and 0.27, respectively). Time spent in moderate- and vigorous-intensity activities showed generally moderate-to-high Spearman correlations between the time use diary and the accelerometer. Total nonoccupational moderate- and vigorous-intensity physical activity showed correlations of 0.45 and 0.69 for the first and second time use diaries, respectively. In general, the first and second administrations of the time use diary showed similar validity Spearman correlation coefficients within the different domain and intensity categories. Subgroup analyses for gender, age (≤45 years, ≥45 years), and body mass index (<25, ≥25) showed differences that were consistent across the 2 diary validations only for total nonoccupational light-intensity and moderate- and vigorous-intensity activities between men and women (results not shown in table). Women showed validity correlations for the first and second diary validations of 0.32 ($P < 0.05$) and 0.40 ($P < 0.05$) for light-intensity activities and 0.55 ($P < 0.05$) and 0.40 ($P < 0.05$) for moderate/vigorous activities, respectively. Men showed correlations of 0.16 ($P > 0.05$) and 0.31 ($P < 0.05$) for light activities and 0.74 ($P < 0.05$) and 0.47 ($P < 0.05$) for moderate/vigorous activities, respectively.

DISCUSSION

The current study suggests that time use surveys have acceptable measurement properties for assessing nonoccupational sedentary and physical activity behavior, but the results are less clear for light-intensity activity. The test-retest reliability of the time use survey was similar to that of traditional physical activity surveillance questionnaires such as the US Behavioral Risk Factor Surveillance System physical activity questionnaire (22), the widely used International Physical Activity Questionnaire (23), and the Global Physical Activity Questionnaire endorsed by the

Table 2. Test-Retest Reliability of Time Use Diary Data for Participants With Similar and Dissimilar Activity Patterns at Test and Retest ($n = 134$), Sydney, Australia, 2006–2007

<table>
<thead>
<tr>
<th>Activity Domain and Intensity Level</th>
<th>Mean Amount of Time Spent in Activity Over 2 Days, minutes (SD)</th>
<th>Comparison Between Diary 1 and Diary 2 (Test-Retest Reliability)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Diary 1</td>
<td>Diary 2</td>
</tr>
<tr>
<td></td>
<td>ICC</td>
<td>95% CI</td>
</tr>
<tr>
<td>Nonoccupational activity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sedentary</td>
<td>0.74</td>
<td>0.60, 0.84</td>
</tr>
<tr>
<td>Light intensity</td>
<td>0.46</td>
<td>0.23, 0.64</td>
</tr>
<tr>
<td>Moderate and vigorous intensity</td>
<td>0.73</td>
<td>0.59, 0.83</td>
</tr>
<tr>
<td>Occupational activity</td>
<td>0.87</td>
<td>0.80, 0.92</td>
</tr>
<tr>
<td>Household activity</td>
<td>0.71</td>
<td>0.56, 0.82</td>
</tr>
<tr>
<td>Leisure activity</td>
<td>0.79</td>
<td>0.67, 0.87</td>
</tr>
<tr>
<td>Transportation activity</td>
<td>0.61</td>
<td>0.43, 0.75</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; ICC, intraclass correlation coefficient; SD, standard deviation.

a Participants who scored 1–4 on the question “How different were the last 2 days compared with the same 2 days last week?” at the end the second diary period (1 = not different at all, 10 = very different).

b Participants who scored 5–10 on the question “How different were the last 2 days compared with the same 2 days last week?” at the end the second diary period (1 = not different at all, 10 = very different).
When compared with an accelerometer, the time use survey performed substantially better (concurrent validity correlations in the 0.45–0.69 range for nonoccupational sedentary and moderate- and vigorous-intensity activities) than the traditional surveillance questionnaires, which have generally shown concurrent validity correlations in the 0.2–0.3 range for moderate- and vigorous-intensity physical activity and sitting time (22–25). This finding seems to support the notion that time use surveys would be less susceptible to reporting errors and social desirability bias than traditional physical activity surveillance questionnaires (6).

The time use diary exhibited stronger validity coefficients for leisure- and transportation-related moderate- and vigorous-intensity physical activity and sitting time (22–25). This finding seems to support the notion that time use surveys would be less susceptible to reporting errors and social desirability bias than traditional physical activity surveillance questionnaires (6).

### Table 3. Concurrent Validity of Time Use Diary Information Compared with Accelerometer Data (n = 129), Sydney, Australia, 2006–2007

<table>
<thead>
<tr>
<th>Activity Domain and Intensity Level</th>
<th>Mean Amount of Time* Spent in Activity Over 2 Days, minutes (SD)</th>
<th>Spearman’s ( r ) (Diary 1 vs. Accelerometer 1)</th>
<th>Mean Amount of Time* Spent in Activity Over 2 Days, minutes (SD)</th>
<th>Spearman’s ( r ) (Diary 2 vs. Accelerometer 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Diary 1</td>
<td>Accelerometer 1</td>
<td></td>
<td>Diary 2</td>
</tr>
<tr>
<td>Household activity</td>
<td></td>
<td></td>
<td>Leisure activity</td>
<td></td>
</tr>
<tr>
<td>Sedentary</td>
<td>144 (91)</td>
<td>210 (104)</td>
<td>Sedentary</td>
<td>164 (138)</td>
</tr>
<tr>
<td>Light intensity</td>
<td>224 (149)</td>
<td>157 (96)</td>
<td>Light intensity</td>
<td>1 (12)</td>
</tr>
<tr>
<td>Moderate and vigorous intensity</td>
<td>79 (105)</td>
<td>74 (58)</td>
<td>Moderate and vigorous intensity</td>
<td>54 (72)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leisure activity</td>
<td>213 (172)</td>
<td>164 (133)</td>
<td>Leisure activity</td>
<td>164 (138)</td>
</tr>
<tr>
<td>Sedentary</td>
<td>20 (60)</td>
<td>67 (47)</td>
<td>Sedentary</td>
<td>1 (12)</td>
</tr>
<tr>
<td>Light intensity</td>
<td>208 (142)</td>
<td>138 (88)</td>
<td>Light intensity</td>
<td>227 (159)</td>
</tr>
<tr>
<td>Moderate and vigorous intensity</td>
<td>76 (100)</td>
<td>68 (58)</td>
<td>Moderate and vigorous intensity</td>
<td>48 (65)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transportation activity</td>
<td></td>
<td></td>
<td>Transportation activity</td>
<td></td>
</tr>
<tr>
<td>Sedentary</td>
<td>164 (138)</td>
<td>120 (94)</td>
<td>Sedentary</td>
<td>164 (138)</td>
</tr>
<tr>
<td>Light intensity</td>
<td>20 (60)</td>
<td>67 (47)</td>
<td>Light intensity</td>
<td>20 (60)</td>
</tr>
<tr>
<td>Moderate and vigorous intensity</td>
<td>54 (72)</td>
<td>49 (49)</td>
<td>Moderate and vigorous intensity</td>
<td>54 (72)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total nonoccupational activity</td>
<td></td>
<td></td>
<td>Total nonoccupational activity</td>
<td></td>
</tr>
<tr>
<td>Sedentary</td>
<td>523 (246)</td>
<td>495 (194)</td>
<td>Sedentary</td>
<td>523 (246)</td>
</tr>
<tr>
<td>Light intensity</td>
<td>244 (158)</td>
<td>277 (121)</td>
<td>Light intensity</td>
<td>227 (159)</td>
</tr>
<tr>
<td>Moderate and vigorous intensity</td>
<td>185 (124)</td>
<td>168 (90)</td>
<td>Moderate and vigorous intensity</td>
<td>175 (118)</td>
</tr>
</tbody>
</table>

Abbreviations: NA, not applicable; SD, standard deviation.
* \( P < 0.05 \).
\( a \) Only including time that the accelerometer was considered to have been worn during the 2 measurement days.
light-intensity activity, especially in specific activity domains. However, total nonoccupational light-intensity activity still showed reasonable correlations (0.39 and 0.27 for the 2 diary validations), and perhaps future estimations of light-intensity activity with time use diaries need to be delimited to this domain.

A related issue with the time use survey is the lumping of activities into 1 category, which makes the assignment of 1 specific MET value to each activity code difficult. For example, the code “organized sport” lumps all organized sport activities together regardless of their intensity, which can easily range from 3 METs to 12 METs. For this reason, we opted to classify each activity code with the 3 intensity categories “sedentary,” “light,” and “moderate and vigorous,” rather than make a specific MET classification. However, for all occupational codes, the activity codes are so general (e.g., “main job—usual hours at work”) that the assignment of intensity categories is almost impossible. This knowledge gap with regard to occupational activities is probably the main limitation of time use surveys, and it makes the estimation of total sedentary and physical activity levels more problematic. However, some rough indications about activity intensity levels during work could still be inferred on the basis of job titles (14). The level of detail provided in the activity codes of the time use diary also did not allow us to distinguish between moderate- and vigorous-intensity activities. Hence, a combined moderate-and-vigorous category was used in the intensity classification, which encompasses the nonoccupational health-enhancing spectrum of physical activity that is important for public health.

The somewhat low test-retest reliability correlation coefficients in the total study sample (~0.5) were probably due to actual individual changes in behavior between the first and second administrations of the diary. When splitting the reliability analysis into people who changed and did not change their behavior much between the first and second assessments, the reliability correlations were much better (~0.75) for persons who did not change their behavior. Hence, the time use survey seems reliable, especially for sedentary and moderate- and vigorous-intensity activities, but it seems to be sensitive to individual changes in behavior, possibly because of the relatively short assessment period of just 2 days. It is known from the literature that at least 3–4 days and preferably 1 week of accelerometer measurements is needed to obtain a representative picture of a person’s physical activity behavior (26). However, the time use surveys were conducted for the purpose of population surveillance, not individual assessment. The large sample sizes generally used in time use surveys would most likely level out individual behavioral misclassifications that resulted from the short assessment period, and would consequently give accurate population-level estimates. A limitation of the time use diary would be that it would be more complicated to determine the proportion of the population meeting the physical activity recommendations of at least 30 minutes of moderate-intensity physical activity on at least 5 days (but preferably all days) of the week. Since most time use surveys consist of just 1 or 2 diary days, the proportion meeting the recommendation could probably be best estimated by translating it to 30 minutes of moderate-intensity physical activity per day.

Because the majority of time use surveys in the world are based on the internationally established activity classification that was developed under the leadership of Szalai (7), the current study results obtained using the Australian time use survey are generalizable to time use surveys used in other countries. Nevertheless, differences in data collection between and within countries can make direct comparisons of reported physical activity and sedentary behavior levels more complicated. In addition, the high level of education in the current study sample could potentially have led to some overestimation of the measurement properties of time use surveys. However, the Behavioral Risk Factor Surveillance System and International Physical Activity Questionnaire validation studies also had more highly educated samples (22, 23), suggesting that the relatively better validity of the time use diaries was not a result of selection bias.

A strength of the study was that the time use survey was validated twice, with the second validation of the time use diary mostly confirming the results from the first assessment. The double validation was also instrumental for determining differences in the measurement properties of the time use survey between subgroups. Most differences between subgroups in the first validation were reversed in the second validation. The results consistently suggested only that nonoccupational light-intensity activity was somewhat better estimated for women than for men and that moderate- and vigorous-intensity activity was less well estimated in women than in men.

A limitation of the current study was the potential mismatch in the temporal linkage for the validity analyses of the activities recorded in the diary with the accelerometer. The linkage was based on the self-reported starting and stopping times recorded in the time use diary, which might have led to some temporal mismatching with the accelerometer data. This could have caused underestimation of the concurrent validity of the domain-specific activities, as these might have been matched with the wrong intensities recorded by the accelerometer. However, this matching of self-reported and accelerometer data is an important strength of the current study, as it allowed us to determine the concurrent validity of sedentary and physical activity assessments by domain. This validation with accelerometers by different domains is unique, because accelerometers do not provide any information about domains, and temporal linkage with questionnaires is not possible, since these generally do not record the starting and stopping times of activities. Furthermore, the temporal linkage also allowed us to exclude any missing data (mostly from the accelerometer), which increased the quality of the validation in comparison with validation studies of traditional surveillance questionnaires. It is likely that this higher-quality validation method did account for some proportion of the better concurrent validity of the time use survey as compared with traditional physical activity surveillance surveys.

Although accelerometers are frequently used as a criterion measure in physical activity validation studies, they are not a gold standard for the assessment of free-living physical activity and have some limitations of their own. An
accelerometer underestimates upper body movements and activities without a strong vertical component, such as cycling (27). Water-based activities are not recorded, because the accelerometer is not worn in the water. Furthermore, there is a large range of different regression models for the conversion of activity counts into time spent in intensity levels, but all of these regression models produce quite different absolute estimates (19). Hence, comparison between the time use diary and the accelerometer for the absolute amounts of time spent in different intensity categories (as presented in Table 3) is not useful, as the absolute accelerometer estimates would have been quite different if another regression model had been used.

In conclusion, the results of this study suggest that time use surveys can be used to study sedentary behavior and health-enhancing physical activity across the leisure, transportation, and household domains. Time use surveys have the potential to provide more detailed and accurate estimates of population-level nonoccupational sedentary and physical activity than traditional self-report surveillance systems. This opens possibilities to utilize national time use surveys to retrospectively study nonoccupational sedentary and physical activity behavior over the past 5 decades, which could be of great importance to public health, since sedentary and physical activity behaviors are important population-wide health priorities.

ACKNOWLEDGMENTS

Author affiliations: Cluster for Physical Activity and Health, Sydney School of Public Health, University of Sydney, Sydney, New South Wales, Australia (Hidde P. van der Ploeg, Dafna Merom, Josephine Y. Chau, Adrian E. Bauman); Discipline of Sociology, School of Behavioural, Cognitive and Social Sciences, University of New England, Armidale, New South Wales, Australia (Michael Bittman); and Department of Nutrition and Exercise Sciences, College of Health and Human Sciences, Oregon State University, Corvallis, Oregon (Stuart G. Trost).

This work was supported by a National Health and Medical Research Council of Australia program grant (grant 301200).

The authors thank Belinda von Hofe, Dr. Judith Clarke, Sylvia Groen, and Marloes van Duuren for their assistance with data collection.

Conflict of interest: none declared.

REFERENCES


