Walking Objectively Measured: Classifying Accelerometer Data with GPS and Travel Diaries

Bumjoon Kang[1], **Anne V. Moudon**[1], **Philip M. Hurvitz**[1], **Lucas Reichley**[2], and **Brian E. Saelens**[2,3]

[1]Urban Form Lab and the Department of Urban Design and Planning, University of Washington, Seattle, WA

[2]Seattle Children's Research Institute, Seattle, WA

[3]Department of Pediatrics, University of Washington, Seattle, WA

Abstract

Purpose—This study developed and tested an algorithm to classify accelerometer data as walking or non-walking using either GPS or travel diary data within a large sample of adults under free-living conditions.

Methods—Participants wore an accelerometer and a GPS unit, and concurrently completed a travel diary for 7 consecutive days. Physical activity (PA) bouts were identified using accelerometry count sequences. PA bouts were then classified as walking or non-walking based on a decision-tree algorithm consisting of 7 classification scenarios. Algorithm reliability was examined relative to two independent analysts' classification of a 100-bout verification sample. The algorithm was then applied to the entire set of PA bouts.

Results—The 706 participants' (mean age 51 years, 62% female, 80% non-Hispanic white, 70% college graduate or higher) yielded 4,702 person-days of data and had a total of 13,971 PA bouts. The algorithm showed a mean agreement of 95% with the independent analysts. It classified physical activity into 8,170 (58.5 %) walking bouts and 5,337 (38.2%) non-walking bouts; 464 (3.3%) bouts were not classified for lack of GPS and diary data. Nearly 70% of the walking bouts and 68% of the non-walking bouts were classified using only the objective accelerometer and GPS data. Travel diary data helped classify 30% of all bouts with no GPS data. The mean duration of PA bouts classified as walking was 15.2 min (SD=12.9). On average, participants had 1.7 walking bouts and 25.4 total walking minutes per day.

Conclusions—GPS and travel diary information can be helpful in classifying most accelerometer-derived PA bouts into walking or non-walking behavior.

Keywords

Physical activity; walk trip; algorithm; classification

Corresponding author: Brian E. Saelens, Seattle Children's Research Institute, P.O. Box 5371, M/S: CW8-6, Seattle, WA 98145-5005; Phone: (206) 884-7800; Fax: (206) 884-7801; brian.saelens@seattlechildrens.org.

Introduction

Walking is the most popular means of being physically active and is beneficial to health. Public health researchers and practitioners have given increasing attention to policies to encourage people to walk more (19). The evaluation of interventions that specifically target walking requires accurate estimation of walking as a discrete form of physical activity. Self-report instruments can inaccurately estimate walking. The International Physical Activity Questionnaire (IPAQ) typically over-estimates walking frequency and duration (14, 27). Transportation surveys, generally designed to capture motorized trips, tend to underreport short trips, and walking trips in particular (28). Objective instruments, such as accelerometers and pedometers, provide more accurate assessments of physical activity (PA) intensity and duration, but are limited in identifying specific types of PA such as walking (1).

Recent studies have attempted to identify walking activity using time-based integration of accelerometry (i.e., identification of PA occurring at certain times) and GPS data (i.e., speeds consistent with walking during that time) (25, 30). The integrated data provide objective estimates of walking activities' duration, speed, and amount of PA gained. Moreover, GPS can specify the location of walking activities, thus defining the spatial and temporal context in which walking occurs (25). We found two studies that combined accelerometer and GPS data for an identification of walking: one used data from 10 adults walking under controlled conditions (30), and another used data from 42 girls' under free-living conditions (25). Clearly more investigation with larger samples is needed to further develop the integration of accelerometry and GPS to identify walking. Also, although GPS data provide objective location and speed information, they have limitations in data completeness. GPS units have missing data due to lost signals in urban canyons or inside buildings, signal drop-out, warm start/cold start and power interruptions (12, 20). A recent review of 24 studies of general PA (not focused on specific PA like walking) using GPS units in combination with accelerometers or travel/activity diaries reported that 17 studies had missing or unusable GPS data ranging from 2.5% to 92% of the observed time (13). Studies using only accelerometers and GPS might therefore yield biased results if the missing GPS data are not randomly distributed in time and space. Combining travel diary data with accelerometer and GPS data might help assess the distribution of missing GPS data and identify possible walking behavior when GPS data are missing.

The present study aimed to develop and test an automated algorithm for classifying accelerometer data as walking or non-walking using either GPS or travel diary data within a large adult population under free-living conditions. Accelerometer data were considered to be the complete catchment source for PA bouts.

Methods

Participants and data collection

Data came from phase 1 of the Travel Assessment and Community (TRAC) project. Between July 2008 and July 2009, 750 participants were recruited within the greater Seattle area. The spatial sampling frame covered 773 Census block groups with a uniform range of household income, race, home values, net residential density, housing type, availability of proximate neighborhood services, and levels of bus ridership (18). Participants were instructed to wear a hip-mounted accelerometer, to carry a GPS unit, and to record their travel in a diary for 7 consecutive days. The accelerometer (GT1M, Actigraph LLC, Fort Walton Beach, FL) was configured to acquire uniaxial activity counts in 30 s epochs, and XY coordinates, altitude, and instantaneous speed were measured with the DG-100 GPS data logger (GlobalSat, Taipei, Taiwan), also set to record at 30 s intervals. The travel diary

was modified from the National Household Travel Survey (NHTS) place-based format (8). Participants were instructed to record places visited, activities, arrival and departure times, and travel modes for all daily destinations. Participants were asked to re-wear both instruments and complete additional travel diaries up to two additional times until their data met the day-level initial data screening criteria (at least 5 days with any GPS data, 6 days with any data in the travel diary, and 6 days with accelerometry data   8 h after removing data with consecutive zeros for at least 20 min). In total, 52 participants were asked to re-wear; 3 were asked to re-wear twice; 5 continued to have only partial data after initial re-wearing but were included in the analyses; and one decided to drop from the study after re-wearing. Participants provided informed consent and the study was approved by the Seattle Children's Hospital IRB.

Data processing

Data were combined into a "LifeLog," which is an individual-level master table for all study participants with one record per 30 second epoch, spanning the assessment period and indexed by timestamp. Accelerometer data were directly joined to the LifeLog based on the same epoch times. Each GPS record was joined to match the LifeLog record closest in time (GPS records are not always recorded at consistent intervals, such as during re-acquisition of signal). The travel diary data were converted into a place table and a trip table, with trip records constructed by linking two temporally adjacent places. Each LifeLog record was then populated with the characteristics of its contemporaneous diary-based place or trip record. The complete LifeLog thus consisted of accelerometer counts, GPS XY coordinates and speed, and associated travel diary place or trip characteristics for each 30 s epoch, called LifeLog units.

This study only included data on complete days, defined by having at least one place record in the travel diary and an accelerometer wearing time of   8 h. Accelerometer periods of 20 min with continuous zeros were considered as non-wearing times (16, 25). A complete day may or may not have had GPS data. Accelerometer data on complete days were assumed to record all PA during wake time except for aquatic activities (i.e., swimming, showering). These data defined the temporal frame of PA for the classification of walking or non-walking (e.g., walking recorded in the travel diary but without accompanying accelerometer data were not part of the present analysis).

Definition of walking

Walking was defined *a priori* as non-mechanical (e.g., not cycling) and human-powered travel associated with sustained light or moderate intensity PA for at least 7 min with a 2-min tolerance of lower PA intensity (thus a walking bout must be at least 5 min in duration). Walking activity could be for utilitarian, recreational, or both purposes; it could start and end at the same place, but could not occur continuously or mostly continuously at the same location in space (i.e., walking on a treadmill was excluded). Walking was also differentiated from more vigorous or very slow movement. This operational definition of walking served to isolate "walking as travel in space" from other types of PA.

Physical activity bouts

Accelerometer data served to identify PA bouts, some of which were classified as walking. For present purposes, based on prior accelerometry evidence about walking (4, 17), PA bouts were defined as time intervals having accelerometer counts > 500 counts per 30 s epoch (cpe) for at least 7 min, allowing for up to 2 min of epochs below that threshold during the 7 min interval. Multiple time intervals with breaks   2 min were considered as one bout if the entire sequence of counts satisfied the count criteria. The count threshold of 500 cpe was chosen to capture light PA that might be associate with slow walking,

corresponding to the average of 500 cpe recorded by the GT1M for walking at an average speed of 3 kmh^{-1}(4, 11).

Developing the algorithm

An algorithm was developed and tested to sort PA bouts into walking and non-walking bouts (henceforth referred to as walking and non-walking). It was based on existing literature and evolved from a learning process using an algorithm-development sample of 100 randomly selected PA bouts. First, criteria were reviewed to determine GPS-derived walking speed, valid GPS temporal coverage, and if bouts occurred within a small spatial extent ("dwell bouts"). Second, seven scenarios, with sequential decisions, were developed to integrate the criteria into the process of classifying PA bouts as walking or non-walking. Third, the scenarios were structured into a decision-tree algorithm that could differentiate walking and non-walking identified from objective (accelerometry and GPS based) data from bouts identified by a mix of objective and subjective (accelerometry and diary-based) data. Finally, the algorithm was tested relative to independent raters using a verification sample of 100 randomly selected bouts, which were distinct from the algorithm development bouts sample.

To investigate the algorithm-development and the algorithm-verification samples, bout data were displayed using aerial photos and time series graphics (Figure 1) that provided information about their spatial and temporal context. The static aerial photos on which GPS points were overlaid were downloaded in February and March 2012 from Google Maps using RgoogleMaps 1.2.0 (15) for R 2.13.2 (22).

Criteria for GPS data

Selected GPS-derived walking speeds ranged between 2 kmh^{-1} and 6 kmh^{-1}. Two studies had defined walking trips as continuous movement without a break and having instantaneous GPS speeds within the wider ranges of 2 kmh^{-1} and 8 kmh^{-1} (6) and 1.6 kmh^{-1} and 9.6 kmh^{-1}, respectively (25). In the present study, we allowed walking bouts to have breaks 2 min within a 7-minute rolling window; and bout speed was defined as the median of available GPS speeds within that bout. The definition accounted for non-continuous movement often associated with walking (e.g., walking in an urban area where one might stop at an intersection) and for distinguishing walking from running or very slow movement. Median speed was selected because it is more robust than mean speed which could be biased by few GPS records from poor signals.

The review of GPS-derived tracks and speeds in the algorithm-development sample suggested that a GPS data coverage ratio of 20%, with at least 5 GPS records (2.5 min) provided reasonably sufficient spatial context information to classify a bout as a walking or non-walking. The entire set of PA bouts had a U-shaped distribution of GPS coverage ratio, which was similar to that of the algorithm-development sample. Nearly 84% of all PA bouts and 85% of the development sample bouts had GPS coverage ratios either below 20% or above 80%. So, in the algorithm, bouts with < 20% of GPS time coverage or fewer than 5 GPS observations were considered to have incomplete GPS data.

PA bouts occurring within a single location were considered as "dwells," which were considered non-walking by definition. Identifying a dwell bout was accomplished by (1) calculating the sum of distances from each point to all other points within the bout; (2) selecting points having sum distance below the 95th percentile of the sum distances of all points in the bout; (3) generating a minimum bounding circle fully containing the selected points; (4) finally obtaining the circle's radius. Bouts with radii 66 ft were considered as dwell bouts. Because some non-dwell bouts with few GPS observations were likely to have

radii 66 ft, dwell bouts were defined as having 10 GPS points. The cutoff of 66 ft was selected following an observational study that reported 95% of GPS points measured within of 66 ft of a fixed location using the same GPS model (33).

Classification scenarios in the algorithm

Based on the above criteria, we developed the following scenarios that select available data and classify PA bouts as walking or non-walking. The scenarios used first accelerometer data, then GPS data if available, and then travel diary data, based on the assumptions that accelerometers were more complete and accurate than GPS units, and that GPS data were more reliable and accurate than travel diary data. Thus data from the more reliable and accurate instruments were always used first. For example, a bout with complete GPS data and a median GPS speed of 0.25kmh^{-1} was classified as non-walking even though it might be a declared walking trip in the travel diary. In the algorithm-development sample, we found 15 cases of data conflicts between GPS and travel diary data.

Four scenarios using GPS and/or travel diary data served to define accelerometer-derived PA bouts as walking:

- *Walk1-GPS. GPS-derived non-dwell and walking speed*: Bouts with complete GPS data, with non-dwell GPS points, and with GPS speed medians within the acceptable walk speed range [2 kmh^{-1}, 6kmh^{-1}] (Figure 1 A).

- *Walk2-Diary. Overlap with diary-based walking trip*: Bouts with no or incomplete GPS data, but having any overlap in time with a walking trip recorded in the travel diary (Figure 1 B).

- *Walk3-Diary. Overlap with diary-based non-walking trip*: Bouts with no GPS data, but having any time overlap with a non-walking trip in the travel diary (Figure 1 C). Because non-walking trips (e.g., car, transit, and bike trips) are not PA bouts by definition, it was assumed that walking trips were typically not recorded in the diary when walking was not the primary travel mode, and that bouts adjacent to non-walking trips represented unreported walking trips (e.g., walking to and from transit stops).

- *Walk4-Diary. Overlap within a 10-minute tolerance of a diary-based trip*: Bouts overlapping a 10-minute time buffer around a trip reported in the travel diary (Figure 1 D). Reported times in the travel diary typically did not accurately match accelerometer time, likely due to recall errors or to the fact that times in the travel diary are often rounded to the nearest 5, 10 or 15 minutes (28). In the algorithm-development sample, the tolerance was not likely to result in false-positive errors based on the diary context. As in the Walk3-Diary scenario, bouts overlapping with non-walking modes were considered walking. Bicycling was assumed not to produce a PA bout because of the typically low counts obtained by uniaxial accelerometers while bicycling (24).

Three scenarios using GPS and travel diary data helped classify PA bouts as non-walking:

- *NonWalk1-ACC. Upper bound of accelerometer counts*: Bouts of vigorous PA with mean count 2,863 cpe (9) (no example shown). This scenario was validated with 3 PA bouts occurring while participants reported indoor exercising, with count means between 2,874 and 3,360 cpe.

- *NonWalk2-GPS. GPS-derived dwell and speed*: Dwell bouts or bouts with a GPS median speed outside the defined walk speed range. (Figure 1 E).

- *NonWalk3-Diary. Occurring within a diary-based place*: Bouts with no or incomplete GPS data, but with bout durations completely within a reported single place (e.g., home) and not within a 10-minute tolerance of a declared trip (Figure 1 F).

The characteristics of the algorithm-development PA bouts used to generate the criteria and scenarios are summarized in Table 1. Each column in the table shows PA bouts used for the corresponding scenario. In total, 52 PA bouts in the algorithm-development sample served for scenario development for walking, and 45 bouts for non-walking. The Walk1-GPS and Walk2-Diary scenarios identified likely unreported walking trips and time errors in the travel diary; 17 of 40 PA bouts in Walk1-GPS did not have an overlapping declared walking trip (Figure 1 A). Furthermore, 8 overlapped with a car or transit trip, supporting Walk3-Diary and Walk4-Diary assumptions that walking trips missing from the travel diary might be linked to other primary travel modes. NonWalk2-GPS scenario had 3 of 26 PA bouts overlapping with diary-based walk trips, but their GPS data showed that those bouts were dwell bouts or had low median speeds < 2 kmh^{-1}. This further illustrated inaccuracies in the recording of time in the travel diary.

The scenarios were organized by the level of confidence in the reliability and accuracy of the measures. Results of the scenarios using objective data—accelerometer only (NonWalk1-ACC) and the combination of accelerometer and GPS data (Walk1-GPS and NonWalk2-GPS)—should be more reliable than scenarios using a combination of accelerometer and travel diary data (Walk2-Diary, Walk3-Diary, Walk4-Diary and NonWalk3-Diary). Yet among this latter group, Walk2-Diary should be more reliable than the others because its bouts overlapped with declared walking trips, and the others were based on assumption of unreported walking trips and/or false time reporting of walking trips.

Algorithm

The classification algorithm was designed to combine the seven scenarios into one procedure with mutually exclusive bout classes, and to classify all of the accelerometer-derived PA bouts into walking and non-walking. A decision-tree model served to apply the scenarios sequentially, ranked by order of confidence. Figure 2 depicts the decision tree used in the algorithm.

Algorithm verification

The reliability of the algorithm results was examined by two trained analysts external to the research team. While blind to the algorithm results, each analyst independently classified the same algorithm-verification sample of 100 randomly selected PA bouts. The algorithm-verification sample was a separate sample from the algorithm-development sample. The analysts were provided the same instructions and visual materials (aerial photos with GPS points, time series graphics, and numerical bout summaries) (see Figure 1) as those previously used to establish the classification scenarios, but they were free to question the instructions. Analysts were asked to document their classification reasons.

Sensitivity analyses

Sensitivity analyses were conducted by changing data processing/algorithm parameter values. The accelerometer count threshold for vigorous activity was changed by ±5% (3,006 and 2,720 cpe); speed range shifted by ±0.5 kmh^{-1} ([2.5, 6.5 kmh^{-1}] and [1.5, 5.5 kmh^{-1}]); radius size for dwell bouts shifted by ±5% (69.3 and 62.7 ft); and time tolerance for trip overlap with diary-based trip shifted by ±5 min (5 min and 15 min). The percentage of PA bouts classified by each algorithm as walking versus non-walking was examined.

Results

Data processing and demographics

The final sample consisted of 13,528,634 LifeLog units, spanning 4,702 complete person-days for 706 participants (average 6.7 d per person). The sample had mean age of 50.9 yr (SD=13.3); 62% were females, 80% non-Hispanic whites, and 70% college graduate or higher. For 30% of the participants, the average annual household income was < $40,000; 50% had between $40,000 and $100,000, and 20% had > $100,000.

Wearing time and PA bout identification

Per day, participants had a mean accelerometer wearing time of 12.8 h (SD=1.6); an average of 11.9 h (SD=7.2) of GPS data; and 20.2 h (SD=5.2) of recorded time in the travel diary. In total, 13,971 PA bouts were identified through accelerometry. On average, PA bouts lasted 14.3 min (SD=12.2) and had 1,296 accelerometer counts per 30 s epoch (SD=556). On average, participants had 3.6 bouts and spent total 44.5 min in bouts per day. Approximately 30% of bouts had no GPS data. Approximately half of the bouts had 80% or more of their duration with GPS coverage.

Verification of algorithm results

Using the 100 PA bouts in the algorithm-verification sample, analyst A classified 64 as walking bouts, 34 as non-walking bouts, and 2 as unknown while analyst B classified 65 as walking bouts, 32 as non-walking bouts, and 3 as unknown. There were 92 agreements and 8 disagreements on individual bouts (Table 2), demonstrating good to excellent agreement between the two analysts (κ=.831, p-value <.0001). For the same sample of PA bouts, the algorithm classified 65 as walking bouts, 32 as non-walking bouts, and 3 as unknown. The algorithm had 92 agreements out of 100 with analyst A (κ=.831, p-value <.0001) and 98 agreements with analyst B (κ=.958, p-value <.0001).

Algorithm-based classification results

Based on all PA bouts, the algorithm classified 8,170 as walking and 5,337 as non-walking; 464 bouts were not classified for lack of GPS and travel diary data. The average walking bout duration was 15.2 min (SD=12.9). Walk1-GPS (69.8%) and Walk2-Diary scenarios (16.9%) classified the majority of the PA bouts that were identified as walking, compared to Walk3-Diary (7.7%) and Walk4-Diary (5.6%). Walking bouts of the Walk1-GPS and Walk2-Diary group were longer (mean +3.5 min) and had higher activity intensity (mean +241 cpe) than those of the Walk3-Diary and Walk4-Diary group (p<.0001). Non-walking was classified mainly with NonWalk1-ACC (3%) and NonWalk2-GPS (65.2%), which used only objective data. NonWalk1-ACC bouts had a mean duration of 32.4 min, 2 to 3 times longer than other non-walking bouts. Their mean accelerometer count was 3,518.9 cpe, equivalent to running at 7 kmh^{-1} (11). Non-walking bouts of the NonWalk2-GPS and NonWalk3-Diary group had significantly different mean intensity from all walking bouts (p<.0001). The mean of count means of the non-walking bouts in that group (NonWalk2-GPS and NonWalk3-Diary) was 429 cpe lower than the mean of all walking bouts' (Table 3). On average, at the person level, participants had 1.7 walking bouts (total 25.4 min) per day (Table 4). Out of 706 participants, 644 (91.2%) had at least one walking bout over the course of data collection. For this subsample of those with any identified walking, participants had 1.8 walking bouts (total 27.9 min) per day.

Sensitivity analyses

The base algorithm classified 58.5% of the sample bouts (n=13,971) as walking. Changing various algorithm parameters classified between 53.4 % (−5.1%) and 61.5% (+3.0%) of the

sample bouts as walking. When applied to the verification sample (n=100 bouts), only one algorithm (vigorous activity threshold +5%) presented an agreement rate as high as that of the base algorithm. Other adjusted algorithms decreased the agreement rates by up to 5% and 6% for analyst A and B, respectively.

Comparison with travel diary data

The travel diary had 8,201 declared walking trips; of these, 7,704, trips (93.9%) had complete data on reported travel duration; and 5,800 (70.7%) had a duration 5 min, which could be roughly comparable to algorithm-identified walking bouts (n=8,170) also defined to be 5 min. The number of algorithm-identified walking bouts was 2,370 more than the number of declared walking trips with a duration 5 min. There were significant differences (p<.0001) between the algorithm-identified walking and the travel diary reported walking trips in terms of frequency and duration at the person level (Table 4). The average frequency of walking bouts identified by the algorithm was 42% higher than travel diary walk trips with a duration 5 min (1.7 bouts versus 1.2 reported trips). There is a corresponding difference in estimated walking time between the algorithm and the diary.

Discussion

This study's large sample of adults observed over a week under free-living conditions provided a unique opportunity to estimate walking behavior using a combination of objective and participant-reported measures, which were compiled in a LifeLog integrating accelerometer, GPS, and travel diary data by common time stamps. Of the 13,971 PA bouts of at least 5 min derived from accelerometer data, the developed algorithm classified 58.5% as walking and 38.2% as non-walking (3.3% could not be classified). Overall, 57% of PA bout minutes were classified as walking (25.4 of 44.6 min) per person per day for the entire participant sample (n=706) and 59% (27.9 of 47.6 min) for the subsample of participants having any walking bouts (n=644). These percentages of walking relative to overall PA are consistent with the 60-65% estimate of daily walking relative to daily total PA time based on time use data collected among U.S. adults (32). However, the present study's estimate of relative contribution of walking to overall PA is markedly higher than the ~30% estimate observed among U.S. adults based on the self-reported International Physical Activity Questionnaire (IPAQ) (2). The low contribution of IPAQ-estimated walking relative to overall PA might be the result of overestimation of overall PA in the retrospective survey assessment (14).

In the current study, the estimated mean frequency of walking was 1.7 times per day, and the mean duration of walking was 15.2 min per episode. For the U.S. adult population based-on 2001-2009 NHTS data, the frequency of walking trips was 0.4-0.5 per day and the duration was 5.4-6.2 min per episode (21). Among U.S. adults, the 1998 Behavioral Risk Factor Surveillance System (BRFSS) data, which estimated leisure-time physical activity, indicated a small frequency of 0.4 walking episodes per day but a very long duration of 34.5 min per walking episode (23). Combining frequencies and per-episode durations, the total walking time per day was about 2.6 min in NHTS and 13.8 min in BRFSS, which is markedly lower than the present study's estimate of time spent in walking (25.4 min). It is not clear whether these differences reflect true behavioral differences between samples or potential measurement error.

The present study suggested that GPS was a necessary, but not sufficient, objective instrument to identify walking behavior among accelerometer-derived physical activity bouts. The Walk1-GPS scenario in the algorithm, using the combination of accelerometer and GPS data, captured a mean of 1.2 walking bouts (total 18.5 min) per person per day, missing the 0.5 walking bouts (total 6.9 min) obtained from Walk2-Diary, Walk3-Diary, and

Walk4-Diary, which used the combination of accelerometer and travel diary data. Also, Walk1-GPS still failed to capture the 0.3 walking bouts or 4.1 min of walking per person per day identified under the Walk2-Diary scenario (itself deemed more reliable than Walk3-Diary and Walk4-Diary). PA studies using existing GPS devices will have missing data due to poor satellite signal reception, some of which may be associated with urban form (e.g., urban canyons where signal is lost or location inaccurate) (33). This study used among the best and most feasible-to-wear GPS units (DG-100 GPS data logger) available at the time of study initiation, but these limitations might be minimized by newer GPS models, but could not likely be entirely eliminated.

Travel diaries have been criticized for self-report bias and missing trips, especially for short and/or incidental walking trips (e.g., walking to/from public transit) (28, 31). Studies using GPS data confirm that travel diaries substantially under-report trips, especially short trips (3). In the present study, the per capita per day frequency (1.2) and duration (21.6 min) of walking trips reported in travel diaries were higher than corresponding data (0.5, 5.4-6.2 min) from the 2001 and 2009 U.S. NHTS (21). However, travel diary data alone appears insufficient to adequately capture walking. In the algorithm-development bout sample, 43% of 40 walking bouts served for Walk1-GPS scenario were not reported in the travel diary (Figure 1 A). Across participants, travel diary data resulted in a 0.5 lower frequency (1.2 versus 1.7) and 3.8 fewer total minutes of walking per day (21.6 versus 25.4) than results from the algorithm that included GPS data (Table 4). This adds to the evidence that travel diaries underestimate walking behavior. Indeed, participants in the present study might have been especially diligent in recording activities in a complete and precise fashion in the travel diary because they knew they were being monitored with GPS, but travel diaries still resulted in under-reporting of walking.

Nearly 77% of the algorithm results were from scenarios with high confidence levels (Walk1-GPS, Walk2-Diary, NonWalk1-ACC, and NonWalk2-GPS). Scenarios with lower confidence levels, which yielded 20% of the results (Walk3-Diary, Walk4-Diary, and NonWalk3-GPS), are likely to have more misclassification errors. However, the sensitivity analyses showed that tolerance adjustment for diary-based trip overlap from 10 min to either 5 or 15 min only minimally changed the proportion of classified walking, from 58.5% to 57.4% (−1.1%) or to 59.4% (+0.9%), respectively. This suggests that the algorithm results were robust with different assumptions and use of the subjective travel diary data.

There was high agreement between the classifications of walking and non-walking bouts between the two independent analysts and between the analysts. The minimal disagreements that did exist between the analysts showed that GPS and travel diary data could be interpreted subjectively. For example, 3 bouts, which had very slow GPS median speeds (1.7-1.9 kmh^{-1}) but had GPS traces which suggested walking occurred, were classified differently by the two analysts. Also, place names were interpreted differently. For instance, a bout that lacked GPS data and was associated with "Bumbershoot," a local music/arts festival, was classified differently based on the analysts' local knowledge. Agreement between the algorithm and analysts was also high. Among all disagreements with the two analysts (total=9, overlap=1), 5 resulted from the deterministic criteria used (i.e., specific walking speed range and necessary percentage of GPS coverage), for which no gold standards exists. The remaining 4 disagreements stemmed from the inability of the current algorithm to use semantic information from the travel diary. As in the case of "Bumbershoot" above, the analysts used place names associated with bouts to infer bout characteristics. Employing semantic pattern recognition methods could improve the algorithm, but it would require an extensive database and business logic to link behaviors with place names.

Automated algorithms offer the benefit of reducing classification time and increasing transparency. Approximately 20 h were required per analyst for training, classification, and documentation for the verification sample of 100 bouts (0.7% of the total 13,971 bouts). Furthermore, it was not always clear why some analyst decisions were made. Computer-based deterministic algorithms, on the other hand, work very rapidly, and classify bouts without ambiguity.

This study had limitations. The participants are likely not representative of the general U.S. population and were not sampled to be representative of the study region (Seattle/King County, WA). This study was part of a larger study investigating the impact of light rail on travel behavior and physical activity, and participants had more transit access than the general U.S. population. Thus, the algorithm requires further testing for use in other populations. The results might also be sensitive to the instruments used (GT1M accelerometer and DG 100 GPS). However, the algorithm could be easily adjusted for different instruments by changing the criteria for PA bout identification and GPS data processing.

The algorithm and the results could potentially be biased due to missing accelerometer data, although the study sample had low accelerometer non-wearing time that was comparable to previous studies. The present study participants had a mean accelerometer non-wearing time of 11.2 h per day (wearing time: 12.8 h), and previous studies had between 9.8 and 11.5 h of daily non-wearing time (7, 10, 26, 29). It is possible that even after excluding legitimate non-wearing (e.g., sleeping, showering, etc.), some erroneous non-wearing time (i.e., forgot wearing at home) might exist. Some PA bouts could be completely missing (5), possibly affecting the algorithm and the results. External monitoring such as devise wearing logs or heart monitors could be used to more precisely isolate accelerometer non-wearing time, although themselves are not without limitations. In addition, the present algorithm is likely underestimating walking behavior, particularly within short bouts and if done within a relatively small spatial extent such as an indoor workplace. Other measures are needed to better capture this type of walking.

In summary, the integration of GPS and travel diary data with accelerometer data can help to identify walking behavior. It is likely, although requiring further testing, that this classification method is more accurate and complete than relying either solely on any one of these instruments alone or on retrospective report of walking behavior. Using GPS alone allowed the classification of nearly 66% of PA bouts as either walking or non-walking. The inclusion of travel diary data was helpful in those instances of PA bouts with insufficient GPS information. Approximately 30% of algorithm-derived walking bouts either lacked or had incomplete GPS data, and were identified through travel diary data. Travel diaries also provide useful contextual information on likely walk trips, including their associations with other modes of transport and their purpose (e.g., recreation, commuting), but as seen herein, clearly underestimate walking. Investigators and practitioners should capitalize on the strengths of using these three data streams in isolation and in combination, and align the aims of their inquiry to best measure those aspects of walking that are of interest.

Acknowledgments

References

1. Bassett DR Jr. Mahar MT, Rowe DA, Morrow JR Jr. Walking and measurement. Medicine and Science in Sports and Exercise. 2008; 40(7):529–36. [PubMed: 18379217]

2. Bauman A, Bull F, Chey T, Craig CL, Ainsworth BE, Sallis JF, Bowles HR, Hagstromer M, Sjostrom M, Pratt M. The International Prevalence Study on Physical Activity: results from 20 countries. International Journal of Behavioral Nutrition and Physical Activity. 2009; 6:21. [PubMed: 19335883]

3. Bohte W, Maat K. Deriving and validating trip purposes and travel modes for multi-day GPS-based travel surveys: A large-scale application in the Netherlands. Transportation Research Part C: Emerging Technologies. 2009; 17(3):285–97.

4. Brage S, Wedderkopp N, Franks PW, Andersen LB, Froberg K. Reexamination of validity and reliability of the CSA monitor in walking and running. Medicine and Science in Sports and Exercise. 2003; 35(8):1447–54. [PubMed: 12900703]

5. Catellier DJ, Hannan PJ, Murray DM, Addy CL, Conway TL, Yang S, Rice JC. Imputation of missing data when measuring physical activity by accelerometry. Medicine and Science in Sports and Exercise. 2005; 37(11):555–62.

6. Cho GH, Rodríguez DA, Evenson KR. Identifying walking trips using GPS data. Medicine and Science in Sports and Exercise. 2011; 43(2):365. [PubMed: 20581721]

7. Cradock AL, Wiecha JL, Peterson KE, Sobol AM, Colditz GA, Gortmaker SL. Youth recall and TriTrac accelerometer estimates of physical activity levels. Medicine and Science in Sports and Exercise. 2004; 36(3):525. [PubMed: 15076797]

8. Federal Highway Administration. 2001 National Household Travel Survey: user's guide. US Department of Transportation; Washington, DC: 2004. Available from: US Department of Transportation

9. Freedson PS, Melanson E, Sirard J. Calibration of the computer science and applications, Inc. accelerometer. Medicine and Science in Sports and Exercise. 1998; 30(5):777–81. [PubMed: 9588623]

10. Hagströmer M, Oja P, Sjöström M. Physical activity and inactivity in an adult population assessed by accelerometry. Medicine and Science in Sports and Exercise. 2007; 39(9):1502–8. [PubMed: 17805081]

11. John D, Tyo B, Bassett DR. Comparison of four ActiGraph accelerometers during walking and running. Medicine and Science in Sports and Exercise. 2010; 42(2):368–74. [PubMed: 19927022]

12. Kerr J, Duncan S, Schipperjin J. Using Global Positioning Systems in health research: a practical approach to data collection and processing. American Journal of Preventive Medicine. 2011; 41(5):532–40. [PubMed: 22011426]

13. Krenn PJ, Titze S, Oja P, Jones A, Ogilvie D. Use of Global Positioning Systems to study physical activity and the environment: A systematic review. American Journal of Preventive Medicine. 2011; 41(5):508–15. [PubMed: 22011423]

14. Lee PH, Macfarlane DJ, Lam TH, Stewart SM. Validity of the international physical activity questionnaire short form (IPAQ-SF): A systematic review. International Journal of Behavioral Nutrition and Physical Activity. 2011; 8:115. [PubMed: 22018588]

15. Loecher, M.; Berlin School of Economics and Law. RgoogleMaps: Overlays on Google map tiles in R. R package version 1.2.0. 2012. http://CRAN.R-project.org/package=RgoogleMaps

16. Mâsse LC, Fuemmeler BF, Anderson CB, Matthews CE, Trost SG, Catellier DJ, Treuth M. Accelerometer data reduction: a comparison of four reduction algorithms on select outcome variables. Medicine and Science in Sports and Exercise. 2005; 37(11):544–54. [PubMed: 15809550]

17. Matthews CE. Calibration of accelerometer output for adults. Medicine and Science in Sports and Exercise. 2005; 37(11):512–22.

18. Moudon, AV.; Saelens, BE.; Rutherford, S.; Hallenbeck, M. A report on participant sampling and recruitment for travel and physical activity data collection. Transportation Northwest; Seattle, Wash.: 2009. Available from: Transportation Northwest

19. Ogilvie D, Foster CE, Rothnie H, Cavill N, Hamilton V, Fitzsimons CF, Mutrie N. Interventions to promote walking: systematic review. British Medical Journal. 2007; 334(7605):1204. [PubMed: 17540909]

20. Oliver M, Badland H, Mavoa S, Duncan MJ, Duncan J. Combining GPS, GIS, and accelerometry: methodological issues in the assessment of location and intensity of travel behaviors. Journal of Physical Activity and Health. 2010; 7(1):102–8. [PubMed: 20231761]

21. Pucher J, Buehler R, Merom D, Bauman A. Walking and cycling in the United States, 2001-2009: Evidence from the National Household Travel Surveys. American Journal of Public Health. 2011; 101(SUPPL. 1):S310–S7. [PubMed: 21551387]

22. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing; Vienna, Austria: 2012.

23. Rafferty AP, Reeves MJ, McGee HB, Pivarnik JM. Physical activity patterns among walkers and compliance with public health recommendations. Medicine and Science in Sports and Exercise. 2002; 34(8):1255–61. [PubMed: 12165679]

24. Reilly JJ, Kelly LA, Montgomery C, Jackson DM, Slater C, Grant S, Paton JY. Validation of Actigraph accelerometer estimates of total energy expenditure in young children. International Journal of Pediatric Obesity. 2006; 1(3):161–7. [PubMed: 17899634]

25. Rodriguez D, Cho G, Elder J, Conway T, Evenson K, Ghosh-Dastidar B. Identifying walking trips from GPS and accelerometer data in adolescent females. Journal of Physical Activity and Health. 2012; 9(3):421–31. [PubMed: 21934163]

26. Rodríguez DA, Cho GH, Evenson KR, Conway TL, Cohen D, Ghosh-Dastidar B, Pickrel JL, Veblen-Mortenson S, Lytle LA. Out and about: Association of the built environment with physical activity behaviors of adolescent females. Health and Place. 2012; 18(1):55–62. [PubMed: 21945085]

27. Rzewnicki R, Auweele YV, De Bourdeaudhuij I. Addressing overreporting on the International Physical Activity Questionnaire (IPAQ) telephone survey with a population sample. Public Health Nutrition. 2003; 6(3):299–306. [PubMed: 12740079]

28. Stopher PR, Greaves SP. Household travel surveys: Where are we going? Transportation Research Part A: Policy and Practice. 2007; 41(5):367–81.

29. Troiano RP, Berrigan D, Dodd KW, Mâsse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. Medicine and Science in Sports and Exercise. 2008; 40(1):181–8. [PubMed: 18091006]

30. Troped PJ, Oliveira MS, Matthews CE, Cromley EK, Melly SJ, Craig BA. Prediction of activity mode with global positioning system and accelerometer data. Medicine and Science in Sports and Exercise. 2008; 40(5):972. [PubMed: 18408598]

31. Tudor-Locke C, Bittman M, Merom D, Bauman A. Patterns of walking for transport and exercise: a novel application of time use data. International Journal of Behavioral Nutrition and Physical Activity. 2006; 7(2):55–64.

32. Tudor-Locke C, van der Ploeg HP, Bowles HR, Bittman M, Fisher K, Merom D, Gershuny J, Bauman A, Egerton M. Walking behaviours from the 1965-2003 American Heritage Time Use Study (AHTUS). The International Journal of Behavioral Nutrition and Physical Activity. 2007; 4(45)

33. Wu J, Jiang C, Liu Z, Houston D, Jaimes G, McConnell R. Performances of different global positioning system devices for time-location tracking in air pollution epidemiological studies. Environmental Health Insights. 2010; 4:93–108. [PubMed: 21151593]
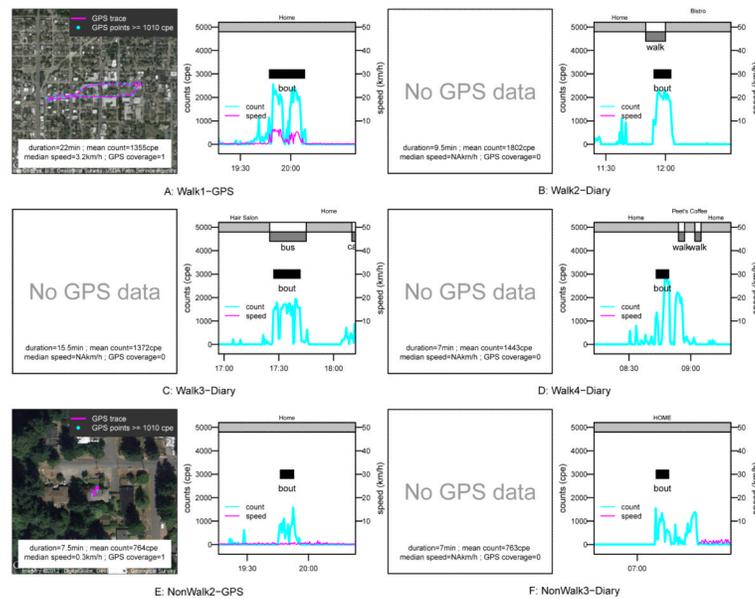
Figure 1.
Each panel (A-F) has an aerial photo with GPS points (left) and a time series graphic (right). Basic numerical summaries are presented in the bottom of the photo. If bouts have no GPS data at all, a blank photo is shown. Above the graphic is diary-based place and trip information. Accelerometer counts (cyan) and GPS speeds (magenta) are plotted in the graphic. A black box shows bout duration, derived from the accelerometer, in the middle of the graphic. [A: Walk1-GPS example] The PA bout was non-dwell with GPS median speed of 3.2 kmh^{-1} and therefore was classified as a walking bout, even though no walking trip was reported in the diary when the bout occurred. [B: Walk2-Diary example] The PA bout had no GPS data, but had time overlap with a diary-based walking trip, so was classified as a walking bout. [C: Walk3-Diary example] The PA bout had no GPS data, no walking trip was recorded in the diary near the bout, but since it had time overlap with a diary-based bus trip it was classified as a walking bout assumed to be associated with the bus trip. [D: Walk4-Diary example] The PA bout had no GPS data and no overlap with a diary-based trip, but had a diary-based walking trip 4.5 min after it and was thus classified as a walking bout assumed to be associated with the walking trip but with diary time errors. [E: NonWalk2-GPS example] The PA bout was a dwell bout (the GPS point cluster circle radius = 12.8ft) and thus was not a walking bout. [F: NonWalk3-Diary example] The PA bout had no GPS data, but occurred within a diary-based place and had no other diary-based trips close in time.
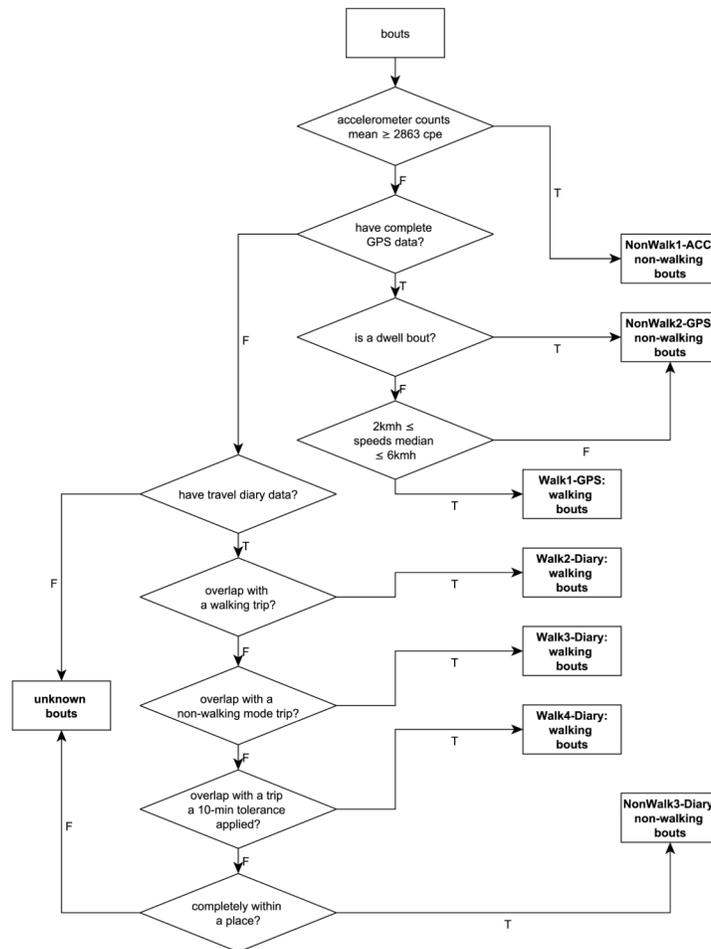
Figure 2.
The decision tree algorithm shows sequential application of the seven scenarios to classify PA bouts as walking or non-walking.

Table 1

Characteristics of PA bouts (n=97) used for scenario development (3 bouts with incomplete GPS and no travel data were excluded.)

Scenario	Walk1-GPS (n=40)	Walk2-Diary (n=5)	Walk3-Diary (n=3)	Walk4-Diary (n=4)	NonWalk1-ACC (n=3)	NonWalk2-GPS (n=26)	NonWalk3-Diary (n=16)
Accelerometer counts mean range (counts per 30 s epoch)	[663.4, 2464.6]	[881.6, 1802.3]	[805.4, 1832.5]	[735, 1442.6]	[2874.2, 3360.3]	[548.1, 2198.4]	[701.6, 2158]
GPS coverage ratio range	[0.25, 1]	[0, 0.14]	[0, 0]	[0, 0]	[0.05, 0.89]	[0.62, 1]	[0, 0.19]
GPS median speed range (kmh^{-1})	[2.15, 5.3]	[2.55, 2.55]	[NA, NA]	[NA, NA]	[0.5, 4.64]	[0.1, 3.15]	[NA, NA]
Number of bouts overlapping with travel diary reported:							
• Walking trip	23	5	0	1*	0	3	0
• Bike trip	0	0	0	1*	0	0	0
• Car trip	6	0	2	1*	0	2	0
• Transit trip	2	0	1	0	0	0	0
• Other/unknown mode trip	0	0	0	1*	0	1	0

*
Bouts in Walk4-Diary overlapped with a trip with a 10-minute tolerance applied around diary-based trips.

Table 2

Comparison of walking versus non-walking classification among the algorithm and independent analysts

		Analyst A				Algorithm			
		Walking	Non-walking	Unknown	Total	Walking	Non-walking	Unknown	Total
Analyst B	Walking	**61**	4	0	65	**64**	1	0	65
	Non-walking	3	**29**	0	32	1	**31**	0	32
	Unknown	0	1	**2**	3	0	0	**3**	3
	Total	64	34	2	100	65	32	3	100
	Agreement	(61+29+2)/100=**0.92** Cohen's Kappa=0.831, p <.0001				(64+31+3)/100=**0.98** Cohen's Kappa=0.958, p <.0001			
Algorithm	Walking	**61**	4	0	65				
	Non-walking	3	**29**	0	32				
	Unknown	0	1	**2**	3				
	Total	64	34	2	100				
	Agreement	(61+29+2)/100=**0.92** Cohen's Kappa=0.831, p <.0001							

Table 3

Bout classification results from the algorithm

Scenario	Number	Overall %	Within group %	Duration (min) Mean	(SD)	Mean count (counts per 30 s epoch) Mean	(SD)
Walking bouts							
• Walk1-GPS	5,704	40.8	69.8	15.7	(13.3)	1,467.6	(484.0)
• Walk2-Diary	1,378	9.9	16.9	15.2	(11.9)	1,491.5	(457.0)
• Walk3-Diary	632	4.5	7.7	11.7	(10.2)	1,257.9	(448.5)
• Walk4-Diary	456	3.3	5.6	12.8	(12.9)	1,194.5	(474.3)
All walking bouts	**8,170**	**58.5**	**100.0**	**15.2**	**(12.9)**	**1,440.2**	**(483.5)**
Non-walking bouts							
• NonWalk1-ACC	159	1.1	3.0	32.4	(18.3)	3,518.9	(748.0)
• NonWalk2-GPS	3,479	24.9	65.2	12.7	(10.6)	1,009.4	(406.8)
• NonWalk3-Diary	1,699	12.2	31.8	11.9	(9.7)	1,013.2	(389.8)
All non-walking bouts	**5,337**	**38.2**	**100.0**	**13.1**	**(11.2)**	**1,085.4**	**(595.5)**
Unknown bouts	464	3.3	-	12.2	(9.8)	1,171.5	(475.0)
All bouts	**13,971**	**100.0**	-	**14.3**	**(12.2)**	**1,295.7**	**(556.1)**

Table 4

Comparison between algorithm-identified walking bouts and travel diary-based walking trips at per person per day level

	Frequency	Total minutes
Algorithm-identified walking bouts		
• Walk1-GPS	1.2	18.5
• Walk2-Diary	0.3	4.1
• Walk3-Diary	0.1	1.6
• Walk4-Diary	0.1	1.2
All walking bouts	**1.7**	**25.4**
Travel diary-based walking trips　5 minutes	**1.2** [*]	**21.6** [**]

Paired T-test between algorithm-identified walking and diary-based walking:

[*] t = 13.67, df = 705, p < .0001

[**] t = 5.39, df = 705, p < .0001