Objectively Measured Walking: Temporal, Spatial, and Built Environmental Characteristics

Bumjoon Kang

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2013

Reading Committee:

Anne Vernez Moudon, Chair

Brian Ernest Saelens

Qing Shen

Program Authorized to Offer Degree:

Urban Design and Planning Group

University of Washington

Abstract

Objectively Measured Walking: Temporal, Spatial, and Built Environmental Characteristics

Bumjoon Kang

Chair of the Supervisory Committee:
Professor Anne Vernez Moudon
Urban Design and Planning

Walking and environmental correlates have been increasingly investigated in the field of public health and transportation. Recent development in real-time activity and location tracking technology provides a unique opportunity to explore the relationship between walking and environments. This dissertation developed a methodology to measure PA and to classify types objectively (walking versus non-walking and utilitarian versus recreational) using such new-technology data (accelerometer and GPS). Based on the objective measures of walking and PA, the inter-relationships among PA components were analyzed. Findings showed that increased walking contributed to total PA without reducing other types of PA. This suggested that no substitutive behaviors existed between walking and other types of PA, and that more walking would correspond to a net increase in overall PA. In addition, differences between utilitarian and recreational walking were investigated. Utilitarian and recreational walking can be determined more objectively by travel diary and GPS data. It was shown that walking is not a one-kind behavior, and utilitarian walking should be distinguished from recreational walking. Environmental variables strongly predicted utilitarian walking but not recreational walking. Correlates of utilitarian walking seem to also apply to overall PA.

TABLE OF CONTENTS

Chapter 1 Introduction 1

Background 1

References 3

Chapter 2 Walking Objectively Measured 5

Chapter Abstract 6

Introduction 7

Methods 8

Participants and data collection 8

Data processing 9

Data management and visualization 10

Definition of walking 14

Physical activity bouts 14

Developing the algorithm 14

Criteria for GPS data 17

Criteria for trip overlap 19

Classification scenarios in the algorithm 20

Algorithm 23

Algorithm verification 23

Sensitivity analyses 23

Results 27

Data processing and demographics 27

Wearing time and PA bout identification 27

Verification of algorithm results 28

Algorithm-based classification results 29

Sensitivity analyses 30

Comparison with travel diary data 32

Discussion 32
Conclusion 38
References 39

Chapter 3 The Contribution of Walking to Total Physical Activity 42

Chapter Abstract 42
Introduction 43
Methods 45
Data collection 45
Data processing 46
Data analysis 50
Results 51
Data collection and processing 51
Data analysis 54
Discussion 61
Conclusion 64
References 65

Chapter 4 Utilitarian and Recreational Walking 69

Chapter Abstract 69
Introduction 71
Methods 73
Participants and data collection 73
Data processing 73
Classification of utilitarian and recreational walking 74
PA bout measurements 75
PA perceptions 76
Subjective neighborhood perception 76
Objective built environments 77
Analysis 79

Results 80

Participants..80

Classification...82

Characteristics of utilitarian versus recreational walking at the bout level 82

Person-level distributions of PA measurements 83

Analysis..84

Discussion 89

Conclusion 93

References..94

Chapter 5 Dissertation Conclusion 98

LIST OF FIGURES

Figure 2-1. Workflow to create LifeLogs....11
Figure 2-2. Map of GPS points in a 7-day LifeLog....12
Figure 2-3. Time-series graphics of a 7-day LifeLog....13
Figure 2-4. Graphical examples of the classification....16
Figure 2-5. GPS statistics of PA bouts....18
Figure 2-6. Example of a dwell bout....19
Figure 2-7. PA bouts and declared walking bouts....20
Figure 2-8. The decision tree algorithm....26

Figure 3-1. Geocoded home points of the sample (n=657)....52
Figure 3-2. Scatter plot of WPD versus OPD for 657 participants....56

Figure 4-1. Walking bouts classified as recreational....75
Figure 4-2. Recreational walking bout determined by GPS data....75
Figure 4-3. Classification of bouts....82

LIST OF TABLES

Table 2-1. Characteristics of PA bouts (n=97) 25
Table 2-2. Demographic characteristics of the sample (n=706) 27
Table 2-3. Comparison among the algorithm and independent analysts 28
Table 2-4. Bout classification results from the algorithm....................30
Table 2-5. Comparison between algorithm-identified walking bouts and travel diary-based walking trips at per person per day level 31
Table 2-6. Sensitivity analysis summary 31

Table 3-1. PA perception variables....................48
Table 3-2. Demographic characteristics of the sample....................53
Table 3-3. Distributions of WPD and OPD at the person-level....................54
Table 3-4. Full sample model results at the person-level 55
Table 3-5.Subgroup model results at the person-level....................58
Table 3-6. Full sample model results at the day-level 59
Table 3-7. Subgroup model results at the day-level 60

Table 4-1. Neighborhood perception variables....................77
Table 4-2. Built environment variables, measures, and data sources 79
Table 4-3. Demographic, psycho-social, neighborhood perception and environmental measures....................81
Table 4-4. Bout level comparison....................83
Table 4-5. Person level distribution of PA measurements....................84
Table 4-6. Pearson’s correlation among BE variables....................85
Table 4-7. Results of non-environmental models (Model NE)....................87
Table 4 8. Results of environmental models (Model E) 88
Table 4 9. Comparison of the analysis with previous studies....................91

ACKNOWLEDGEMENTS

This dissertation was made possible by the wonderful mentorship of Professor Anne Vernez Moudon and the rest of my Ph.D. Reading Committee: Professors Brian E. Saelens, and Qing Shen. Technical support was provided by Professor Phillip M. Hurvitz.

For both technical and personal support, I wish to thank Ruizhu Huang, Phillip M. Hurvitz, Jason Y. Scully, Amir J. Sheikh, Orion Stewart, and Jared Ulmer in the Urban Form Lab. For data collection and technical support, I wish to thank Lucas Reichley and Chuan Zhou in the Seattle Children's Research Institute.

The most special thanks go to my family for their sympathetic support. In particular, my wife Seung-un always provided necessary intellectual, spiritual, and emotional support and encouragement during the time of my Ph.D. study.

Financial support was generously provided by the University of Washington Department of Urban Design and Planning, the National Institute for International Education of South Korea Study Abroad Scholarship Program, the National Institutes of Health (Grants R01HL091881 and R01DK076608), the Washington State Department of Transportation, and Centers for Disease Control and Prevention (Grant U48DP000050-04).

DEDICATION

This work is dedicated to my family.

Chapter 1 Introduction

Background

Physical inactivity is estimated to cause 3.2 million deaths globally (6%) and it was identified as the fourth leading mortality risk factor by World Health Organization.[1, 2] It increases the risk of cardiovascular diseases, diabetes, colon and breast cancer, and depression. Yet physical inactivity is becoming increasingly prevalent in most industrialized countries.[3] The 2005 Behavioral Risk Factor Surveillance System survey (BRFSS) data showed that fewer than 50% adult Americans achieved adequate levels of physical activity (PA).[4] The 2003–2004 National Health and Nutritional Examination Survey (NHANES) indicated that less than 5% of the US population achieved the recommended PA levels using objective PA measurement data.[5] Recommended PA levels range from a daily average of 30 min of moderate-intensity PA for at least 5 days per week or 20 min of vigorous-intensity PA for at least 3 days a week.[6, 7]

Walking is the most common and popular form of moderate PA.[8] Walking is known to be an appropriate form of exercise providing cardiovascular capacity and endurance, bone strength, and psychological and social benefits.[9] It is accessible to most adults, does not require special equipment or skills, and is conveniently accommodated in daily routines. For these reasons, and because walking also brings many transportation and environmental benefits [10], health researchers and practitioners have increasingly recognized that promoting walking is an important public health goal.[11, 12]

To date, most studies have used subjective data for walking and PA measures. These measures have been criticized for their inaccurate outcomes.[13, 14] For example, a common walking measure is calculated from the multiplication of average duration and average frequency of self-reported walking trips per week. Such estimates are easily biased by participants' recall and social desirability biases. Recent technological advances promise more accurate and objective measurement of PA thanks to a new generation of activity and location measurement devices. The present research used PA data derived from such devices. Specifically, it benefits from objective activity data obtained from accelerometers and location data from GPS units. These data are integrated in a continuous spatio-temporal framework that is augmented by travel logs containing reported information on places visited and trips made between places.

This dissertation has three research objectives:

1. Develop a methodology measuring PA objectively and classifying the measured PA by types;
2. Examine the extent to which walking contributes to overall PA and determine whether walking has positive or negative effects on other types of PA; and
3. Define more objectively the differences between utilitarian and recreational walking and identify their respective environmental correlates

Chapters 2, 3, and 4 discuss the research objectives 1, 2, and 3 respectively.

References

1. World Health Organization. Global recommendations on physical activity for health. World Health Organization; 2010.
2. World Health Organization. Global health risks : mortality and burden of disease attributable to selected major risks. Geneva, Switzerland: World Health Organization; 2009.
3. Pate RR, Pratt M, Blair SN, Haskell WL, Macera CA, Bouchard C, et al. Physical activity and public health. Journal of the American Medical Association 1995;273(5):402.
4. Kruger J, Kohl Iii HW, Miles IJ. Prevalence of Regular Physical Activity Among Adults--United States, 2001 and 2005. JAMA: Journal of the American Medical Association 2008;299(1).
5. 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):181-188.
6. Haskell WL, Lee IM, Pate RR, Powell KE, Blair SN. Physical activity and public health. Updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. Medicine and Science in Sports and Exercise 2007;39(8):1423-34.
7. Sallis JF, Frank LD, Saelens BE, Kraft MK. Active transportation and physical activity: Opportunities for collaboration on transportation and public health research. Transportation Research Part A: Policy and Practice 2004;38(4):249-268.
8. Simpson ME, Serdula M, Galuska DA, Gillespie C, Donehoo R, Macera C, et al. Walking trends among U.S. adults: the Behavioral Risk Factor Surveillance System, 1987-2000. American Journal of Preventive Medicine 2003;25(2):95-100.
9. Morris J, Hardman A. Walking to health. Sports Medicine 1997;23(5):306.
10. Chapman L. Transport and climate change: a review. Journal of Transport Geography 2007;15(5):354-367.
11. Ogilvie D, Foster CE, Rothnie H, Cavill N, Hamilton V, Fitzsimons CF, et al. Interventions to promote walking: systematic review. British Medical Journal 2007;334(7605):1204.
12. United States. Dept. of H, Human S. Healthy people 2010. Washington, DC: U.S. Dept. of Health and Human Services : For sale by the U.S. G.P.O., Supt. of Docs.; 2000.

13. 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(1).
14. 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.

Chapter 2 Walking Objectively Measured

NOTE:

This chapter was included in an article, Walking Objectively Measured: Classifying Accelerometer Data with GPS and Travel Diaries in *Medicine and Science in Sports and Exercise* (to be published in July 2013). As the first author of the article, I led research design, analysis, and writing. Co-authors are 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

[2] Seattle Children's Research Institute

[3] Department of Pediatrics, University of Washington

Chapter 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.

Key words: Physical activity; walk trip; algorithm; classification

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.[1] 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.[2, 3] Transportation surveys, generally designed to capture motorized trips, tend to underreport short trips, and walking trips in particular.[4] 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.[5]

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).[6, 7] 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.[6] 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.[7], and another used data from 42 girls' under free-living conditions.[6] 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.[8, 9] 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.[10] 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.[11] 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.[12] 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.[6, 13] 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).

Data management and visualization

The accelerometers (in .DAT format) and GPS data (in .CSV or GSD format) were downloaded using the instrument drivers. Completed paper travel diaries were digitized into a Microsoft Access database file by research specialists in the Seattle Children's Research Foundation. The Urban Form Lab. at the University of Washington received these raw data and processed using PostgreSQL 9.1.1 (PostgreSQL Global Development Group, 2012), PostGIS 2.0.1 (Refractions Research, 2012), and R 2.15.1 (R Core Team, 2012). For each participant, accelerometer, GPS, and travel diary data were temporally joined and merged into a master table, called "LifeLog" (Figure 2-1). One LifeLog unit is a record in LifeLog tables, which contains information of 30 s epoch.

A typical one day's LifeLog consists of 2,880 LifeLog units (24 h/min × 60 min/h × 2 LifeLog units/min). The entire dataset theoretically has 5,250 days (750 participants × 7 d/participant), yielding more than 15 million LifeLog unit observations (2,880 × 5,250). Therefore, it is very difficult to review data individually. Data were visualized to check errors and to provide temporal and spatial contexts of PA assessments. GPS maps (Figure 2-2) and time-series graphics (Figure 2-3) and were developed.

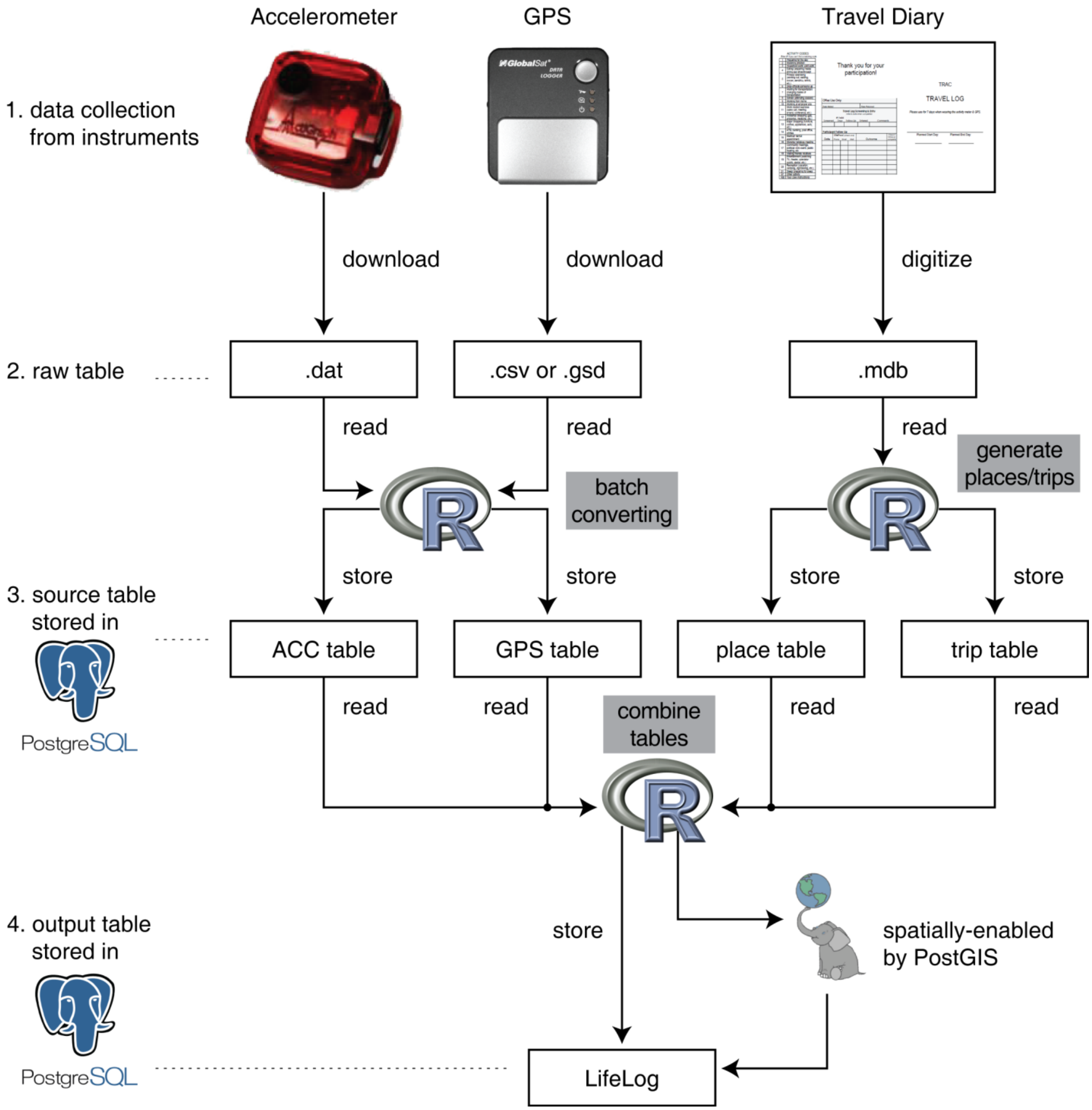

Figure 2-1. Workflow to create LifeLogs

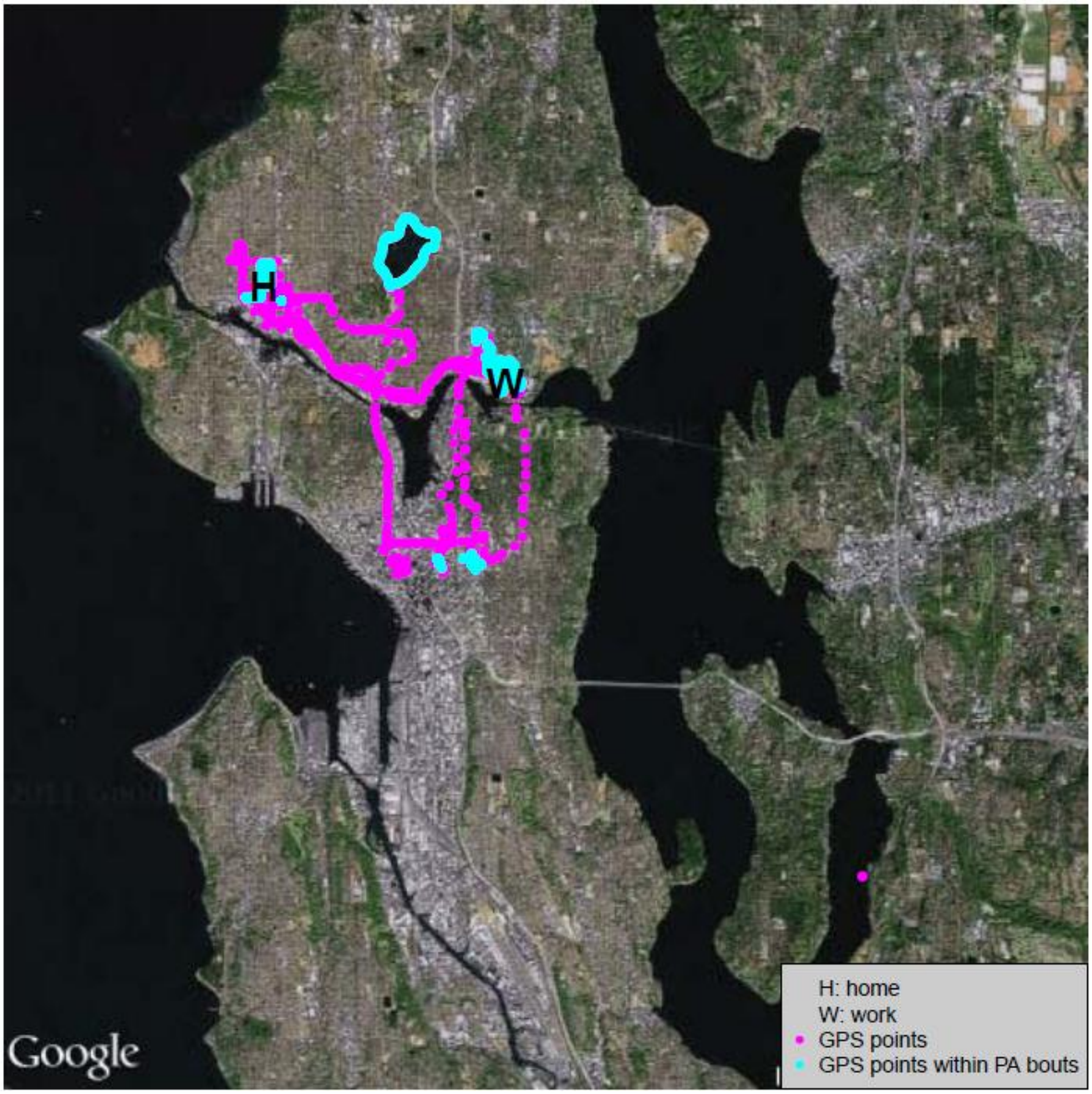

Figure 2-2. Map of GPS points in a 7-day LifeLog (the same participant's data shown in Figure 2-3)

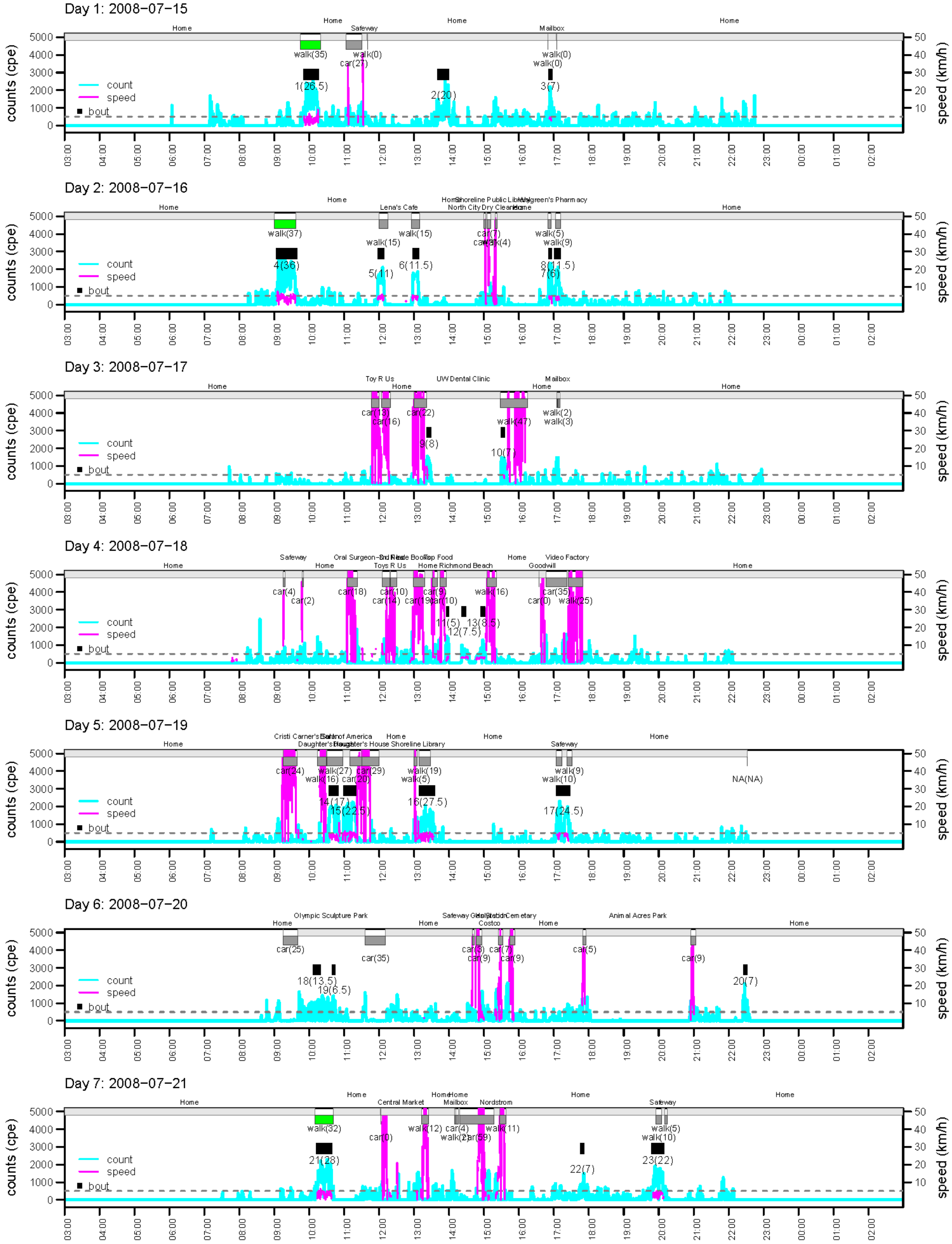

Figure 2-3. Time-series graphics of a 7-day LifeLog (the same participant's data shown in Figure 2-2).

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.[14, 15], 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}[14, 16]

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 2-4) 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[17] for R 2.13.2.[18]

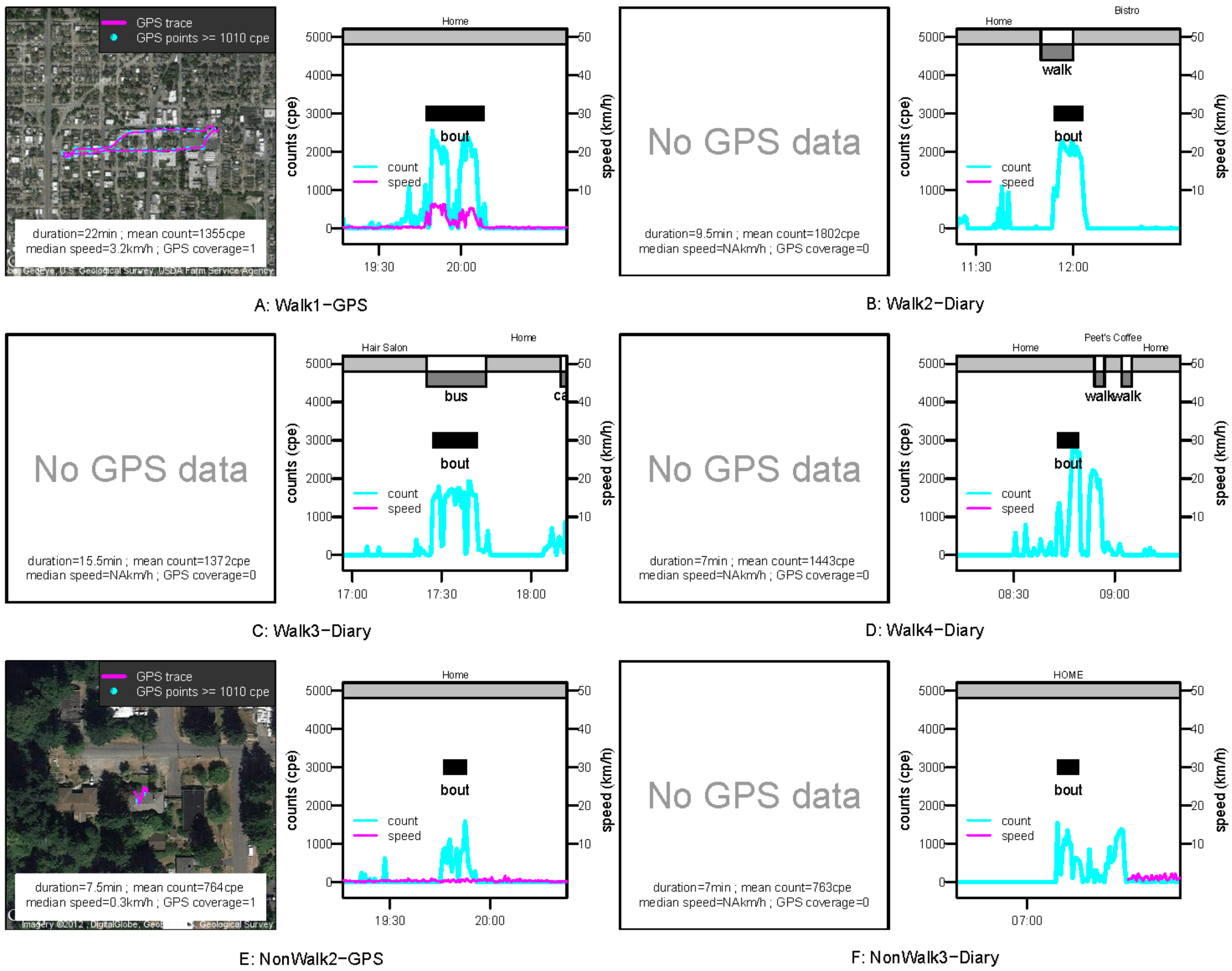

Figure 2-4. Each panel (A-F) has an aerial photo with GPS points (left) and a time series graphic (right). Basic numerical summaries are presented in the bottom of the photo. If bouts have no GPS data at all, a blank photo is shown. Above the graphic is diary-based place and trip information. Accelerometer 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.

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}.[19] and 1.6 kmh^{-1} and 9.6 kmh^{-1}, respectively.[6] 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 (Figure 2-5). 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.

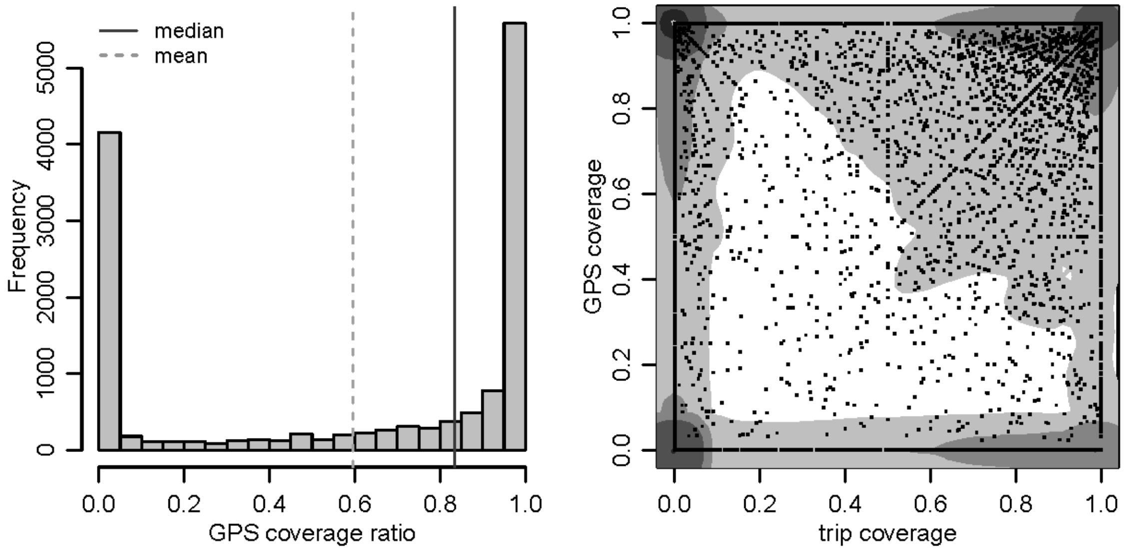

Figure 2-5. (Left) GPS coverage ratio distribution of PA bouts; (Right) Joint distribution of GPS coverage % and overlap % with declared trip of PA bouts

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.[20] Figure 2-6 shows an example of an identified dwell bout. One outlier GPS point was excluded to determine whether it is a dwell bout.

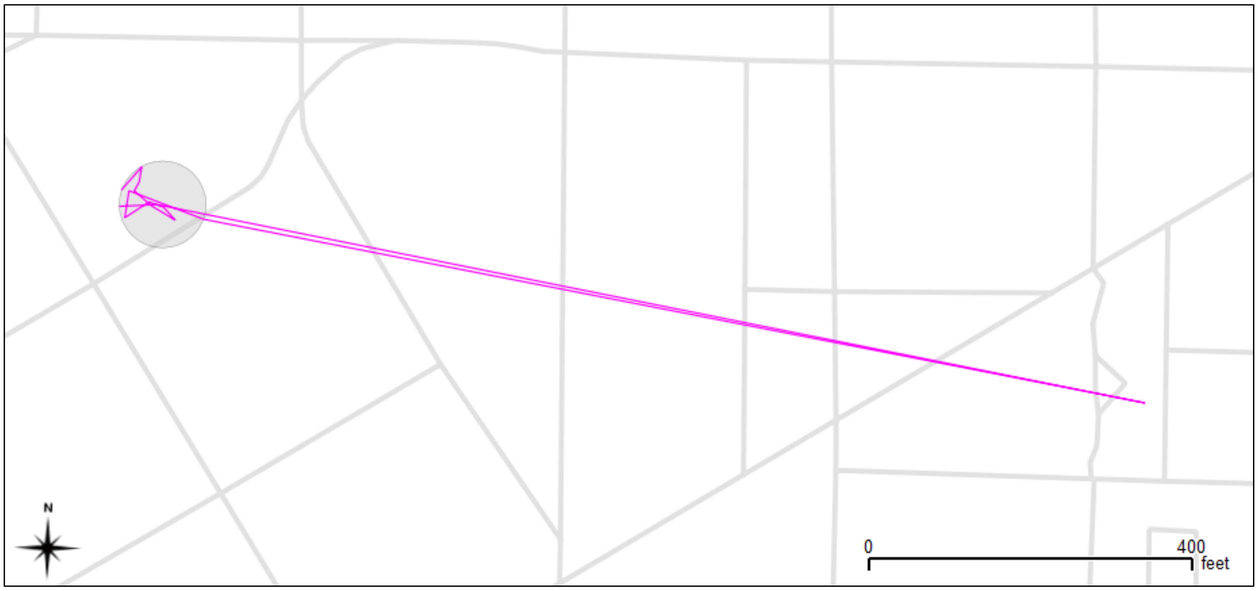

Figure 2-6. Example of a dwell bout. One potential error GPS point (located in East) was excluded to create the minimum bounding circle (grey) to determine whether this set of GPS points were recorded within a dwell bout

Criteria for trip overlap

PA bouts overlapping with declared walking trips from travel diary were likely to be walking bouts. However, participants might record trips, particularly short walking trips with time errors. For example, a 12-min walking trip in reality could be recorded as a 10-min one in the diary. Time errors may make true negative errors in identification of walking bouts. Therefore, a time tolerance was considered to determine bouts overlapping with trips. A preliminary study with a subset PA bout sample (n=1,534) was conducted to investigate time differences. In brief, each of the selected PA bouts had a time overlap with one declared walking trip and had similar duration with the overlapped walking trip. Durations of a walking bout and its associated walking trip were to be between 80% and 120% of their mean value. The left panel in Figure 2-7 shows a contour plot of the joint density of time differences in starting time versus ending time. More than 95% pairs of a bout and a walking trip had < 10 min time differences.

Based on this preliminary study, a 10-min time tolerance was considered. This time tolerance value was further studied and verified in the following process.

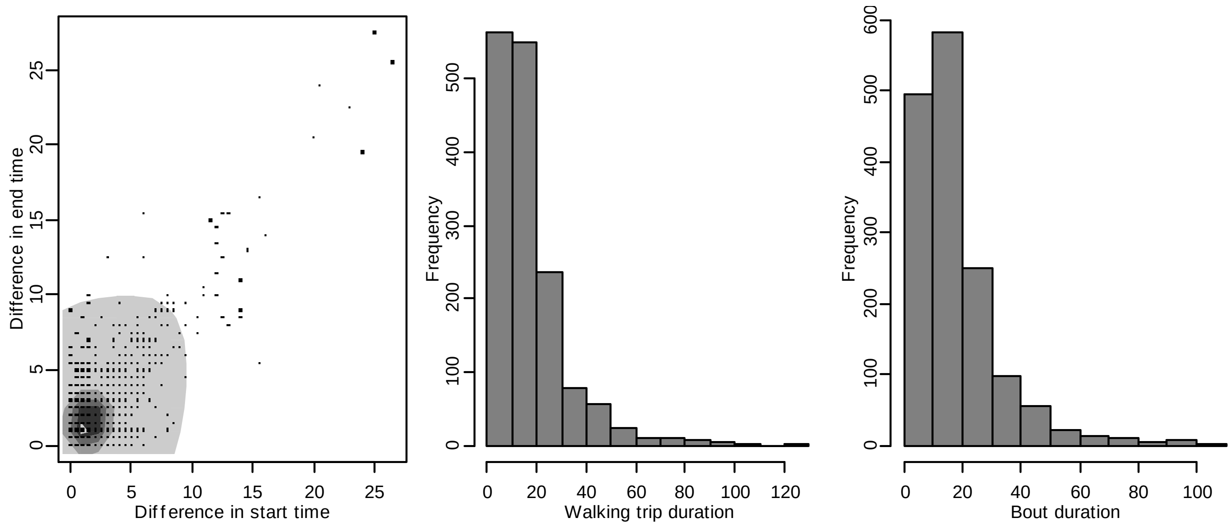

Figure 2-7. (Left) Time differences between PA bouts and declared walking bouts (x-axis is the time difference of start time and y-axis is the time difference of end time); (Center) Distribution of walking trips; (Right) Duration of PA bouts in the subset

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 2-4 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 2-4 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 2-4 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 2-4 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.[4] 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.[21]

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.[22] (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 2-4 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 2-4 F).

The characteristics of the algorithm-development PA bouts used to generate the criteria and scenarios are summarized in Table 2-1. 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 2-4 A). Furthermore, 8 overlapped with a car or transit trip, supportingWalk3-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 2-8 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 2-4) 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.

Table 2-1. 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.

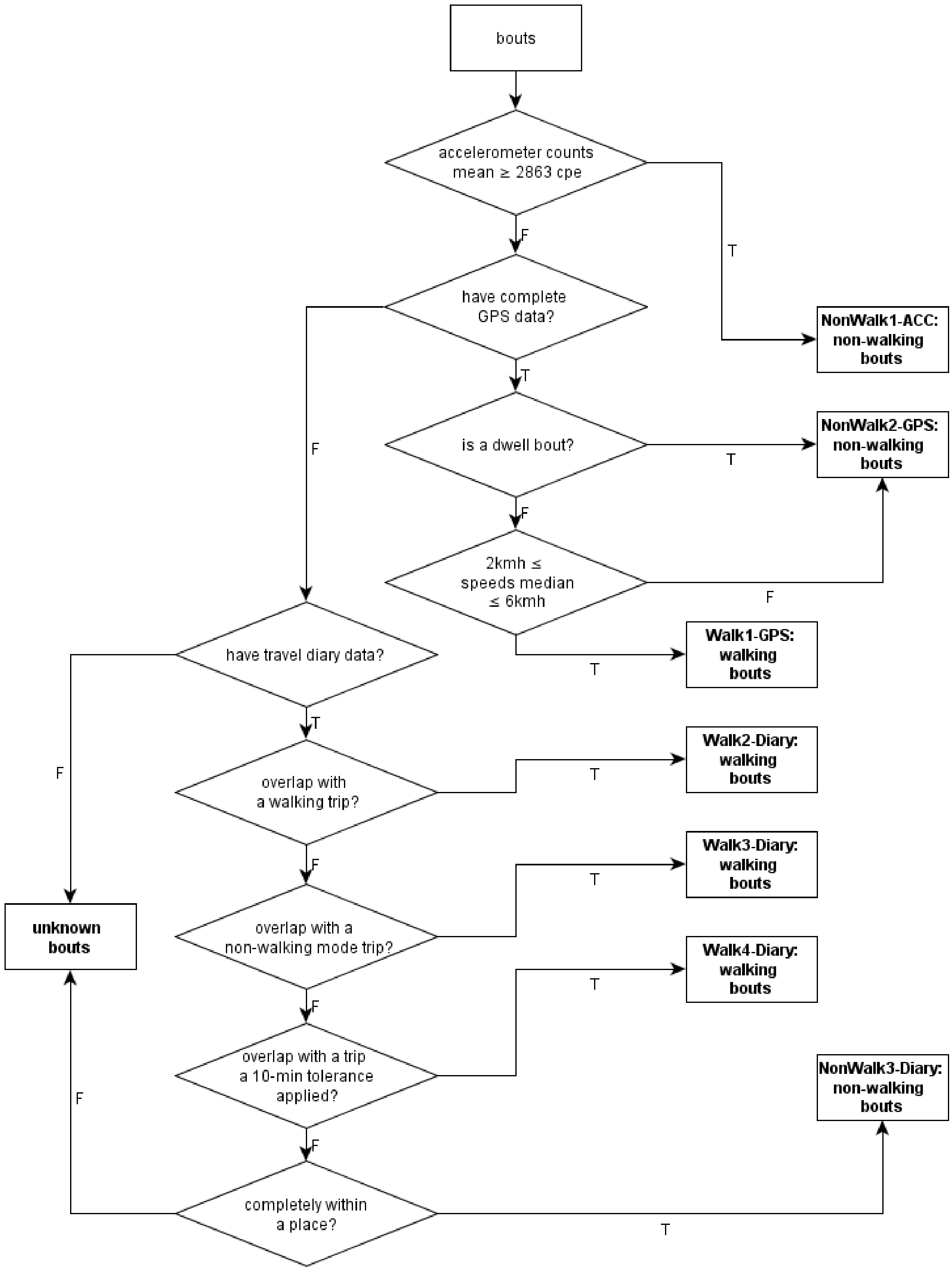

Figure 2-8. The decision tree algorithm shows sequential application of the seven scenarios to classify PA bouts as walking or non-walking.

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 (Table 2-2).

Table 2-2. Demographic characteristics of the sample (n=706)

		Mean, Count	**SD, %**
Age		51.0	13.1
Sex	Female	444	63%
Ethnicity(non-Hispanic white)		550	79%
Household income	<49K USD	260	39%
	50-99K USD	267	40%
	>100K USD	138	21%
Education (college graduate or above)		472	70%

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.

<u>Verification of algorithm results</u>

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 2-3), 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).

Table 2-3. 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							

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}.[16] 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 2-4). On average, at the person level, participants had 1.7 walking bouts (total 25.4 min) per day (Table 2-5). 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 (Table 2-6). 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.

Table 2-4. Bout classification results from the algorithm

Scenario	**Number**	**Overall %**	**Within group %**	**Duration (min)**		**Mean count (counts per 30 s epoch)**	
				Mean	**(SD)**	**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 2-5. 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

Table 2-6. Sensitivity analysis summary

Algorithm	**Changed parameter value (difference Δ)**	**% of classified as walking in the sample (Δ)**	**% of agreements in the algorithm-verification sample (Δ)**	
			Analyst A	**Analyst B**
Base	NA	58.5	92	98
1	VL=3,006 cpe (+5%)	58.6 (+0.1)	92 (0.0)	98 (0.0)
2	VL=2,720 cpe (-5%)	58.1 (-0.4)	91 (-1.0)	97 (-1.0)
3	SR= [2.5, 6.5 km/h] (+0.5 km/h)	53.4 (-5.1)	87 (-5.0)	93 (-5.0)
4	SR= [1.5, 5.5 km/h] (-0.5 km/h)	61.5 (+3)	92 (0.0)	92 (-6.0)
5	DR=69.3 ft (+5%)	53.4 (-5.1)	87 (-5.0)	93 (-5.0)
6	DR=62.7 ft (-5%)	53.4 (-5.1)	87 (-5.0)	93 (-5.0)
7	TT=5 min (-5 min)	57.4 (-1.1)	92 (0.0)	94 (-4.0)
8	TT=15 min (+5 min)	59.4 (+0.9)	91 (-1.0)	97 (-1.0)

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 2-5). 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.[23] 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).[24] 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.[2]

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.[25] 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.[26] 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).[20] 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).[4, 27] Studies using GPS data confirm that travel diaries substantially under-report trips, especially short trips.[28] 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.[25] 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 2-4 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 2-54). 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.[29-32] 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.[33], 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.

Conclusion

GPS and travel diary information can be helpful in classifying most accelerometer-derived PA bouts into walking or non-walking behavior. Objective measures of walking from the study sample were substantially different from subjective measures based on national samples of other studies although they were not completely comparable. Using two objective instruments (accelerometer and GPS unit) and one subjective instrument (travel diary) provides more objective and accurate estimates of walking.

In future research, the algorithm will be improved using statistical methods of multivariate classification. This would address the issue of rule-based deterministic criteria of the current algorithm because that approach considers multiple variables simultaneously. In addition, objectively measured walking should be further analyzed in comparison with subjectively measured walking (i.e., from IPAQ). It would answer many important questions about whether disagreements are influenced by individual characteristics and how walking measures are different at person- and activity-levels.

References

1. Ogilvie D, Foster CE, Rothnie H, Cavill N, Hamilton V, Fitzsimons CF, et al. Interventions to promote walking: systematic review. British Medical Journal 2007;334(7605):1204.
2. 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(1).
3. 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.
4. Stopher PR, Greaves SP. Household travel surveys: Where are we going? Transportation Research Part A: Policy and Practice 2007;41(5):367-381.
5. 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.
6. 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):421-431.
7. 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.
8. 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):532-540.
9. 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):102-108.
10. 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):508-515.
11. Moudon AV, Saelens BE, Rutherford S, Hallenbeck M. A report on participant sampling and recruitment for travel and physical activity data collection. [Seattle, Wash.]: Transportation Northwest; 2009. Report No.: TNW2009-05.

12. Federal Highway Administration. 2001 National Household Travel Survey: user's guide. Washington, DC: US Department of Transportation; 2004.
13. Mâsse LC, Fuemmeler BF, Anderson CB, Matthews CE, Trost SG, Catellier DJ, et al. 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.
14. 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.
15. Matthews CE. Calibration of accelerometer output for adults. Medicine and Science in Sports and Exercise 2005;37(11):512-22.
16. 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):368-374.
17. Loecher M, Berlin School of Economics and Law. RgoogleMaps: Overlays on Google map tiles in R. R package version 1.2.0. http://CRAN.R-project.org/package=RgoogleMaps. In; 2012.
18. R Core Team. R: A Language and Environment for Statistical Computing. In. Vienna, Austria: R Foundation for Statistical Computing; 2012.
19. Cho GH, Rodríguez DA, Evenson KR. Identifying walking trips using GPS data. Medicine and Science in Sports and Exercise 2011;43(2):365.
20. 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.
21. Reilly JJ, Kelly LA, Montgomery C, Jackson DM, Slater C, Grant S, et al. Validation of Actigraph accelerometer estimates of total energy expenditure in young children. International Journal of Pediatric Obesity 2006;1(3):161-167.
22. 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):777-781.
23. Tudor-Locke C, van der Ploeg HP, Bowles HR, Bittman M, Fisher K, Merom D, et al. Walking behaviours from the 1965-2003 American Heritage Time Use Study (AHTUS). The International Journal of Behavioral Nutrition and Physical Activity 2007;4(45).

24. Bauman A, Bull F, Chey T, Craig CL, Ainsworth BE, Sallis JF, et al. The International Prevalence Study on Physical Activity: results from 20 countries. International Journal of Behavioral Nutrition and Physical Activity 2009;6(21).
25. 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-S317.
26. 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.
27. 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.
28. 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):285-297.
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):181-188.
30. Rodríguez DA, Cho GH, Evenson KR, Conway TL, Cohen D, Ghosh-Dastidar B, et al. Out and about: Association of the built environment with physical activity behaviors of adolescent females. Health and Place 2012;18(1):55-62.
31. 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.
32. 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):1502-1508.
33. Catellier DJ, Hannan PJ, Murray DM, Addy CL, Conway TL, Yang S, et al. Imputation of missing data when measuring physical activity by accelerometry. Medicine and Science in Sports and Exercise 2005;37(11):555-62.

Chapter 3 The Contribution of Walking to Total Physical Activity

Chapter Abstract

Background: To assess the contribution of walking to total physical activity (PA), it is necessary to measure PA related to walking separately from other types of PA and to examine the possible interactions between them. Most studies have focused on specific types of PA (e.g., exercise) without analyzing their inter-type relationships. Moreover, they used subjective and therefore inaccurate PA measures.

Purpose: The relationships between walking and other types of PA were analyzed using objective and classified PA measures to determine whether walking had a positive, negative, or null effect on other types of activity.

Methods: PA measures classified into walking, other types of PA, and total PA, were obtained from 657 participants' 7-day accelerometer, GPS, and travel diary data. The effect of walking-related PA on other types of PA was modeled at the person-level, adjusting for demographic, socioeconomic, psycho-social, and environmental density covariates and at the day-level, adjusting for weather and day of the week effect.

Results: On average, 60% of total PA came from walking. A 15-min increase in walking was associated with a 1.6-min increase in other types of PA at the person-level. Among moderate walkers (walk 15–30 min per day), the same increase in walking was associated with a 7.8-min increase in other types of PA.

Conclusions: Increased walking actually contributed to total PA without reducing other types of PA.

Introduction

According to the National Health and Nutritional Examination Survey (NHANES), less than 5% of American adults satisfied the recommended level of physical activity (PA).[1] Many public health interventions promote walking because walking is one of the most popular and accessible forms of PA.[2-4] Also, as frequent short bouts of PA (≥ 10 min) may provide substantial health benefits[5], walking trips that are part of daily life might be a significant source of PA.[6] While most studies have assumed that more walking results in more total PA, none have addressed the question of whether walking produces a net increase in total PA or whether it reduces other types of PA.

It is important to understand whether and to what extent walking contributes to total PA for public health policy development. If more walking has a negative impact on other types of PA, health interventions simply promoting walking may not effectively maximize total PA. In contrast, if more walking synergistically increases other types of PA, promoting walking would also serve to increase total PA. The effects of walking on other PA may be selectively shown among certain population classes. A better understanding of this contribution of walking may serve to tailor public health interventions specifying target population groups.

In spite of the importance of this issue, no study has directly addressed the actual contribution of walking to PA. A few studies focused on PA gained from exercise, but not general walking. Structured exercise interventions were found to increase total PA or total energy expenditure in children and young adults[7, 8] but not in elderly persons.[9, 10] The young may become more physically active after participation in exercise while the elderly may seek to regulate total PA and reduce non-exercise PA when prompted to exercise. As in the case of structured exercise, walking may have certain effects on other types of PA. Yet little is known about the possible effects of walking on PA.

One of the major reasons is the lack of objective and type-specific PA measurements. While it is relatively easy to isolate the amount of PA obtained from structured exercise sessions (e.g., using pedometer count difference before and after sessions[11]), the amount of PA derived from walking is difficult to estimate because most walking occurs in unfixed space and time of daily life. Questionnaire-based subjective estimates of walking could be a solution. But subjective data are inaccurate and incomparable.[12, 13] To address the issue of estimating walking objectively, it is required to monitor subjects all waking times with the capability of distinguishing between different PA types.

Recent technical advances in real-time behavior tracking sensors, including accelerometers and GPS units, provide a unique opportunity to precisely measure and classify PA objectively in space and time. Combining data from accelerometers (PA intensity) and GPS units (location and speed) provides more objective and precise identification of walking under free-living conditions.[14]

Based on data on type-specific PA, the current study examined the inter-relationships between walking and other types of PA within total PA at the person-level as well as at the day-level. The primary aim of this study was to determine whether the levels of walking-related PA have a positive or negative effect on the levels of other PA types, after adjusting for sociodemographic, psychosocial, and environmental correlates. Environmental correlates were included in the analyses because of the known associations between walking and environment.[15, 16]

Methods

Data collection

Data were from the Travel Assessment and Community (TRAC) project. Participants were recruited by random digit dialing methods within 773 Census block groups having the same range of household income ($7,400–$83,000), race (3.1%–81.0% white), home values, net residential density, housing type, availability of proximate neighborhood services, and levels of bus ridership in King County, WA.[17] Of the individuals contacted by phone, 53% refused to participate; 41% were not eligible or provided bad/incomplete information; and 6% agreed to participate. Eligible participants were to be ≥ 20 years old; be able to walk outside their home without assistance; and be English-speaking. The initial recruitment target sample size was 750. Participants provided informed consent to participate and the study was approved by the Seattle Children's Research Foundation Institutional Review Board. Data were collected from recruited participants on a rolling basis between July 2008 and July 2009.

Participants were instructed to wear a uni-axial accelerometer (GT1M, Actigraph LLC, Fort Walton Beach, FL) on their hips, to carry a GPS unit (DG-100 GPS data logger, GlobalSat, Taipei, Taiwan), and to record a travel diary (modified from the National Household Travel Survey place-based formats[18]) for 7 consecutive days from a day selected by participants. The accelerometers and GPS units were configured to record data in 30 s epochs or intervals. Instruments were mailed back to the research team and data were screened. Participants were asked to rewear instruments and to record a travel diary up to two additional times until their data met the 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). Finally, participants completed a self-administered questionnaire, which included demographic information and perceptions of PA.

Data processing

The present analysis included data on complete days of data, defined as 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.[19, 20] Days without a recorded place in travel diary (including a place at the start of day: e.g., home) were excluded because participants were likely not to have worn the instruments when they skipped the travel diary. A complete day might or might not have had GPS data. If and when participants stayed all day in a building with no satellite signal reception, they could have no GPS data at all. Accelerometer data on complete days were assumed to record all PA of interest during wake time except for aquatic activities (i.e., swimming, showering). No data were removed or imputed.

Data were processed to identify walking and non-walking PA bouts by a previously documented algorithm.[14] PA bouts were initially identified for accelerometer counts exceeding 500 counts per 30 s epoch (cpe) within a 7-min rolling time window, allowing for up to 2 min of epochs below that threshold.[21, 22] Next, PA bouts having either GPS speeds and traces constituent to walking or time overlaps with a declared walking trip from the travel diary were classified as walking. In cases where PA bouts had insufficient GPS and no travel diary data, they were classified as unknown. In the present study, person-days with unknown bouts were excluded because they might cause a misclassification bias. They were assumed to be missing at random.

Classified PA bouts were aggregated to the person-day level first, and then they were aggregated again to person level. For each participant, daily minutes spent in each type of PA were calculated as:

- Total physical activity per day (TPD): Total time spent in all PA bouts per average day
- Walking per day (WPD): Total time spent in walking PA bouts per average day
- Other types of PA per day (OPD): Total time spent in non-walking (other types of) PA bouts per average day

Demographic and socioeconomic information came from the survey and included were age, sex, ethnicity (non-Hispanic white or not), education level (college graduate or higher or some college or below), employment status (full time, part time, retired, or unemployed), and household income (<50K, 50−100K, >100K USD).

Perception measures known to be related with moderate PA[23, 24] came from five sections in the survey, each of which included 3 to 16 questions (Table 3-1). The questions were answered on a Likert scale (1=never; 2=rarely; 3=some times; 4=often; and 5=very often). Numerical codes were averaged by survey section to create five PA perception variables: self-confidence, enjoyment, benefits, barriers, and social-support, all ranging from 1 to 5.

Table 3-1. PA perception variables

PA perception **survey section**	**Sub-questions**
(1) Self-confidence: 3 questions Self-confidence for moderate PA: "How sure you are that you could do moderate PA in these situations?"	1. Do PA even though I am feeling sad or highly stressed. 2. Stick to my program of PA even when family or social life takes a lot of time. 3. I will set aside time for regular PA.
(2) Enjoyment: 3 questions Enjoyment of moderate PA	1. I enjoy doing moderate PA. 2. I enjoy the feeling I get while doing moderate PA. 3. I enjoy the feeling I get after doing moderate PA.
(3) Benefits: 10 questions Benefits of regular PA: "If I participate in regular PA, then:"	1. I will feel less depressed and/or bored. 2. I will improve my self-esteem. 3. I will meet new people. 4. I will lose weight or improve my shape. 5. I will build up my muscle strength. 6. I will feel less tension and stress. 7. I will improve my health or reduce my risk of disease. 8. I will do better at my job. 9. I will feel more attractive. 10. I will improve my heart & lung fitness.
(4) Barriers: 16 questions Barriers to Regular PA "How often do the following prevent you from getting regular PA?"	1. Self-conscious about my looks when I exercise. 2. Lack of interest in exercise or PA. 3. Lack of self-discipline. 4. Lack of time. 5. Lack of energy. 6. Lack of company. 7. Lack of enjoyment from exercise or PA. 8. Discouragement. 9. Lack of equipment. 10. Lack of good weather. 11. Lack of skills. 12. Lack of facilities or space. 13. Lack of knowledge on how to exercise. 14. Lack of good health. 15. Fear of injury. 16. Activity facilities or space is too far away.
(5) Social-support: 6 questions Social support	1. Family did PA with me. 2. Friends did PA with me. 3. Family offered to do PA with me. 4. Friends offered to do PA with me. 5. Family encouraged me to do PA. 6. Friends encouraged me to do PA.

Participants' home addresses were geocoded to parcel polygon centroids by ArcGIS 9.3.1 (ESRI, 1999) using the King County, WA address-point GIS layer as a reference. For participants living in large parcels (e.g., apartment complexes), mean XY coordinates were established from GPS data associated with the participants' travel diary reporting their being at home.

Residential unit density and job density within a 10-min walking distance (833m) were measured at geocoded home locations, using residential unit and job density SmartMaps developed by the Urban Form Lab.[25] The residential density SmartMap was created from the King County assessor's data downloaded in January 2010. Residential unit numbers were the sum of units in single houses, apartment and condominium buildings. The job density SmartMap was generated by joining the same assessor's data and the North American Industry Classification System (NAICS) data, downloaded in December 2006. The number of jobs associated with the difference land uses was estimated by the Urban Form Lab.[26]

Meteorological data was obtained from the National Weather Service Office at the airport nearest the study location (Seattle-Tacoma Washington Airport weather station).[27] Daily mean temperature (° F) and total precipitation (inches) were used in this study. Because PA may not be preferred especially outdoor at hot and cold weather,[28] the continuous mean temperature values were categorized into three levels ($(0, 50)$, $[50, 65)$, and $[65, \infty)$ ° F). Days having precipitation > 0.05 inches were considered to have rain.

Data analysis

The relationship between WPD and OPD was analyzed in two steps. First, variances in the relationship were examined between participants at the person-level. Second, variances between days within a participant were further investigated considering day-level variables like weather and day of the week effects.[29] At the person-level, multivariate OLS regression models estimated the association between WPD levels (the explanatory variable) and OPD levels (the response variable). The regression coefficient for WPD was used to determine whether the effect of WPD on OPD was positive (synergistic) or negative (substitutive), after adjusting for sociodemographic, PA perception, and built environment variables.[15, 16, 30-32] In cases where the relationship between WPD and OPD was linear only for certain ranges of WPD, four additional regression models were developed for subgroups determined by participants' WPD values. The four subgroups are insufficient walkers (WPD < 15 min per day), moderate walkers (15 ≤ WPD < 30 min per day), sufficient walkers (30 ≤ WPD < 45 min per day), and active walkers (WPD ≥ 45 min per day).

At the day-level, the relationship between WPD and OPD was analyzed using hierarchical linear regression models (random intercept models). Day-level WPD and OPD values were grouped by participants. For each j-th participant's i days, OPD was estimated by WPD and day-level covariates (weather[28] and day of the week[29]) as:

$$OPD_{i,j} \sim WPD_{i,j} + (Weather_{i,j} + DayofWeek_{i,j}) + \alpha_j,$$

where $\alpha_j \sim N(0, \sigma_j^{\ 2})$, assuming α_j adjusts person-level variances.

Results

Data collection and processing

In total, 706 participants' LifeLogs were used, which yielded 4,702 person-days. Among these, 232 person-days (5%) were dropped because they had PA bouts with no GPS and no travel diary data, which could not be classified as either walking or non-walking PA. This removed data from 5 participants. Next, 25 participants who did not have PA bouts at all during their PA assessment periods were excluded. Finally, 19 participants were dropped because they did not provide survey data. Dropped data for incompleteness were considered as missing at random. The final sample consisted of 657 participants' 4,283 complete person-days (average 6.5 per person, SD=1.7), which contained 12,323,234 LifeLog units. Of the final sample, 370 participants (56%) had data for all days of the week (2,731 days). The reminder of the participants (n=287, 44%) missed some days of the week and 608 (93%) had at least one weekday and one weekend day. Geocoded home points were plotted in Figure 3-1.

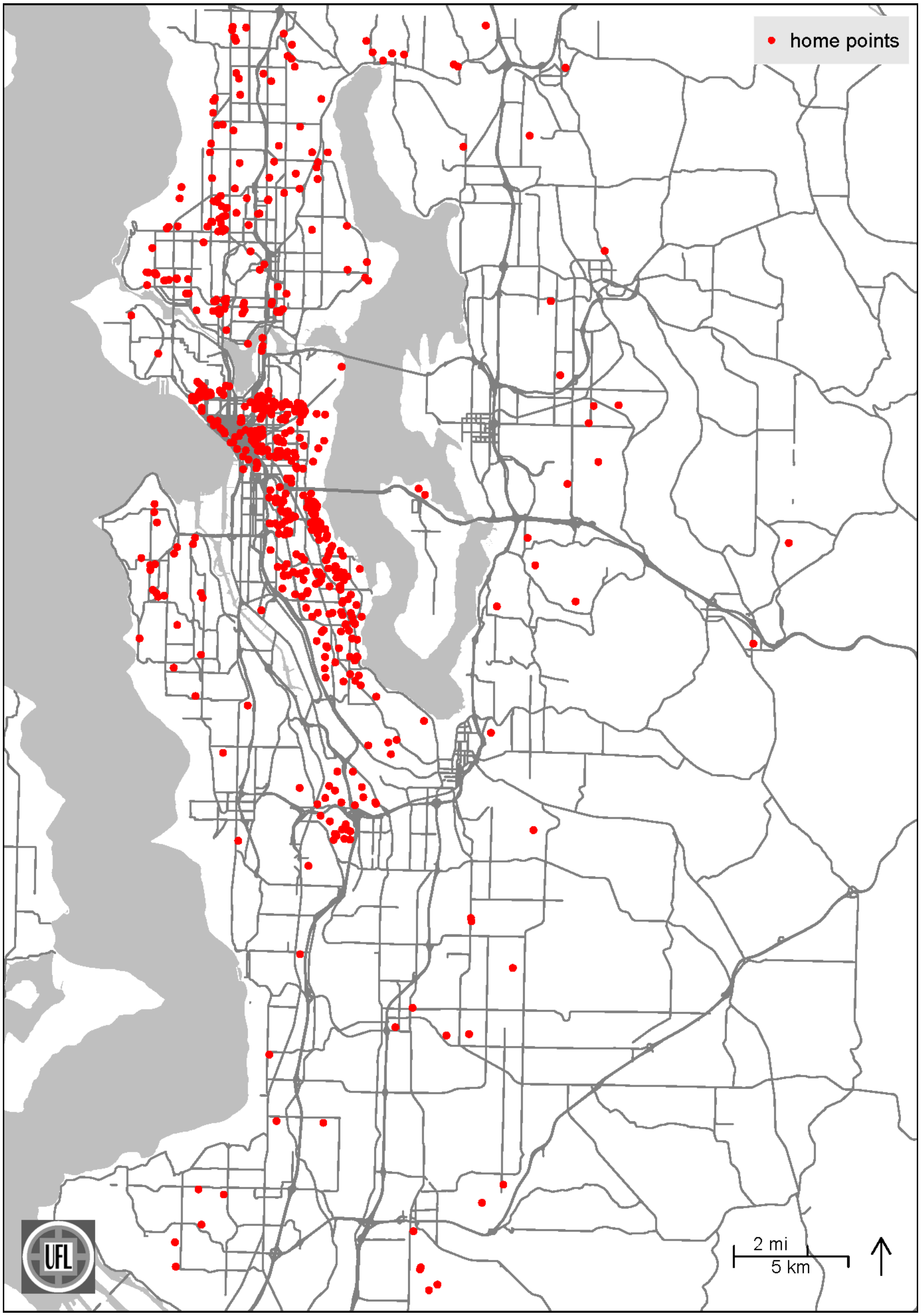

Figure 3-1. Geocoded home points of the sample (n=657)

On average, the sample population was 51 yr (SD=13); 63% were female; 81% non-Hispanic white; 71% college graduate or above; and 54% employed full time (Table 3-2). About 38% of the participants' households had an annual household income < 49,000 USD; 41% between 50,000 and 99,000 USD; and 21% 100,000 USD or more. Measures of perception showed high values for enjoyment and benefits of PA. As expected due to the sampling approach taken, both residential unit and job densities at home locations were above the County's average.

Table 3-2. Demographic characteristics of the sample

		Mean, N	**SD, %**
Age		50.8	13.0
Sex (Female)		412	63.3
Ethnicity(non-Hispanic white)		521	80.5
Household income	<49K USD	240	37.9
	50-99K USD	260	41.0
	≥100K USD	134	21.1
Education (college graduate or above)		461	71.3
Employment	Unemployed	37	5.9
	Full-time	341	54.0
	Part-time	151	23.9
	Retired	102	16.2
PA perception	Self-confidence	3.92	0.93
	Enjoy	4.28	0.84
	Benefit	4.21	0.61
	Barrier	2.20	0.68
	Social-support	2.36	0.91
Built environment	Residential unit density (du/acre)	20.50	14.40
	Job density (job/acre)	34.24	63.15

* NA values are excluded in calculating %.

On average, each participant spent 42.1 min of TPD, 26.6 min of WPD, and 15.5 min of OPD per day (Table 3-3). Approximately 60% of the total amount of PA was from walking. Among the sample, 375 participants (57%) had TPD ≥ 30 min when counting all PA bouts and 273 participants (42%) had TPD ≥ 30 min when including only PA bouts ≥ 10 min.

Table 3-3. Distributions of WPD and OPD at the person-level

	TPD	WPD	OPD
Minimum	1.0	0.0	0.0
Maximum	277.3	163.9	170.2
Median	34.6	19.8	9.8
Mean	42.1	26.6	15.5
SD	31.7	23.8	18.1

(n=657, unit=minute per day per person, all bouts included ≥ 5 min)

Data analysis

The person-level model showed that WPD was positively associated with OPD, after adjusting for sociodemographic and environmental variables (Model 1) (Table 3-4) . The 0.11 coefficient meant that a 15-min increase in WPD was associated with a 1.6-min increase in OPD. Among covariates, sex and residential unit density were significantly associated with OPD. Females had 3.08 min less OPD than males when they had the same level of walking and the values of all other variables were equal. An additional residential unit per acre was associated with a decrease of OPD by 0.13 min (7.8 sec). According to R^2, the model explains 8% of total amount of variance in OPD.

In addition, sex and residential unit density were significantly associated with OPD at the level of p-value <0.05. Females had 3 min less OPD than males and an additional 10 residential units per acre within home neighborhood was associated with 1.3 min less OPD when WPD and other variables adjusted for. None of the PA perception variables were significant.

Table 3-4. Full sample model results at the person-level

Model 1: full sample (n=657)		Est	SE	
(Intercept)		14.34	9.45	
WPD		**0.11**	**0.03**	**
Sex	Female	**-3.08**	**1.53**	*
Age		-0.08	0.07	
Ethnicity	Non-Hispanic white	-0.51	1.85	
Income	[50, 100K)	1.61	1.72	
(base: [0, 50K))	[100K,)	1.81	2.12	
Employment	Full time	-4.67	3.19	
(base: full-time)	Part time	0.08	3.38	
PA perception	Retired	-4.89	3.78	
	Social-confidence	0.77	1.01	
	Enjoy	0.82	1.15	
	Benefit	0.91	1.44	
	Barrier	-0.75	1.42	
	Social-support	0.83	0.88	
Built environment	Residential unit density	**-0.13**	**0.07**	*
	Job density	-0.02	0.01	
R^2		0.08		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

In the scatter plot of participants' WPD and OPD at the person-level, the LOWESS smooth line (bold) along with the fitted regression line (straight) shows that the positive association is more apparent among participants who had WPD roughly between 0 and 40 min per day (Figure 3-2). The regression fitted line has a slope of 0.11, which is drawn with an intercept derived from a theoretic average participant (51 years old, female; non-Hispanic-white; income between 50K and 100K; full-time employed; having mean values for PA perception variables and environmental variables). For participants having roughly more than 40 min of WPD, there was a large gap between the LOWESS and the fitted regression lines.

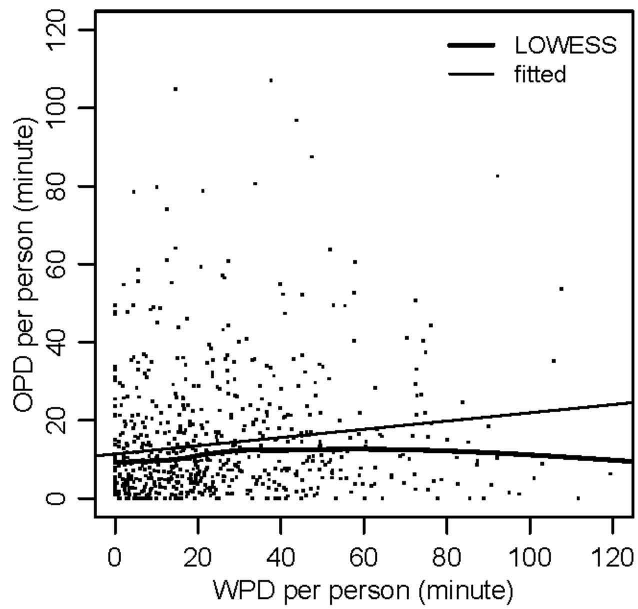

Figure 3-2. Scatter plot of WPD versus OPD for 657 participants. It includes a LOWESS smooth line and a fitted regression line of Model 1. The fitted line is drawn from a theoretic average participant of the sample (51 years old, female; non-Hispanic-white; income between 50K and 100K; full-time employed; having mean values for PA perception variables and environmental variables).

Regression models were estimated for 4 subgroups determined by participants' WPD values (Table 3-5). Model 2-1 was for the insufficient-walker subgroup (WPD < 15 min; n=253 participants), Model 2-2 for the moderate-walker subgroup (15 ≤ WPD < 30 min; n=181 participants), Model 2-3 for the sufficient-walker subgroup (30 ≤ WPD < 45 min; n=92 participants), and Model 2-4 for the active-walker subgroup (WPD ≥ 45 min; n=131). WPD was significant only in Model 2-2. The coefficient was 0.52, almost four times larger than the coefficient of the entire sample (0.11 in Table 3-4), meaning a 15-min increase in WPD was associated with a 7.8-min increase in OPD. In the same model (2-2), household income was also significant. Participants with household income between \$50K and \$100K had 7 min more OPD than those with ≤ \$50K. In Model 2-1, age was significant. In Models 2-3 and 2-4, none of the variables were significant at the level of $p<0.05$.

The day-level model shows that WPD was negatively associated with OPD at the day-level (Model 3 in Table 3-6). Variables whose 95% CI did not cover null value (0) were considered as significant (marked with * in Table 3-6 and Table 3-7). The estimated coefficient of WPD is -0.06 and 95% CI range is [-0.09, -0.03]. This means that on average a 15-min increase in WPD was associated with a 0.9-min decrease in OPD. Rainy days had 1.8 min less OPD than dry days. Saturdays and Sundays had 4 min and 2.7 min more OPD than week days, respectively. When participants were partitioned into subgroups (Models 4 in Table 3-7), WPD was significant only among the moderate-walker and the active-walker subgroups. The coefficient was -0.09 for the moderate-walker subgroup (Model 4-3) and -0.10 for the active-walker subgroup (Model 4-4).

Table 3-5.Subgroup model results at the person-level

		Model 2-1: ***Insufficient walkers*** WPD [0, 15) n=253			**Model 2-2:** ***Moderate walkers*** WPD [15, 30) n=181			**Model 2-3:** ***Sufficient walkers*** WPD [30, 45) n=92			**Model 2-4:** ***Active walkers*** WPD [45, ∞) n=131		
		Est	SE		Est	SE		Est	SE		Est	SE	
(Intercept)		21.65	14.57		11.37	15.24		-8.57	29.64		9.24	37.74	
WPD		0.15	0.24		**0.52**	**0.25**	*	-0.01	0.54		0.11	0.12	
Sex	Female	-1.65	2.39		-0.58	2.27		-0.62	4.85		-9.50	4.87	.
Age		**-0.22**	**0.10**	*	-0.02	0.10		0.33	0.20		-0.13	0.23	
Ethnicity	Non-Hispanic white	-0.13	2.56		-2.10	3.49		0.39	6.17		1.46	6.37	
Income	[50, 100K)	1.00	2.57		**7.04**	**2.72**	*	-7.00	5.47		5.12	5.51	
(base: [0, 50K))	[100K,)	2.77	3.33		6.23	3.16	.	1.34	6.94		-1.80	6.68	
Employment	Full time	-7.18	4.81		-4.26	4.46		10.55	7.88		-10.87	18.48	
(base: full-time)	Part time	-1.39	5.20		2.59	4.77		12.60	8.37		-5.22	18.71	
	Retired	-4.20	5.73		-3.03	5.52		-2.40	9.85		-10.39	19.96	
PA perception	Social-confidence	0.43	1.44		1.37	1.52		2.16	3.38		-1.32	3.85	
	Enjoy	2.05	1.73		-1.08	1.77		-3.87	3.39		2.76	3.95	
	Benefit	0.20	2.16		1.51	2.20		1.68	4.11		3.69	4.90	
	Barrier	-2.43	2.23		-2.97	2.24		3.91	3.49		2.82	5.25	
	Social-support	1.89	1.33		-1.75	1.43		-1.76	2.96		1.46	2.60	
Built environment	Residential unit density	-0.06	0.12		-0.17	0.09	.	0.08	0.19		-0.38	0.20	.
	Job density	-0.03	0.03		0.02	0.02		-0.04	0.04		0.00	0.04	
R^2		0.137			0.138			0.165			0.136		

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 3-6. Full sample model results at the day-level

Model 3: full sample (n=657)		Est.	95% CI Lower	Upper	
(Intercept)		17.23	15.17	19.29	*
WPD		**-0.06**	**-0.09**	**-0.03**	*****
Rain >.05 inches		**-1.83**	**-3.60**	**-0.06**	*****
Temperature (°F)	[55, 65)	-1.03	-3.13	1.07	
(Base: [25, 55))	[68, 80)	-1.60	-4.69	1.49	
Day of week	Saturday	**4.01**	**1.91**	**6.12**	*****
(Base: weekday)	Sunday	**2.70**	**0.57**	**4.84**	*****
Random effects (σ)		14.98			
AIC score		10,770			
BIC score		6,102			

Signif. codes: * 95% CI range covers NULL value(0)

Table 3-7. Subgroup model results at the day-level

		Model 4-1: ***Insufficient walkers*** (n=253)				**Model 4-2:** ***Moderate walkers*** (n=181)				**Model 4-3:** ***Sufficient walkers*** (n=92)				**Model 4-4:** ***Extreme walkers*** (n=131)			
		Est.	95% CI			Est.	95% CI			Est.	95% CI			Est.	95% CI		
			Lower	Upper			Lower	Upper			Lower	Upper			Lower	Upper	
(Intercept)		15.79	12.99	18.59	*	16.74	13.52	19.97	*	18.53	11.95	25.12	*	24.59	17.89	31.29	*
WPD		-0.05	-0.14	0.04		**-0.09**	**-0.14**	**-0.03**	*****	-0.07	-0.15	0.01		**-0.10**	**-0.15**	**-0.05**	*****
Rain >.05 inches		-1.09	-3.49	1.31		-1.08	-3.95	1.80		0.70	-4.93	6.32		**-6.10**	**-11.29**	**-0.92**	*****
Temperature (°F)	[55, 65)	**-3.28**	**-6.22**	**-0.33**	*****	-1.32	-4.65	2.01		0.80	-5.67	7.26		0.14	-5.82	6.10	
(Base: [25, 55))	[68, 80)	-3.16	-7.66	1.34		-2.65	-7.65	2.35		0.12	-9.25	9.49		0.38	-7.81	8.56	
Day of week	Saturday	**3.27**	**0.33**	**6.21**	*****	2.07	-1.43	5.58		4.35	-2.40	11.09		**7.09**	**1.51**	**12.67**	*****
(Base: weekday)	Sunday	-0.06	-2.99	2.88		**4.95**	**1.37**	**8.53**	*****	0.30	-6.53	7.14		5.62	-0.19	11.43	
Random effects (σ)		14.22				10.96				15.33				19.43			
AIC score		14,555				10,770				6,062				8,421			
BIC score		14,604				10,815				6,102				8,464			

Signif. codes: * 95% CI range covers NULL value(0)

Discussion

The present study found that walking had a positive effect on other types of walking at the person-level. A 15-min increase in WPD was associated with a 1.6-min increase in OPD at the person-level (Table 3-4). Thus, walking contributed to total PA without reducing other types of PA. The positive effect was large among the moderate-walker subgroup who walked between 15 and 30 min per day. A 15-min increase in WPD was associated with a 7.8 min increase in OPD (Model 2-2 in Table 3-5). Although the direction of causality is not clear, assuming that a change in walking affects other types of PA at the person-level, a 15-min increase in walking would synergistically bring total PA per day to 22.8 min (=15 + 7.8 min). This large synergistic effect (52%) could be critical for persons who are marginally under the daily threshold of PA recommended in public health (30 min of moderate PA per day[1, 5]). In the study sample, 6.8% of the participants were 5 min short of the 30-min public health standard.

When participants were partitioned by their WPD minutes, the synergistic association was not present among participants who had WPD < 15 or ≥ 30 min per day. These participants had no strong association patterns between the PA types used in this study (WPD and OPD). It is possible that PA types would need to be further classified to identify relationships between PA types. For example, certain patterns may exist only among discretionary-type PA (e.g., related to unstructured activities) after excluding obligatory-type PA (e.g., related to occupational activities). People may easily change discretionary-type walking but not obligatory-type walking. Other types of PA could be also distinguished as discretionary (e.g., recreational) or obligatory (e.g., domestic maintenance). One study that attempted to frame youth’s daily time use with PA levels also distinguished PA types by places where activities were performed (e.g., walking destinations), yielding four types of PA: home-related, school-related, discretionary and obligatory.[33] Simultaneous consideration of PA types as walking versus non-walking, and discretionary versus obligatory, would provide the better understanding of contribution of walking to total PA.

At the day-level, WPD had a very small association with OPD. Although significant variables in the models showed negative associations, their actual effects were only up to -0.10. All day-level models (Model 3 and 4s) have WPD coefficient 95% CI ranges between -0.15 and 0.04. A 15-min increase in WPD was associated with a decrease in OPD between 2.25 and 0.6 min. Therefore, the levels of OPD were nearly constant across the levels of WPD between days grouped by participants. The current analysis included only three day-level covariates (average daily temperature, total precipitation, and effects from days of the week). These variables may not precisely capture the actual conditions of when and where participants' PA occurred. For example, local weather where PA occurred could have been different from the one observed at the nearest airport weather station. Participants' actual working days might not match the conventional calendar weekdays (e.g., nurses working on weekend shifts). Nonetheless, the current evidence suggested that there were no strong substitutive behaviors (negative effect) of walking to other types of PA between days within a participant.

The findings of no substitutive behaviors at the person-level in the current analyses are different from those of two previous studies which concluded that there were substitutive behaviors between walking and other types of PA[34, 35] Their results showed significant differences in walking but not in total PA between respondents living in high- versus low-density neighborhoods, indicating that increased walking might be associated with reduced PA of other types in high density neighborhoods. However, the design of these studies and the data they used differed considerably from those of the present study. A similar analysis with the current data was attempted but the results were still different from those of the two studies. Participants were grouped into 4 quartile subgroups by residential unit density. There were significant differences in both WPD and TPD across the subgroups (one-way ANOVA tests; $p<0.001$; data not shown). On subgroup-level average, as residential increased both WPD and TPD increased. Yet comparisons are limited because these other studies fully or partially relied on subjective PA measures, which are known to be inaccurate and to bias results[12]; and their samples were drawn from areas with about half of the residential densities (~ 6.12 du per acre; ~ 10 persons per acre) than those in the present analyses (sample mean: 20.5 du per acre). It is not clear whether the inconsistent findings resulted from different PA measurements or different study samples. However, the current study using objective PA measures and directly analyzing the relationship between walking and other types of PA showed that there was no evidence for substitutive behaviors at least at the person-level.

It is noteworthy that none of the PA perception variables were significant in the person-level models (Model 1 and 2s) at the level of $p<0.05$. PA perception appeared to be significant in explaining walking and total PA individually in previous studies.[23, 24] The current data also showed some PA perception variables were correlated with TPD, WPD, and OPD individually (data not shown). WPD may be more correlated with OPD than PA perception variables, because adding WPD in the model reduced the significance of PA perception variables.

The sample in the study does not represent the Seattle region. Because the original project aimed to assess the impact of a newly introduced light rail system, participants were recruited from the denser urban and suburban areas of King County, where transit service was available.[17] Accordingly, mean residential unit density of the sample was 20.5 units per gross acre (calculated within the 833 m buffer), which is ~23% higher than the residential densities in two other studies set in the same county: one had a sample mean of 16.6 units per gross acre (within a 0.5 mile buffer)[36], and another had 11.5 household units per residential acre (the data were census derived and averaged within neighborhoods of different sizes).[23] Because walking PA is related to residential density, future studies should consider comparing samples from a range of densities.

This study has numerous strengths including the use of objective and comparable PA measurements, the capability to isolate walking from other types of PA, and the simultaneous consideration of the interaction between walking and other types of PA. It offered the unique opportunity to analyze the actual contribution of walking to and its effects on total PA.

Conclusion

On average, walking contributed to nearly 60% of total PA. This study based on a large adult sample (n=657) showed that walking had a small but significantly positive association with other types of PA at the person-level. Thus, increased walking synergistically contributed to total PA without reducing the levels of other types of PA. The synergistic effect was more apparent among moderate walkers who walked between 15 and 30 min per day than among the entire sample. At the day-level, the effect of walking on other types of PA was less clear. The current study did not find there were substitutive behaviors between walking and other types of PA.

Future research should likely include (1) more detailed classification of PA types under the framework of discretionary versus obligatory activities; and (2) further analyses on day-level relationships with realistic day-level variables. The former would address potential confounding effects from activity types and the latter would enable micro-level analysis of PA dynamics.

References

1. 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):181-188.
2. U. S. Department of Health and Human Services. Physical Activity and Health : A Report of the Surgeon General. McLean: Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion; 1996. Report No.: 188320531X 9781883205317.
3. Ogilvie D, Foster CE, Rothnie H, Cavill N, Hamilton V, Fitzsimons CF, et al. Interventions to promote walking: systematic review. British Medical Journal 2007;334(7605):1204.
4. Ekkekakis P, Hall EE, VanLanduyt LM, Petruzzello SJ. Walking in (affective) circles: Can short walks enhance affect? Journal of Behavioral Medicine 2000;23(3):245-275.
5. Haskell WL, Lee IM, Pate RR, Powell KE, Blair SN. Physical activity and public health. Updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. Medicine and Science in Sports and Exercise 2007;39(8):1423-34.
6. Sallis JF, Frank LD, Saelens BE, Kraft MK. Active transportation and physical activity: Opportunities for collaboration on transportation and public health research. Transportation Research Part A: Policy and Practice 2004;38(4):249-268.
7. Pelclová J, Ansari WE, Vašíčková J. Is participation in after-school physical activity associated with increased total physical activity? A study of high school pupils in the Czech Republic. International Journal of Environmental Research and Public Health 2010;7(7):2853-2865.
8. Westerterp KR. Alterations in energy balance with exercise. The American Journal of Clinical Nutrition 1998;68(4):970S-974S.
9. Meijer EP, Westerterp KR, Verstappen FTJ. Effect of exercise training on total daily physical activity in elderly humans. European Journal of Applied Physiology and Occupational Physiology 1999;80(1):16-21.
10. Goran MI, Poehlman ET. Endurance training does not enhance total energy expenditure in healthy elderly persons. American Journal of Physiology-Endocrinology And Metabolism 1992;263(5):E950-E957.

11. Tudor-Locke C, Jones G, Myers A, Paterson D, Ecclestone N. Contribution of structured exercise class participation and informal walking for exercise to daily physical activity in community-dwelling older adults. Research Quarterly for Exercise and Sport 2002;73(3):350-56.
12. 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(1).
13. 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.
14. Kang B, Moudon AV, Hurvitz PM, Reichley L, Saelens BE. Walking objectively measured: Classifying accelerometer data with GPS and travel diaries. Medicine and Science in Sports and Exercise 2013;22:22.
15. Ewing R, Cervero R. Travel and the built environment. Journal of the American Planning Association 2010;76(3):265-294.
16. Saelens BE, Handy SL. Built environment correlates of walking: a review. Medicine and Science in Sports and Exercise 2008;40(7 Suppl):S550.
17. Moudon AV, Saelens BE, Rutherford S, Hallenbeck M. A report on participant sampling and recruitment for travel and physical activity data collection. [Seattle, Wash.]: Transportation Northwest; 2009. Report No.: TNW2009-05.
18. Federal Highway Administration. 2001 National Household Travel Survey: user's guide. Washington, DC: US Department of Transportation; 2004.
19. Mâsse LC, Fuemmeler BF, Anderson CB, Matthews CE, Trost SG, Catellier DJ, et al. 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.
20. 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):421-431.
21. 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.

22. Matthews CE. Calibration of accelerometer output for adults. Medicine and Science in Sports and Exercise 2005;37(11):512-22.
23. Saelens BE, Sallis JF, Frank LD, Cain KL, Conway TL, Chapman JE, et al. Neighborhood environment and psychosocial correlates of adults' physical activity. Medicine and Science in Sports and Exercise 2012;44(4):637-646.
24. Giles-Corti B, Donovan RJ. The relative influence of individual, social and physical environment determinants of physical activity. Social Science and Medicine 2002;54(12):1793-1812.
25. Hurvitz PM, Moudon AV. Home versus nonhome neighborhood: quantifying eifferences in exposure to the built environment. American Journal of Preventive Medicine 2012;42(4):411-417.
26. Hurvitz PM, Jiao J, Ulmer J, Kang B, Huang R, Moudon AV. Space, time, and movement: Integrating high-resolution GPS, GIS, and accelerometry for estimating built environment exposure. In: 2011 Association of American Geographers Annual Meeting. Seattle, WA; 2011.
27. National Weather Service Forecast Office. Preliminary Monthly Climate Data (WS Form F-6). [cited 2012 September 18]; Available from: http://www.nws.noaa.gov/climate/f6.php?wfo=sew
28. Matthews CE, Freedson PS, Hebert JR, Stanek EJ, Merriam PA, Rosal MC, et al. Seasonal variation in household, occupational, and leisure time physical activity: longitudinal analyses from the seasonal variation of blood cholesterol study. American Journal of Epidemiology 2001;153(2):172-183.
29. Trost SG, Pate RR, Freedson PS, Sallis JF, Taylor WC. Using objective physical activity measures with youth: how many days of monitoring are needed? Medicine and Science in Sports and Exercise 2000;32(2):426-31.
30. Lee C, Moudon AV. The 3Ds+ R: Quantifying land use and urban form correlates of walking. Transportation Research Part D: Transport and Environment 2006;11(3):204-215.
31. Saelens BE, Sallis JF, Frank LD. Environmental correlates of walking and cycling: findings from the transportation, urban design, and planning literatures. Annals of Behavioral Medicine 2003;25(2):80-91.

32. Handy SL, Boarnet MG, Ewing R, Killingsworth RE. How the built environment affects physical activity. American journal of preventive medicine 2002;23(2S):64-73.
33. Krizek KJ, Birnbaum AS, Levinson DM. A schematic for focusing on youth in investigations of community design and physical activity. American Journal of Health Promotion 2004;19(1):33-38.
34. Rodriguez DA, Khattak AJ, Evenson KR. Can new urbanism encourage physical activity? Journal of the American Planning Association 2006;72(1):43-54.
35. Forsyth A, Hearst M, Oakes JM, Schmitz KH. Design and destinations: factors influencing walking and total physical activity. Urban Studies 2008;45(9):1973.
36. Lee C, Moudon AV. Correlates of walking for transportation or recreation purposes. Journal of Physical Activity and Health 2006;3:77.

Chapter 4 Utilitarian and Recreational Walking

Chapter Abstract

Background: A popular form of physical activity (PA), walking is a one-of-a-kind behavior: it has multiple purposes and comes in multiple types. Mechanisms by which environments affect walking behaviors could vary by walking activity types.

Purpose: This study developed a more objective identification of utilitarian and recreational walking and tested different environmental associations with these two different walking behaviors.

Methods: Objective measures of PA derived from utilitarian and recreational walking were obtained from 657 participants' 7-day accelerometer, GPS, and travel diary data. Types of walking distinguished between walking for transportation or not. Multivariate regression models examined the effect of the participants' home environment on utilitarian and recreational walking, adjusting for sociodemographic and psycho-social characteristics.

Results: Of the identified walking bouts, 89% were utilitarian walking and 11% recreational walking. Adding subjective and objective environmental variables substantially improved model fits predicting utilitarian walking, but not those predicting recreational walking. Utilitarian walking was associated with sex (female, -), age (-), self-confidence (+), perceived barriers (-), social support for PA (-), attractiveness (+), residential density (+) and employment density (+). Recreational walking was associated with sex (female, +) and employment density (-). Utilitarian walking appeared to be the main driver for the personal and environmental variables that were also associated with total PA.

Conclusions: Utilitarian walking was well explained by different environmental correlates and should be analyzed separately from recreational walking.

Introduction

For humans, walking is the primal means of moving through space. Since the rise of motorized travel, however, walking has become a personal-level choice for transportation or recreation. A growing body of literature shows that more walking is associated with specific environmental characteristics.[1-5] However, walking is defined not only as a - behavior, but also as an activity serving various functions in people's lives in general and in their daily routines specifically.[6] Walking can be a mode of travel or a means to get exercise, or can mix both travel and exercise. The different purposes behind walking activity may trigger different mechanisms by which the built environment and other factors affect such walking activity.

A growing body of literature shows that the built environment plays an important role in predicting walking. However, most of studies have not differentiated walking by type. A few studies separated utilitarian walking from recreational walking and found they had different environmental correlates.[7-11] For example, one study found that hilly terrain was a barrier for utilitarian walking but an enabler of recreational walking.[7] Possibly, the enjoyable sceneries associated with hilly terrains attracted recreational walkers. Thus the effect of an environment feature could be modified by the type of the activity. Another study found that both utilitarian (transportation) and recreational walking were associated with functional environment factors (i.e., well-maintained walking surface), but only utilitarian walking was associated with destination factors (i.e., shop, public transport).[11] In general, empirical studies have reported that environmental associations with recreational walking were less clear than with utilitarian walking.[12] For example, one empirical study tested more than 27 objectively measured built environment variables (i.e., residential unit density, mean block size, count and distance to destinations) with a large population (n=1,608) and found that recreational (exercise) walking could not be predicted by built environment variables.[13] The study did not analyze utilitarian walking. Overall, previous research showed that walking should not be considered as a single behavior, and that aggregating different types of walking into one category may falsely or spuriously link different type of walking to environments.

Defining walking by type will necessarily remain subjective. Most prior studies obtained physical activity (PA) estimates (including walking) from survey questions. Survey questions were also used to allocate amount of walking by types, which were mostly purpose-based (utilitarian/transportation versus recreational/exercise). For example, one study obtained utilitarian and recreational walking estimates from survey questions in the International Physical Activity Questionnaires (IPAQ) to measure walking activity.[7] Due to many potential biases, such estimates are known to be inconsistent within and between individuals.[14, 15] Different participants may classify the same walking episode differently. For example, walking from home to a park for some recreational activity in that park could be classified as utilitarian walking because that walking trip itself is transportation; or it could be considered as recreational because the major activity at the destination is recreational. Thus, estimates of walking by types are subjective and can involve a misclassification bias between individuals. It is possible that inconsistent results in prior studies stemmed from this misclassification bias.

This study distinguished utilitarian walking from recreational walking in order to address the different environmental associations with these two types of walking. Furthermore, precise definitions were used to identify types of walking. Attributes of home neighborhood environments were investigated as predictors of walking, which included factors tested in past research.[1, 3, 4, 12, 16-18] Both subjective (perceived) and objective environmental factors were considered as both had been shown to play an important roles in predicting walking.[1, 19, 20] Psycho-social factors regarding PA were also included as potential moderators.[18, 19]

Methods

Participants and data collection

Participants' PA measurement and location data, demographic information, neighborhood perception, and PA perception data were from the Travel Assessment and Community (TRAC) project. Participants were recruited from 773 Census block groups in King County, WA, which had similar range of 7 variables (household income, race composition, home values, net residential density, housing type, availability of proximate neighborhood services, and levels of bus ridership).[21] Participants wore a uni-axial accelerometer (GT1M, Actigraph LLC, Fort Walton Beach, FL) on their hips and carried a GPS unit (DG-100 GPS data logger, GlobalSat, Taipei, Taiwan), and recorded a travel diary (modified from the National Household Travel Survey place-based formats[22]) for 7 consecutive days. The accelerometers and GPS units were configured to record data in 30 s epochs or intervals. They also completed a self-administered questionnaire about demographic information and perceptions toward PA. Participants provided informed consent to participate and the study was approved by the Seattle Children's Research Foundation Institutional Review Board.

Data processing

The process by which accelerometer data was integrated with GPS and travel log information to identify walking and non-walking physical activity bouts is described in Chapters 2 and 3. The present study also used PA bouts to measure PA. In brief, PA bouts are time intervals having sustained moderate- or vigorous-intensity PA for 5 min or longer. They were classified as walking or non-walking based on their GPS speeds, GPS tracks, and concurrent information on travel diary. Walking bouts were further classified as utilitarian or recreational walking using a method described below.

Classification of utilitarian and recreational walking

Walking bouts were assumed to be of a single type, either recreational or utilitarian. A utilitarian walking bout was defined as a transportation activity performed to get to a destination, which was a different place from the origin of the trip. A recreational walking bout was defined as a non-transportation activity which started and ended at the same place, with no specific place as a destination. For example, if a participant had a walking bout getting around her neighborhood and ending at the starting location, it was classified as a recreational bout. However, if she stopped by at a coffee shop and had a break > 2 min, she would have two utilitarian walking bouts before and after the coffee shop visit.

Recreational walking was identified via travel log information and GPS traces. Participants were instructed to record "tours" in travel diary when they started and ended a trip at the same location without making any stops in-between. Recorded tours therefore have the same definition as "recreational walking", and served to identify recreational walking bouts. The travel diary did not ask about a participant's purposes or intentions associated with trips. Tours were subjectively reported but determined by a revealed behavior outcome whether a trip returned to its origin or not. Walking bouts having a time overlap with a declared tour were classified as recreational (Figure 4-1).

GPS data were also used for the identification of recreational walking. Walking bouts whose first and last GPS points (the trip origin and destination) were within a distance ≤ 132ft (≈40m) of each other were defined as recreational. The cutoff of 132 ft was selected following an observational study that reported 95% of GPS points measured within a 132 ft-diameter circle using the same GPS device model could be deemed to be in the same location (Figure 4-2).[23] Under this definition, utilitarian walking bouts could be falsely classified as recreational if they had a small number of GPS points. Thus, recreation walking bouts classified by GPS data also had to have ≥ 10 GPS points.

All non-recreational walking bouts were classified as utilitarian.

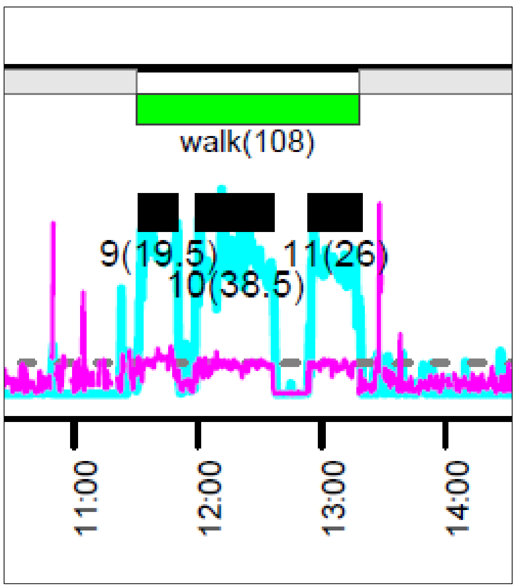

Figure 4-1. Three walking bouts (in black) overlapping with a diary-based tour (with a travel mode of walking) were classified as recreational.

Figure 4-2. A walking bout had GPS points whose first and last points were very close to each other. It was classified as recreational.

PA bout measurements

The person-level analysis used daily average PA bout measurements aggregated from PA bouts based on two steps. First, time spent in PA bouts was aggregated at the person-day-level. Second, they were aggregated between person-days within the same participants. The aggregated measures were TPD, OPD, WPD, WPD-U, and WPD-R by types, described as:

- Total PA per day (TPD): total time spent in any PA bouts per day
- Other types of PA per day (OPD): total time spent in non-walking bouts per day
- Walking per day (WPD): total time spent in walking bouts per day
- Utilitarian walking per day (WPD-U): total time spent in utilitarian walking bouts per day
- Recreational walking per day (WPD-R): total time spent in recreational walking bout per day

PA perceptions

The relationships between the environment and PA may be moderated by psycho-social factors regarding PA.[18, 19] For example, a person who is convinced of the health benefits of PA may be more active than a person who is not, regardless of their sociodemographic characteristics or of the environmental characteristics of their place of residence. Five PA perception variables (self-confidence, enjoyment, benefits, barriers, and social-support) were included in the analysis. Data processing to create these variables were described in detail in Chapter 3. PA perception variables were quantified into a range between 1 and 5. Higher values represent higher perception.

Subjective neighborhood perception

Perceived neighborhood characteristics have been shown significant in explaining PA.[1, 19, 20] The current analysis included common significant variables: quality of walking infrastructure[1, 11], neighborhood attractiveness[24], and safety (traffic and crime).[11, 20] These variables were obtained from self-report survey data. The questions were answered on a Likert scale (1=strongly disagree; 2=somewhat disagree; 3=somewhat agree; and 4=strongly agree). Numerical codes were averaged by survey section to create five neighborhood perception variables: walking infrastructure, attractiveness, traffic risk, and crime risk, all ranging from 1 to 4. Higher values represent better walking infrastructure, more attractive neighborhood, and higher risk of traffic accidents and crime incidents.

Table 4-1. Neighborhood perception variables

Neighborhood perception survey section	**Sub-questions**
(1) Walking infrastructure: 5 questions	1. There are sidewalks on most of the streets in my neighborhood. 2. Sidewalks are separated from the road/traffic in my neighborhood by parked cars. 3. There is a grass/dirt strip that separates the streets from the sidewalks in my neighborhood. 4. My neighborhood streets are well lit at night. 5. There are crosswalks and pedestrian signals to help walkers cross busy streets in my neighborhood.
(2) (Neighborhood) attractiveness: 3 questions	1. There are many interesting things to look at while walking in my neighborhood. 2. There are many attractive natural sights in my neighborhood (such as landscaping, views). 3. There are attractive buildings/homes in my neighborhood.
(3) Traffic risk: 3 questions	1. There is so much traffic along nearby streets that it makes it difficult or unpleasant to walk in my neighborhood. 2. The speed of traffic on most nearby streets is usually slow (30 mph or less). *[* inversely coded]* 3. Most drivers exceed the posted speed limits while driving in my neighborhood.
(4) Crime risk: 3 questions	1. There is a high crime rate in my neighborhood. 2. The crime rate in my neighborhood makes it unsafe to go on walks during the day. 3. The crime rate in my neighborhood makes it unsafe to go on walks at night.

Objective built environments

Environmental characteristics at home neighborhoods were included in the analyses . Variables describing environmental characteristics were selected from previous relevant studies. First, residential and job density measures were included as leading variables describing neighborhood composition and variables affecting PA.[3, 4, 12, 25, 26] For example, one study found that the modal share of walking for shopping trips increased by about 20% as resident/population density increased from a range of 13.1–18 residents per acre to a higher range of 18.1–60.[25] Property value was included to represent neighborhood wealth which may affect PA and health-related behaviors.[27, 28] Street intersections ("intersection")[4, 29], sidewalks[4, 30], slope[7, 30], and clustered neighborhood destinations ("destinations")[17] were considered as they were shown to relate to utilitarian walking. Trails[31, 32], parks[33, 34], and fitness facilities[8] were also included because they have been linked to recreational walking.

Home neighborhoods were defined as areas within a 833 m-radius (10-min walking distance).[16] Variables were calculated as point or line density, count, and area values within home neighborhoods. I used SmartMap data sets developed by the University of Washington's Urban Form Lab[35] except for clustered neighborhood destinations, which were created from convex hulls of a set of destination points. A convex hull included +1 grocery stores, +1 supermarkets, and +1 traditional restaurants within 50 m of each other.[17] Table 4-2 lists the original data sources for these variables.

Home addresses were geocoded by ArcGIS 9.3.1 (ESRI, 1999) using the King County, WA address-point GIS layer.

Table 4-2. Built environment variables, measures, and data sources

	BE measures	Unit	Data source and processing
1	Residential density	Units per acre	2010 KC assessor's data processed by UFL
2	Job density	Jobs per acre	2007 KC assessor's data and 2006 NAIC data processed by UFL
3	Property value	Average residential unit price	2007 KC assessor's data processed by UFL
4	Slope	Average degree	2010 USGS DEM data
5	Sidewalk line density	Mile per acre	Data collected from jurisdictions and processed by UFL in 2012
6	Street intersection density	Count per acre	2006 KC transportation network data processed by UFL
7	Park areas	% within the 833 m buffer	2007 KC assessor's data processed by UFL
8	Trail line density	Mile per acre	2010 KC transportation network data processed by UFL
9	Fitness facility	Count	2008 InfoUSA dataprocessed by UFL
10	Neighborhood destinations	Acre	2008 Seattle and KC public health data processed by UFL

Analysis

Two sets of multivariate OLS regression models were estimated. Each set included 5 models whose response variables were TPD, OPD, WPD, WPD-U, and WPD-R. The first set of models, called “non-environmental models (Model NE)”, were developed with the sociodemographic and psycho-social variables only. The second set of models, called “environmental models (Model E)”, added the subjective neighborhood perception and objective built environmental variables. The first and the second model sets were compared to test how the model fits of models predicting different types of walking were improved by adding the environmental variables. Then, the environmental models were analyzed to compare differences between utilitarian and recreational walking,

Results

Participants

The final sample consisted of 657 participants' 4,283 person-days. Table 4-3 summarizes demographic information, PA perceptions, and environmental characteristics of the sample. The mean age was 50.8 years, 80.5% were non-Hispanic white, 71.3% college graduate or above, and 54% were full-time employees. The household income makeup was 37.9% <49K USD, 41% between 50 and 99K USD, and 21.1% ≥100K USD. The sample perceived higher than neutral levels for PA self-confidence, PA enjoyment and benefits of PA, and lower than neutral levels for barriers of PA and social support. Participants' mean residential unit density in the home neighborhood was 20.5 du per acre.

Table 4-3. Demographic, psycho-social, neighborhood perception and environmental measures

N=657		**Mean, N**	**SD, %***
Age		50.8	13.0
Sex (Female)		412	63.3
Ethnicity (non-Hispanic white)		521	80.5
Household income	<49K USD	240	37.9
	50-99K USD	260	41.0
	≥100K USD	134	21.1
Education (college graduate or above)		461	71.3
Employment (full-time employer)		341	54.0
PA perception	Self-confidence	3.92	0.93
Unit: [1,5]	Enjoy	4.28	0.84
	Benefit	4.21	0.61
	Barrier	2.20	0.68
	Social-support	2.36	0.91
Neighborhood perception	Walking infrastructure	3.17	0.58
Unit: [1,4]	Attractiveness	3.13	0.73
	Traffic risk	2.23	0.66
	Crime risk	1.86	0.66
Built environment	Residential density (du/acre)	20.5	14.4
	Job density (jobs/acre)	34.2	63.1
	Property value (USD 1,000/unit)	247	82
	Slope (degree)	4.3	1.3
	Sidewalk line density (mi/acre)	28.8	15.5
	Street intersection density (count/acre)	215.6	75.7
	Park area (%)	5.0	5.3
	Trail line density (mile/acre)	0.2	0.5
	Fitness facility (count)	5.7	5.0
	Neighborhood destinations (acre)	32.2	58.5

* NA values are excluded in calculating %.

Classification

In total, 12,561 PA bouts were identified from the sample. Nearly 60% of the bouts were walking, of which 89% had utilitarian purposes. The recreational walking bouts, 11% of the walking bouts, were identified as follows: 42% came from overlaps with diary-derived tours; 40% came from GPS data; and 18% were from both diary and GPS data (Figure 4-3).

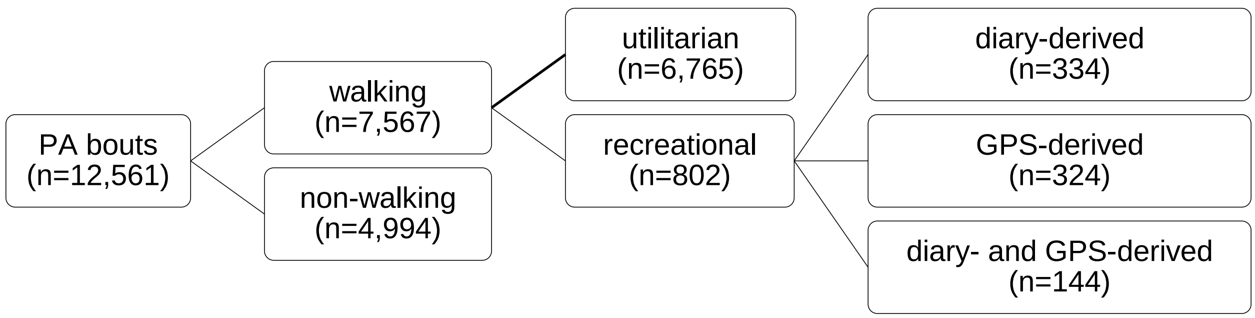

Figure 4-3. Classification of bouts

Characteristics of utilitarian versus recreational walking at the bout level

Table 4-4 provides descriptive statistics of walking bouts by types. The last column of the table presents p-values of ANOVA tests between utilitarian versus recreational walking. On average, walking bouts lasted longer (+2.1 min) and had 135% as high as mean accelerometer counts (+362.7 cpe) than non-walking bouts. Between utilitarian and recreational walking bouts, there were significant differences in duration, GPS speed, and GPS coverage ratio ($p<0.000$). GPS coverage was defined as ratio of time duration covered by GPS data over the total duration of a bout. On average, recreational walking bouts were 91% longer (+12.6 min), slightly slower (-0.3 kmh-1) and covered with more GPS data (+0.22) than utilitarian walking bouts. The standard deviation of recreational walking bout duration was almost twice that of utilitarian walking bouts (20.3 > 11.0 min).

Table 4-4. Bout level comparison

	All bouts		Non-walking		Walking		Utilitarian walking		Recreational walking		U vs. R p-value
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	
N (all)	12,561	-	4,994	-	7,567	-	6,765	-	802	-	NA
Duration (min)	14.4	12.3	13.1	11.2	15.2	12.9	13.9	11.0	26.5	20.3	<0.000
Counts (per min)	1,300.5	559.3	1,082.0	591.4	1,444.7	486.0	1,442.1	483.6	1,466.2	505.1	0.184
N (w/ GPS)	9,224	-	3,566	-	5,658	-	4,935	-	723	-	NA
GPS coverage (ratio)	0.62	0.43	0.64	0.45	0.61	0.41	0.58	0.42	0.80	0.32	<0.000
Speed (kmh^{-1})	3.0	2.4	2.0	3.4	3.6	1.2	3.7	1.3	3.4	0.9	<0.000

Person-level distributions of PA measurements

Time spent in PA bouts was aggregated at the person-level. The aggregated measures were TPD, OPD, WPD, WPD-U, and WPD-R by types. They are summarized in Table 4-5. On average, participants had 26.6 min of WPD, 21.7 min of WPD-U, and 4.8 min of WPD-R.

Of the sample (n=657), 38 participants had no walking bouts at all during their entire assessment period (WPD = 0). Those who had a least one walking bout per day (who had WPD >0; n=619) had 28.2 min of WPD, 23.1 min of WPD-U, and 5.1 min of WPD-R. For the population with at least one walking bout, the person-level mean of the contribution of utilitarian walking to overall walking (WPD-U over WPD) was 81.9%. Participants with at least one recreational walking bout (WPD-R>0; n=281) had 11.3 min of WPD-R.

Table 4-5. Person level distribution of PA measurements

	TPD	OPD	WPD	WPD-U	WPD-R
Minimum	1.0	0.0	0.0	0.0	0.0
Maximum	277.3	170.2	163.9	159.4	70.1
Median	34.6	9.8	19.8	15.3	0.0
Mean	42.1	15.5	26.6	21.7	4.8
SD	31.7	18.1	23.8	22.0	9.4

* Unit = min per day

Analysis

Pearson's correlation coefficients were calculated within the PA perception variable group and within the built environment variable group (subjective and objective). Most highly correlated was the pair of PA barriers and self-confidence (ρ=-0.58). Because of the relatively low correlation coefficients, all PA perception variables were included in the analyses. Within the built environment variable group, 11 of the 91 possible pairs had a correlation coefficient > 0.65 (Table 4-6). For example, the fitness facility variable was highly correlated with intersection density and with sidewalk density, which was also correlated with intersection density. Four variables (sidewalk density, intersection density, fitness facility, and destinations) were then dropped so that none of the remaining pairs had high correlations. . Residential density and job density were kept because they are important variables in describing environments and the correlation between them was not large (ρ=0.62).

Table 4-6. Pearson's correlation among BE variables (only objective variables shown, all pairs with subjective variables have ρ < 0.48)

		Res density	Fitness	Job density	Park	Inter-section	Side-walk	Trail	Pro-perty	Desti-nation	Slop e
1	Res density	**-**	-	-	-	-	-	-	-	-	-
2	Fitness	**0.80**	**-**	-	-	-	-	-	-	-	-
3	Job density	0.62	**0.77**	**-**	-	-	-	-	-	-	-
4	Park	-0.28	-0.27	-0.28	**-**	-	-	-	-	-	-
5	Intersection	**0.79**	**0.78**	**0.74**	-0.18	**-**	-	-	-	-	-
6	Sidewalk	**0.76**	**0.70**	0.46	-0.11	**0.84**	**-**	-	-	-	-
7	Trail	-0.21	-0.08	0.01	0.21	-0.19	-0.27	**-**	-	-	-
8	Property	-0.29	-0.25	-0.15	0.36	-0.15	-0.04	0.19	**-**	-	-
9	Destinations	0.63	**0.78**	**0.95**	-0.30	**0.71**	0.47	0.07	-0.13	**-**	-
10	Slope	-0.12	-0.20	-0.09	0.36	0.01	-0.01	-0.03	0.31	-0.18	**-**

* Correlation values > 0.65 are shown in bold

Two sets of regression models were compared: non-environmental models (Table 4-7) and environmental models (Table 4-8). Columns in the tables show regression coefficients for the different response variables of TPD, OPD, WPD, WPD-U, and WPD-R. Based on AIC and BIC scores, adding environmental variables improved model fits of WPD (AIC decreased by 45; BIC by 1) and WPD-U (AIC decreased by 79; BIC by 35). Models of other response variables (TPD, OPD, WPD-R) had increased BIC scores with adding environmental variables Based on BIC scores, only the fit of WPD-U model decreased substantially.

I selected the significance level as p-value < 0.05. Adding environmental variables did not change significance levels or the direction of influence in the sociodemographic and PA perception variables. All significant variables in the first set of models were also significant in the second set except for sex in the WPD models and highest income in the WPD-R models. All of these significant variables had the same direction of influence in the first and the second sets. In the second set (Model E), the TPD model showed that being female, being older, and having higher PA barriers were negatively associated with higher TPD. Having higher PA self-confidence, living in a higher residential density area, and living in a higher property value area were positively associated with higher TPD. In the OPD model, the association patterns were the same as in the WPD model except for residential density, which was not significant in the OPD model. Magnitudes of coefficients decreased in the OPD model. In addition, non full-time workers had 3.2 fewer min of OPD than full-time workers. The WPD model showed that being older, having higher PA barriers and higher social support were negatively associated with WPD; and having higher PA self-confidence, living in higher residential density and job density areas were positively associated with WPD.

To examine the differences between utilitarian and recreational walking, the WPD-U model was compared to the WPD-R model. The WPD-U model had the same association patterns as those in the WPD model and one more significant variable, sex. Females had daily 5.1 fewer min of utilitarian walking than males. Among neighborhood perception variables, only neighborhood attractiveness was shown to be significant in predicting WPD-U but not in predicting WPD-R. The WPD-R model had only two significant variables. The sign of the coefficients had opposite directions to those of the WPD-U model. Females and those living in a lower job density area were positively associated with WPD-R. In contrast, they were negatively associated with WPD-U.

Table 4-7. Results of non-environmental models (Model NE)

		TPD			OPD			WPD			WPD-U			WPD-R		
		Est	SE		Est	SE		Est	SE		Est	SE		Est	SE	
(Intercept)		52.63	15.11	***	9.25	8.93		43.38	11.40	***	42.43	10.47	***	0.95	4.61	
Sex	Female	**-7.81**	**2.59**	**	-2.96	1.53	.	-4.85	1.96	*	**-7.02**	**1.80**	***	**2.17**	**0.79**	**
Age		**-0.36**	**0.10**	***	**-0.14**	**0.06**	*	**-0.22**	**0.08**	**	**-0.25**	**0.07**	***	0.03	0.03	
Ethnicity	Non-Hispanic white	2.43	3.16		-0.07	1.87		2.51	2.38		0.63	2.19		1.88	0.96	.
Income	[50, 100K)	2.08	2.94		2.15	1.74		-0.07	2.22		0.21	2.04		-0.28	0.90	
(base: [0, 50K))	[100K,)	2.61	3.59		3.02	2.12		-0.41	2.71		-2.61	2.49		2.21	1.10	*
Employment	Full time	-5.17	2.72	.	**-3.36**	**1.61**	*	-1.81	2.05		-1.80	1.88		-0.01	0.83	
PA	Self-confidence	**4.15**	**1.72**	*	1.06	1.02		**3.09**	**1.30**	*	**2.87**	**1.19**	*	0.22	0.52	
perception	Enjoy	2.79	1.96		1.11	1.16		1.68	1.48		1.27	1.36		0.41	0.60	
	Benefit	0.92	2.45		1.54	1.45		-0.62	1.85		-0.22	1.70		-0.40	0.75	
	Barrier	**-6.64**	**2.39**	**	-0.69	1.41		**-5.95**	**1.80**	**	**-4.72**	**1.65**	**	-1.24	0.73	.
	Social-support	-2.19	1.48		0.85	0.88		**-3.04**	**1.12**	**	**-3.44**	**1.03**	***	0.40	0.45	
R2		0.103			0.046			0.093			0.105			0.057		
AIC score		5,917			5,274			5,572			5,468			4,464		
BIC score		5,974			5,331			5,630			5,526			4,522		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 4-8. Results of environmental models (Model E)

		TPD			OPD			WPD			WPD-U			WPD-R		
		Est	SE		Est	SE		Est	SE		Est	SE		Est	SE	
(Intercept)		17.72	18.61		6.17	11.02		11.55	13.60		7.94	12.15		3.61	5.72	
Sex	Female	**-6.41**	**2.64**	*	**-3.40**	**1.56**	*	-3.01	1.93		**-5.01**	**1.72**	**	**2.00**	**0.81**	*
Age		**-0.31**	**0.10**	**	**-0.14**	**0.06**	*	**-0.16**	**0.07**	*	**-0.19**	**0.07**	**	0.02	0.03	
Ethnicity	Non-Hispanic white	0.81	3.19		-0.74	1.89		1.55	2.33		-0.10	2.08		1.65	0.98	.
Income	[50, 100K)	2.21	2.95		1.54	1.75		0.67	2.16		1.06	1.93		-0.39	0.91	
(base: [0, 50K))	[100K,)	2.39	3.71		1.69	2.19		0.70	2.71		-1.03	2.42		1.73	1.14	
Employment	Full time	-3.32	2.73		**-3.20**	**1.62**	*	-0.12	2.00		0.01	1.78		-0.13	0.84	
PA	Self-confidence	**3.54**	**1.71**	*	0.89	1.01		**2.65**	**1.25**	*	**2.42**	**1.12**	*	0.23	0.53	
perception	Enjoy	2.93	1.96		1.37	1.16		1.55	1.44		1.10	1.28		0.45	0.60	
	Benefit	0.64	2.46		0.70	1.46		-0.05	1.80		0.40	1.61		-0.45	0.76	
	Barrier	**-5.91**	**2.40**	*	-1.07	1.42		**-4.84**	**1.75**	**	**-3.64**	**1.57**	*	-1.20	0.74	
	Social-support	-1.72	1.50		0.59	0.89		**-2.30**	**1.09**	*	**-2.50**	**0.98**	*	0.19	0.46	
Neighborhood	Walking infrastructure	1.72	2.40		0.78	1.42		0.94	1.75		0.88	1.56		0.06	0.74	
perception	Attractiveness	3.78	2.08	.	0.53	1.23		**3.25**	**1.52**	*	**3.10**	**1.36**	*	0.15	0.64	
	Traffic risk	2.43	2.14		1.83	1.27		0.60	1.56		0.75	1.40		-0.15	0.66	
	Crime risk	0.91	2.09		0.35	1.24		0.56	1.53		0.97	1.37		-0.41	0.64	
Built	Residential density	**0.28**	**0.12**	*	-0.08	0.07		**0.35**	**0.09**	***	**0.35**	**0.08**	***	0.00	0.04	
environment	Property value	0.03	0.02	.	**0.03**	**0.01**	**	0.01	0.01		0.00	0.01		0.00	0.01	
(objective)	Employment density	0.01	0.03		-0.02	0.02		**0.04**	**0.02**	*	**0.06**	**0.02**	**	**-0.02**	**0.01**	*
	Trail density	-0.78	2.83		-1.23	1.68		0.45	2.07		-0.60	1.85		1.05	0.87	
	Park percentage	-0.11	0.27		-0.30	0.16	.	0.19	0.20		0.13	0.18		0.06	0.08	
	Slope	-1.54	1.08		-0.84	0.79		-0.70	0.64		-0.54	0.70		-0.29	0.33	
R2		0.141			0.082			0.185			0.239			0.080		
AIC score		5,910			5,270			5,527			5,389			4,469		
BIC score		6,012			5,372			5,629			5,491			4,571		

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Discussion

This study provides clear evidence that the built environment plays a strong role in explaining utilitarian walking and general walking. The built environment domain may explain walking better than any other domain. Sociodemographic and psycho-social factors explained 10.5% of the total variance in utilitarian walking (Model NE: WPD-U). The amount of explained variance was increased by 12.5% by adding environment factors (Model E: WPD-U). AIC scores changed from 5,468 to 5,389 (-79) and BIC scores from 5,526 to 5,491 (-35). One study based on a large Australian adult population (n=1,773) found that the relative influence of the built environment on walking was almost equal to that of demographic and psycho-social factors using subjective PA measures.[36] Results from these analyses apparently suggest that the built environment plays an important role in predicting utilitarian walking more so than they did recreation walking.

Recreational walking was not well predicted by environmental variables. Adding environmental variables increased AIC and BIC scores, thus worsening the model fit. Although employment density was significant, its coefficient was quite small. An increase of a 100 jobs within a 10-min distance from home locations was associated with a 2-min decrease in WPD-R. These results are consistent with a previous study showing that environments were not significant in predicting recreational walking.[13]

Recreational walking behaviors were different from utilitarian walking, particularly in their association with sex and job density. Results suggest that environmental effects on walking activity were modified by types of walking. One study using travel diary data from Northern California (n=1,553) compared walking frequency obtained from survey data in traditional versus suburban neighborhoods, and found that neighborhood effects accounted for 61% of the difference in utilitarian walking and 86% of the difference in recreational walking after controlling for residential self-selection.[9] In this study, the built environment accounted for more variance in utilitarian walking than in recreational walking. Different measures of walking may explain the different results from the present study: in the Northern California study, utilitarian walking was captured as ‘walking-to-store’ and recreational walking as ‘strolling’. However, both studies detected the presence of an effect modification based on walking activity purposes and suggested that specifying walking purpose was needed to precisely assess the link between walking behaviors and the built environment. In addition, coefficients of traffic and crime risk for WPD-R were negative although they were not significant (Table 4-8). In contrast, the coefficients for WPD-U were positive.

Results from the environmental models predicting utilitarian and recreational walking were not always consistent with previous studies. Key findings from prior studies selected from recent literature reviews[4, 12] are summarized in Table 4-9, and compared to the present study’s results. Job density was consistently positively associated with utilitarian walking. In contrast, residential density was positively associated with utilitarian walking in the present analyses while two other studies found negative or null associations. On the other hand, job density was negatively associated with recreational walking. No other study seemed to have directly investigated job density with respect to recreational walking. It is noteworthy that all studies in the reviews used subjective measures of walking (e.g., travel diary), which might explain the inconsistencies with the present analyses.

Job density was highly correlated with intersection density (ρ=0.74), destinations (ρ=0.95), and fitness facility (ρ=0.77) (Table 4-6), which were dropped from the final models. Replacing job density with these variables also showed significant associations with both utilitarian walking and recreational walking with only one exception: fitness facility was not significantly associated with recreational walking. Therefore, it appears that intersection density, destination, and fitness facility were positively associated with utilitarian walking. And intersection density and destination were negatively associated with recreational walking when these variables were used instead of job density.

Table 4-9. Comparison of the analysis with previous studies

	Utilitarian walking			**Recreational walking**		
	Result	Consistent	Inconsistent	Result	Consistent	Inconsistent
Residential density	+	Frank et al.[25] Zhang[37]	Rodriguez et al.[30] (-) Lee et al.[7] (0)	0	Lee et al.[7] (0)	
Property value	0	NA		0	NA	
Job density	+	Frank et al.[25] Zhang[37]		-		
Trail	0	Lee et al.[7] (0)		0	Lee et al.[7] (0)	Duncan et al.[33] (+)
Park	0	Lee et al.[7] (0) Plaut[38] (0)		0	Lee et al.[7] (0)	
Slope	0		Rodriguez et al.[30] (-) Lee et al.[7] (-)	0		Lee et al.[7] (+)

+ = positive association; - = negative association; 0 = insignificant association

Built environment measures were taken at the home neighborhood. This measure of exposure needs to be further tested as walking bouts also occur away from the residential neighborhood, in environments that may have different characteristics and may also influence walking behaviors. Further investigation is required to consider actual location of walking bouts. However, about 26% of all walking bouts did not have GPS data at all and their locations could not be identified.

PA behaviors could be further understood with reading Table 4-8 across models. For example, females had 6.66 min less TPD than males (the second column). This effect could be an additive result from the -3.44 min difference in OPD (the third column) and the -3.22 min difference in WPD (the forth column) between males and females. Again, the difference of -3.22 min in WPD came from - 5.12 min difference in WPD-U (the fifth column) and +1.9 min difference in WPD-R (the sixth column). Similarly, residential density had a strong positive influence on WPD-U. The influence was not changed in WPD because residential density had almost null influence on WPD-R. Finally, the influence was slightly reduced in TPD because residential density might have some negative influence in OPD. This approach provides a hierarchical way of explaining the additive effects on PA by different types.

This study defined recreational walking from an activity outcome (whether walking bouts had destinations or not) without considering the actual intention behind the activity. This definition avoided complications which would have arisen had individual participants been asked to define the purpose of their walk. For example, different participants may consider a walking trip to a coffee shop as recreational or utilitarian. Also, the decision to define walking with a single rather than multiple purposes (e.g., combined utilitarian and recreational), while limiting the full understanding of behavior, allowed for determining bout’s purpose based on an objectively measurable outcome.

Most importantly, PA was measured objectively in this study. Although the measurements had issues related to data incompleteness (e.g., missing GPS data), they were less biased by participants’ errors than survey derived measures. Based on more objective and accurate objective data, the present study suggests (1) utilitarian walking is primarily explained by environmental factors; and (2) job density and sex have different influences on utilitarian and recreational walking. Therefore, environmental studies on walking should specify walking types.

Conclusion

Utilitarian and recreational walking can be differentiated by using travel diary ("tour") and GPS track data. Built environment factors were useful in predicting utilitarian walking but not recreational walking. Correlates of utilitarian walking were very different from those of recreational walking. Being males and employment density were positively associated with utilitarian walking but negatively with recreational walking. Self-confidence in PA and neighborhood attractiveness were positively associated and age, perceived barriers of PA, social support for PA were negatively associated with utilitarian walking. Only sex and employment density were significant for recreational walking.

Comparing the results across models suggested that the activity of walking for utilitarian purposes drove the identification of individual demographic, perception, and environmental variables also associated with total PA.

Future research should examine actual locations where walking activities occurred. The current study assumed spatial exposure at home determined walking behaviors. But people may walk for recreation away from home where they can enjoy walking activity itself. Spatially matched behavior and environmental characteristics would provide a better understanding of walking behaviors.

References

1. Saelens BE, Sallis JF, Frank LD. Environmental correlates of walking and cycling: findings from the transportation, urban design, and planning literatures. Annals of Behavioral Medicine 2003;25(2):80-91.
2. Handy SL, Boarnet MG, Ewing R, Killingsworth RE. How the built environment affects physical activity. American journal of preventive medicine 2002;23(2S):64-73.
3. Ewing R, Cervero R. Travel and the built environment: a synthesis. Transportation Research Record: Journal of the Transportation Research Board 2001;1780(-1):87-114.
4. Ewing R, Cervero R. Travel and the built environment. Journal of the American Planning Association 2010;76(3):265-294.
5. Cervero R, Kockelman K. Travel demand and the 3Ds: density, diversity, and design. Transportation Research Part D: Transport and Environment 1997;2(3):199-219.
6. Morris J, Hardman A. Walking to health. Sports Medicine 1997;23(5):306.
7. Lee C, Moudon AV. Correlates of walking for transportation or recreation purposes. Journal of Physical Activity and Health 2006;3:77.
8. Hoehner CM, Brennan Ramirez LK, Elliott MB, Handy SL, Brownson RC. Perceived and objective environmental measures and physical activity among urban adults. American journal of preventive medicine 2005;28(2):105-116.
9. Cao X. Exploring causal effects of neighborhood type on walking behavior using stratification on the propensity score. Environment and Planning A 2010;42(2):487-504.
10. Troped PJ, Saunders RP, Pate RR, Reininger B, Addy CL. Correlates of recreational and transportation physical activity among adults in a New England community. Preventive Medicine 2003;37(4):304-310.
11. Pikora TJ, Giles-Corti B, Knuiman MW, Bull FC, Jamrozik K, Donovan RJ. Neighborhood environmental factors correlated with walking near home: using SPACES. Medicine and Science in Sports and Exercise 2006;38(4):708.
12. Saelens BE, Handy SL. Built environment correlates of walking: a review. Medicine and Science in Sports and Exercise 2008;40(7 Suppl):S550.
13. Lovasi GS, Moudon AV, Pearson AL, Hurvitz PM, Larson EB, Siscovick DS, et al. Using built environment characteristics to predict walking for exercise. International Journal of Health Geographics 2008;7(1):10.

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(1).
15. 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.
16. Hurvitz PM, Moudon AV. Home versus nonhome neighborhood: quantifying eifferences in exposure to the built environment. American Journal of Preventive Medicine 2012;42(4):411-417.
17.Lee C, Moudon AV. The 3Ds+ R: Quantifying land use and urban form correlates of walking. Transportation Research Part D: Transport and Environment 2006;11(3):204-215.
18. Moudon AV, Lee C, Cheadle AD, Garvin C, Johnson D, Schmid TL, et al. Operational definitions of walkable neighborhood: theoretical and empirical insights. Journal of Physical Activity and Health 2006;3:99.
19. Heath GW, Brownson RC, Kruger J, Miles R, Powell KE, Ramsey LT. The effectiveness of urban design and land use and transport policies and practices to increase physical activity: a systematic review. Journal of Physical Activity and Health 2006;3:55.
20. McCormack G, Giles-Corti B, Lange A, Smith T, Martin K, Pikora T. An update of recent evidence of the relationship between objective and self-report measures of the physical environment and physical activity behaviours. Journal of Science and Medicine in Sport 2004;7(1):81-92.
21. Moudon AV, Saelens BE, Rutherford S, Hallenbeck M. A report on participant sampling and recruitment for travel and physical activity data collection. [Seattle, Wash.]: Transportation Northwest; 2009. Report No.: TNW2009-05.
22. Federal Highway Administration. 2001 National Household Travel Survey: user's guide. Washington, DC: US Department of Transportation; 2004.
23. 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.
24. Michael YL, Green MK, Farquhar SA. Neighborhood design and active aging. Health & Place 2006;12(4):734-740.

25. Frank LD, Pivo G. Impacts of mixed use and density on utilization of three modes of travel: single-occupant vehicle, transit, and walking. Transportation research record 1994:44-44.
26. Frank L, Bradley M, Kavage S, Chapman J, Lawton TK. Urban form, travel time, and cost relationships with tour complexity and mode choice. Transportation 2008;35(1):37-54.
27. Gordon-Larsen P, Nelson MC, Beam K. Associations among Active Transportation, Physical Activity, and Weight Status in Young Adults. Obesity 2005;13(5):868-875.
28. Moudon AV, Cook AJ, Ulmer J, Hurvitz PM, Drewnowski A. A neighborhood wealth metric for use in health studies. American Journal of Preventive Medicine 2011;41(1):88-97.
29. Frank LD, Saelens BE, Powell KE, Chapman JE. Stepping towards causation: Do built environments or neighborhood and travel preferences explain physical activity, driving, and obesity? Social Science and Medicine 2007;65(9):1898-1914.
30. Rodriguez DA, Joo J. The relationship between non-motorized mode choice and the local physical environment. Transportation Research Part D: Transport and Environment 2004;9(2):151-173.
31. 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.
32. Fitzhugh EC, Bassett DR, Evans MF. Urban trails and physical activity: A natural experiment. American Journal of Preventive Medicine 2010;39(3):259-262.
33. Duncan M, Mummery K. Psychosocial and environmental factors associated with physical activity among city dwellers in regional Queensland. Preventive Medicine 2005;40(4):363-372.
34. Kaczynski AT, Potwarka LR, Saelens BE. Association of park size, distance, and features with physical activity in neighborhood parks. American Journal of Public Health 2008;98(8):1451.
35. Hurvitz PM, Jiao J, Ulmer J, Kang B, Huang R, Moudon AV. Space, time, and movement: Integrating high-resolution GPS, GIS, and accelerometry for estimating built environment exposure. In: 2011 Association of American Geographers Annual Meeting. Seattle, WA; 2011.
36. Giles-Corti B, Donovan RJ. The relative influence of individual, social and physical environment determinants of physical activity. Social Science and Medicine 2002;54(12):1793-1812.

37. Zhang M. The role of land use in travel mode choice: evidence from Boston and Hong Kong. Journal of the American Planning Association 2004;70(3):344-360.

38. Plaut PO. Non-motorized commuting in the US. Transportation Research Part D: Transport and Environment 2005;10(5):347-356.

Chapter 5 Dissertation Conclusion

This dissertation analyzed physical activity (PA) and walking behaviors using objective data from accelerometers, GPS units, and travel logs. PA and walking were precisely measured and objectively classified. Key findings include:

- GPS and travel diary information can be helpful in classifying most accelerometer-derived PA bouts into walking or non-walking behavior. While accelerometers provide reliably continuous data at small temporal increments, GPS data are typically discontinuous in time and sometimes inaccurate in estimates of speed. Self-reported travel logs are therefore needed to determine start and end points for accelerometer data; to provide information on temporal breaks in GPS data; and to distinguish between and verify movement and stationary activity in time and space.
- Increased walking contributed to total PA without reducing other types of PA. This suggested that no substitutive behaviors existed between walking and other types of PA, and that more walking would correspond to a net increase in overall PA.
- Utilitarian and recreational walking can be determined more objectively by travel diary and GPS data. Environmental variables strongly predicted utilitarian walking but not recreational walking. Correlates of utilitarian walking seem to also apply to overall PA.

One of the most important contributions of the dissertation lies in the use of objective measures of PA and the development of methods to distinguish between different types of PA and specifically to isolate the elements of PA that correspond to walking. Given the explicit measurement of PA in time and space, other researchers will be able to replicate the methods and produce comparable PA measures. As past research produced inconsistent results likely caused by the use of subjective and incomparable PA and walking measures, the data and methods used in this dissertation can serve to guide future explorations of the relationships between different types of PA and between PA and environment. The findings of the present research can be a reference for future studies.

More specifically, the findings have important repercussions for both future research and interventions in PA and health. Chapter 3 concluded that walking contributed to total PA without reducing other types of PA. This meant that public health policy promoting walking would be effective in promoting overall PA. Thus if, for example, interventions to promote PA were targeted to moderate walkers, their effect could help a large proportion of the population reach recommended levels of PA. Chapter 4 suggested that environmental interventions might be effective in promoting utilitarian walking but not recreational walking. Thus environmental interventions should be focused in areas that support utilitarian walking.

Future research will be needed in the following areas:

(1) Further comparisons between subjectively and objectively obtained walking estimates will help the development of survey or other self-report instruments that yield more accurate information on PA and walking. Such instruments are needed to continue to periodically monitor PA levels at a population level;

(2) Improvement of the walking identification algorithm using statistical approach (multivariate classification) based on bouts' characteristics. A more parsimonious algorithm will be easier to use in multiple applications;

(3) Development of a research framework to understand the allocation of walking to discretionary and obligatory activities. This will help refine public health interventions promoting more walking by focusing on individual PA time budgets and their corresponding elasticities;

(4) Micro-scale spatial analyses of PA using GPS location data (e.g., recreational walking and its actual locations). The present analyses took into account the influences of the home environment on PA and walking, but did not examine the likely effects of environments directly associated with the activities. The finding that the home environment has little if any influence on recreational walking suggests that further work is needed on where people walk for recreation and on the characteristics of environments at locations where recreational walking takes place;

(5) Analysis on the relationship between person-level spatial exposures derived from GPS data and health behaviors and/or outcomes (i.e., PA, weight status). Continuous GPS tracking enables to measure people's continuous temporal and spatial exposure to environments. Adequate concepts and methodologies are needed to describe and summarize a person's continuous environmental exposure; and

(6) Temporal variance analysis in PA. There is a large temporal variability in PA (e.g., day-level, week-level). Most studies have used PA measures obtained from a 7-day monitoring. If there are large variances in PA across weeks, longer observation may be required. If there are some consistencies within a week, researchers may need to monitor during the time < a week. Also, day-level variability would be interesting. Public health researchers may identify sedentary time windows and may be able to develop interventions to promote PA during those time windows.

Overall, the rich dataset combining data from accelerometers GPS units, and travel diary which was available from the TRAC study, serves as an example of the types of data needed to answer many of the questions that remain to be addressed in order to reverse trends toward physical inactivity.