Assessing the Built Environment using GPS, Physical Activity Monitors and Geospatial Surveys

Seeger, C. J.,[1] Welk, G. J.[2], and Erickson, M. S.,[1]
[1]Iowa State University, Dept. of Landscape Architecture, 146 College of Design, Ames, Iowa 50011, USA
E-mail: cjseeger@iastate.edu, susaneri@iastate.edu
[2]Iowa State University, Department of Kinesiology, 257 Forker Building, Ames, Iowa 50011, USA
E-mail: gwelk@iastate.edu

Abstract

Understanding what aspects of the built environment contribute to health concerns such as obesity is not an easy task, due to the scale of the built environment and inconsistencies in community design, climate and culture. It is therefore necessary to look at the detailed layers of the built environment to better understand what components may have the largest impact. This can be a daunting task. Through the use of participatory research methods, geospatial tools and physical activity recording devices, it is possible to create models of how people interact with their environment. This paper discusses how geospatial technologies along with physical activity monitors have been used to determine if the urban forest, a component of the built environment, influences an individual's decision to use a particular walking/running route in Ames, Iowa, USA.

1. Introduction

The Centers for Disease Control and Prevention (CDC) reports that approximately 66 percent of the U.S. adult population is either overweight (Body Mass Index of 25–29.9) or obese (BMI >= 30) (CDC, 2007a). According to data from a National Health and Nutrition Examination Survey (NHANES), these percentages are approximately twice the amount reported in health surveys taken in the mid-1970s (CDC, 2007b). Many factors have been implicated as contributing to this epidemic, but low levels of physical activity are viewed as perhaps the most important modifiable factors. The CDC's Behavioral Risk Factor Surveillance System (BRFSS) found that in 2005 the national average of individuals reporting an insufficient amount of physical activity was 37.7 percent and that 14.2 percent reported they were completely inactive (CDC, 2007c). There has been considerable interest among public health researchers in determining factors that may explain the low levels of physical activity in the population. Social-ecological models are the most commonly used framework for this research since they take into account social and environmental factors that may influence physical activity behaviors (Sokols and Bellingham, 1996). The use of these models has led to a flurry of research on possible associations between elements of the built environment and levels of physical activity and obesity in the population. A summary report from the 2004 conference "Obesity and the Built Environment: Improving Public Health through Community Design," in Washington, DC, concluded that the "rapid increase in obesity over the past 30 years strongly suggests that environmental influences are responsible for this trend." A subsequent report published by the Transportation Research Board in January 2005 reviewed additional research and concluded that there is "available empirical evidence" linking the built environment with a person's physical activity (Transportation Research Board, 2005). The report recommended that additional studies are needed to explore the "causal relationship between the built environment and physical activity." A specific recommendation was made for studies to examine "residential location preferences, and characteristics of the built environment as determinants of physical activity" (Transportation Research Board, 2005). Research into these topics necessitates accurate measures of physical activity as well as detailed geospatial information. Spatial analysis and data collection tools such as geographic information systems (GIS) and global positioning systems (GPS) are well suited for research on the built environment. These tools can provide an accurate map with which proximity, spatial distribution and connectedness of individual elements within the built environment can be measured and related. Accelerometry-based activity monitors, on the other hand, provide a way to collect objective data on physical activity under free living conditions. These physical monitors are small (about the size of a

ISSN 1686-6576 / © Geoinformatics International

pedometer), non-invasive monitors that record information about the amount of movement taking place on a minute-by-minute basis. By combining information from GPS and activity monitors it is possible to directly examine environmental factors influencing physical activity behavior. This paper describes the process used to merge these various types of data and demonstrates how the resulting dataset can be visualized and used to examine the relationship.

2. Project Background

Funded by the National Urban and Community Forestry Advisory Council, Iowa State University Extension studied the role vegetation plays on an individual's choice in selecting the use of specific community recreation trails with data collected in Ames, Iowa, USA between 2005 and 2007. The study looked at the following questions.

1) Is the use of a trail impacted by adjacent vegetation?
2) Is variety in vegetation an influencing factor of route selection?
3) Is there a relationship between trail selection and weather conditions?
4) Do physical activity rates (exertion) correspond directly to the type of trail surface, length or presence of vegetation?

Information for the study was collected from 48 Ames residents chosen from a pool of 500 people who volunteered to participate in the study. Ames is a rural university community with a population just over 50,000. The residents identified themselves as physically active adults that use community trails on at least three occasions throughout the week. The participants were selected based on gender, location of residence and age. They were divided into three age groups: 18-30, 30-55 and 55+. The data collection portion of the study took place over the course of a year. During that time there were four one-week data collection periods, one each in November, January, April and August. During the one-week periods, participants were asked to wear a GPS device on their wrist when they were actively walking or running. They were also asked to wear a physical activity monitor (attached to their waistband) throughout the day in order to monitor physical activity. The activity monitor was programmed to record data on a minute-by-minute basis while the GPS recorded its position every ten seconds. In addition, participants were asked to keep a logbook tallying their daily physical activity by recording the time of day that they walked or ran and whether or not they were wearing the device. Weather conditions during the study were collected, and an extensive inventory of the characteristics of the trails and their adjacent landscape was created in a GIS.

3. Data Collection

To conduct the study, a significant amount of data had to be collected to answer the four study questions.

3.1 Trail Characteristics

Trail characteristics such as surface material, width, lighting, amenities and adjacent landuse and vegetation were collected in the field using Trimble's Pocket Pathfinder GPS and a HP iPaq PDA running ESRI's ArcPad 6 software. The ArcPad/PDA solution displayed a map containing the road and trail network along with the location of sample points that were pre-located based on a linear sampling distribution of 100 meters (Figure 1).The sampling points were pre-located on an ArcMap data layer using an aerial base with a resolution of two feet. Two graduate students walked each of the trails and stopped at each of the sampling points to photograph and record the trail characteristics. This process of recording information was simplified using six ArcPad form pages that contained pull down lists and check boxes for recording the inventory. The GPS-enabled PDA was the primary source of finding the general location of the sampling points in the field. Although the GPS-enabled PDA had an accuracy of 10 meters, it was sufficient for the study as the purpose was to record the general attributes at particular locations.

3.2 Community Vegetation

On occasion, participants at times used trails not inventoried in the study thus making the trail data incomplete. It was therefore necessary to create a community-wide land-cover layer. The existing land-cover data for the community was limited to a 15-meter resolution data set that was interpolated from color infrared aerials flown in 2002. This resolution was inadequate for the study, so a new land-cover map was created from the city's sub-meter 2003 photography to make a more accurate vegetation map. Four categories of land cover were digitized: deciduous, coniferous, agriculture fields and water.

3.3 Recording Participant Routes with GPS

One of the two GPS devices considered for study participants to wear was the Garmin Foretrex 101 (Figure 2). This GPS was selected because of its minimal weight, small form factor, accuracy, affordable cost and use of two AAA batteries. The other model considered, the Foretrex 201, offered the same functionality, but it had a higher price tag and used rechargeable batteries, making it unsuitable for use in a long term study. The Foretrex 101 measures 3.3” wide, 1.7” high and 0.9” deep (8.4 x 4.3 x 2.3 cm). The device weighs only 2.75 oz. (78g). The GPS itself is very easy to operate as participants only had to turn the device on and off. Enabling the Wide Area Augmentation System (WAAS) increased the accuracy of the Foretrex 101 from 10 meters to approximately 3 meters. Prior to the study, 47 of the GPS units were tested for accuracy by concurrently placing the devices on the ground at a known geodetic point and collecting data for a period of 10 minutes. Our inability to place all devices at the exact center of the known geodetic point at the same time introduced an error of approximately nine inches (22.9 cm). The test found that 36 of the devices had an average location within 2.5 meters of the known point, 9 devices were between 2.5 and 5 meters, 1 device was between 5 and 7.5 meters, and 1 device was just over 10 meters. It was determined that the device over 10 meters had the WAAS feature disabled.

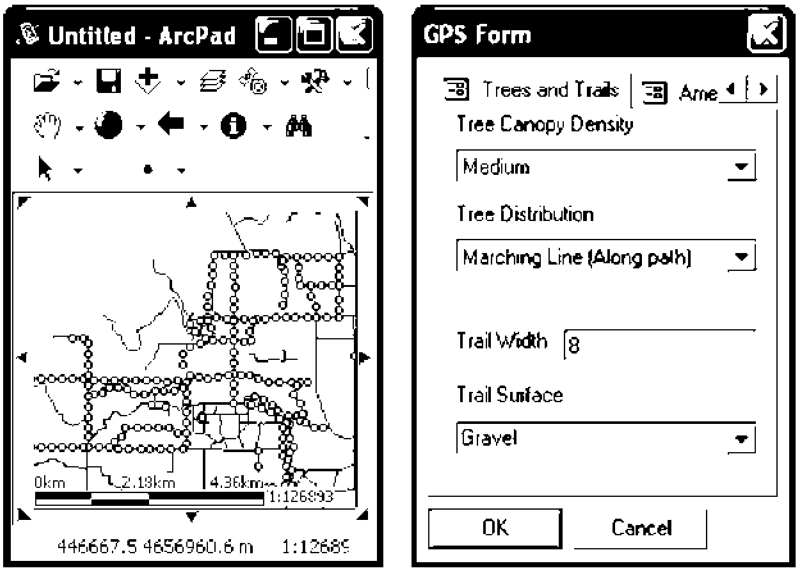

Figure 1: Sample ArcPad inventory forms

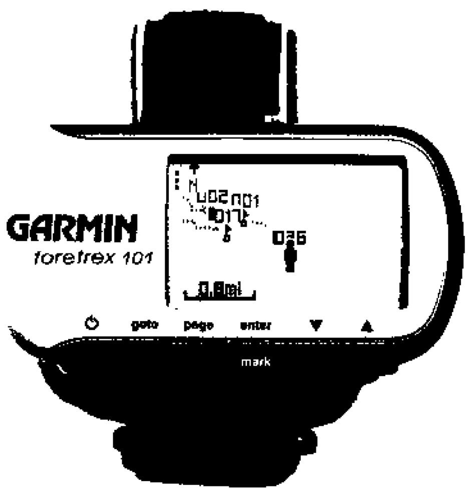

Figure 2: Garmin Fortrex 101

ISSN 1686-6576 / © Geoinformatics International

The results of the test corresponded with the results of Daniel Rodríguez's tests of the Foretrex 201 where he found 81.1 percent of the 726 points collected to be within 3.02 meters of the actual location (Rodríguez, Brown, and Troped, 2005). The most significant feature resulting in the selection of the GPS device was its ability to store a tracklog which records the location the participant walked/ran. The Foretrex 101 was able to record at intervals as short as one second for a maximum of 10,000 points. The study utilized a 10-second interval, sufficient for recording points every 220 feet (67 meters) for a fast four-minute mile or every 44 feet (13.4 meters) for a person walking an average three miles per hour. The tracklog data included latitude, longitude, UTM coordinates, elevation and time stamp. The tracklog also indicated when it was on and in use. The tracklogs' time-stamp values were contained in a single field using the year/month/day-hour:minute:second (2005/11/02-22:02:56) file format. The findings of Rodríguez et al. showed the importance of the location of the device on the body congruent with the impact to the quality of the data collected (Rodríguez, Brown, and Tropod, 2005). In keeping with that paper's recommendations, participants were instructed to wear the device on their wrists over clothing with the LCD facing up. The most significant limitation of the device was the life of the battery, providing close to 12 hours of use when using WAAS. In cold weather the life of the battery was significantly reduced, as the device would ultimately turn off after less than 30 minutes of use. This issue appears to be a result of the quality of the battery used. Due to this limitation, participants were advised to only wear the device when they were walking or running and to not leave the GPS device outside.

3.4 Accelerometry-based Activity Monitors

Accelerometry-based activity monitors, or physical activity monitors (PAM), are often used in health research because of their ability to measure physical activity in free living environments and their ability to digitally record physical activity as numeric values over a specified period of time. The PAM selected for this study was the BioTrainer- Pro by IM Systems (Figure 3). The Biotrainer-Pro is a biaxial activity monitor that has been demonstrated to provide valid and reliable indicators of physical activity (Welk, 2002). Data can be collected (and stored in internal memory) at intervals ranging between 15-second to 5-minute epochs. The data is stored in dimensionless units referred to as counts, which can be converted to estimates of energy expenditure using calibration equations (Welk, et al. 2003). With a sampling rate of 60 seconds, the device can hold 22 days of data. Downloaded device data includes a count value representing the amount of physical activity since the last interval point.

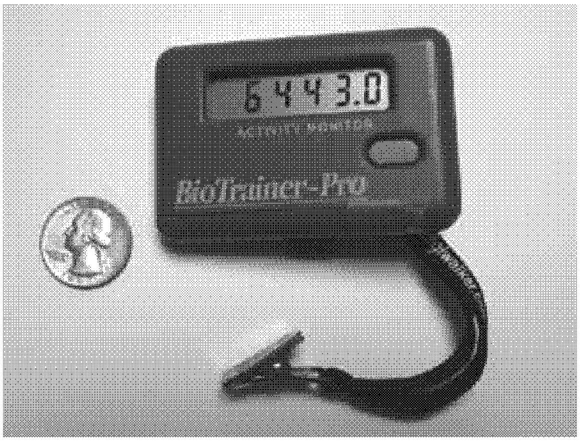

Figure 3: IM Systems Biotrainer-Pro

This data is graphed in a variety of ways showing the amount of physical activity an individual gets over a series of days, a week or as a daily comparison over each of the study weeks (Figure 4).The ability to temporally analyze and visualize the count data is facilitated through the time-stamp field that is stored with each count value. Whereas the time stamp recorded by the GPS is based on coordinated universal time (UTC), the Biotrainer-Pro's time stamp is based on the clock of the computer used to initialize the device therefore requiring that the computer's time be synchronized with UTC. The Biotrainer-Pro's time-stamp information is organized in a column/row format with the first column in each row containing the time expressed in hour: minute format and the following columns containing the count values for a particular day of the study.

3.5 Weather Conditions

Weather condition data for the study was accessed from the Iowa State University Department of Agronomy's Iowa Environmental Mesonet server (http://mesonet.agron.iastate.edu/schoolnet/dl/), which archives data that is automatically collected on a one-minute basis from a weather station at an Ames elementary school. This information included air temperature, wind direction, dew point, wind speed, relative humidity, solar radiation and altimeter pressure. The data obtained contained a time-stamp field in month/date/year 24 hour:minute format (11/2/2004 22:10).

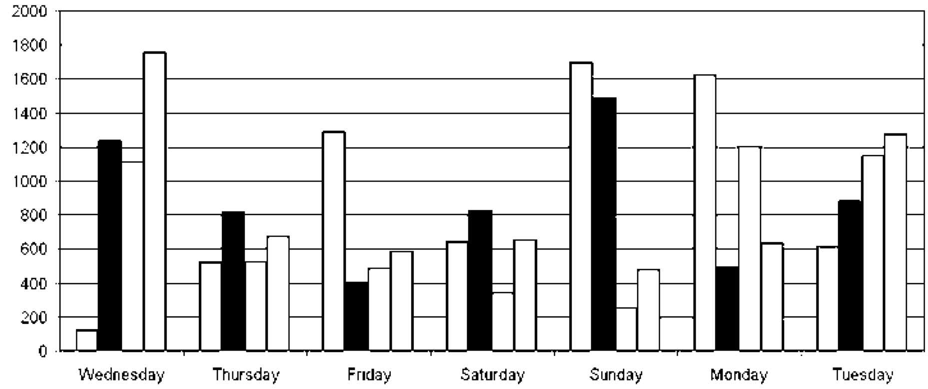

Figure 4: Daily physical activity totals for a single participant over the course of the four study weeks

4. Data Processing

All of the data collected via GPS and PAM was downloaded and reviewed at the end of each weekly study. The data was also cleaned and checked for errors prior to processing for analysis in a GIS.

4.1 GPS and PAM Data Cleaning

After downloading the tracklogs from the devices, the data was filtered to show only information recorded during the seven-day study period. Doing this reduced the number of points to be synchronized, thus making the files significantly easier to manage.

4.2 Testing for Errors

Three potential areas for introducing error into the data set were identified as part of the collection process. The first was a result of the GPS collecting an incorrect point. This typically resulted in a spike in the route path and was corrected using observational and mathematical techniques. The mathematical technique calculated the average distance between neighboring temporal points in order to identify the speed needed to get from point A to point B in 10 seconds. If this speed was significantly higher than the speed calculated for the previous two points the points were identified as suspicious and in need of potential correction. The second type of error resulted when the participant forgot to turn the device according to the given rules for when to use the device. times the participant left the device on when traveling in a vehicle or when riding a bicycle. These errors were noted and identified in the participants' paper logbooks. The final type of error resulted when the physical activity monitors were not initialized using a computer that had its internal clock synchronized to UTC. Correcting this problem was very difficult, requiring an extensive comparison of the raw GPS and PAM data along with the paper data log to determine the correct offset to be applied to the timestamp.

4.3 Data Joining

Joining the physical activity and weather information to the GPS data set first required the conversion of the time stamps into a common format. The selected format converted the time into a simple 4-digit integer representing the number of minutes past midnight. Each participant's ID number, trial number and day of the week were then appended to the beginning of the time value to create a unique code for every collected point in the study. For example, the record location for participant 6 during trial 1, day 2 at 2:36 PM would be identified by 6120876. A similar conversion format was applied to the GPS and weather data prior to conducting a table joins in ArcMap.

5. Analysis and Visualization

The data enabled the reconstruction of a participant's route on each study day. It was then possible to identify the trails used and the amount of physical activity the participant exerted between each recorded point. This is illustrated in Figure 5 where greater levels of physical activity are illustrated using larger dots that are shaded red. When looking at the collective set of data, patterns emerge showing the preferred trails and areas where individuals experienced the greatest amount of physical activity. While Figure 5 shows an example from just one participant, the data from all participants was aggregated into a single shapefile.

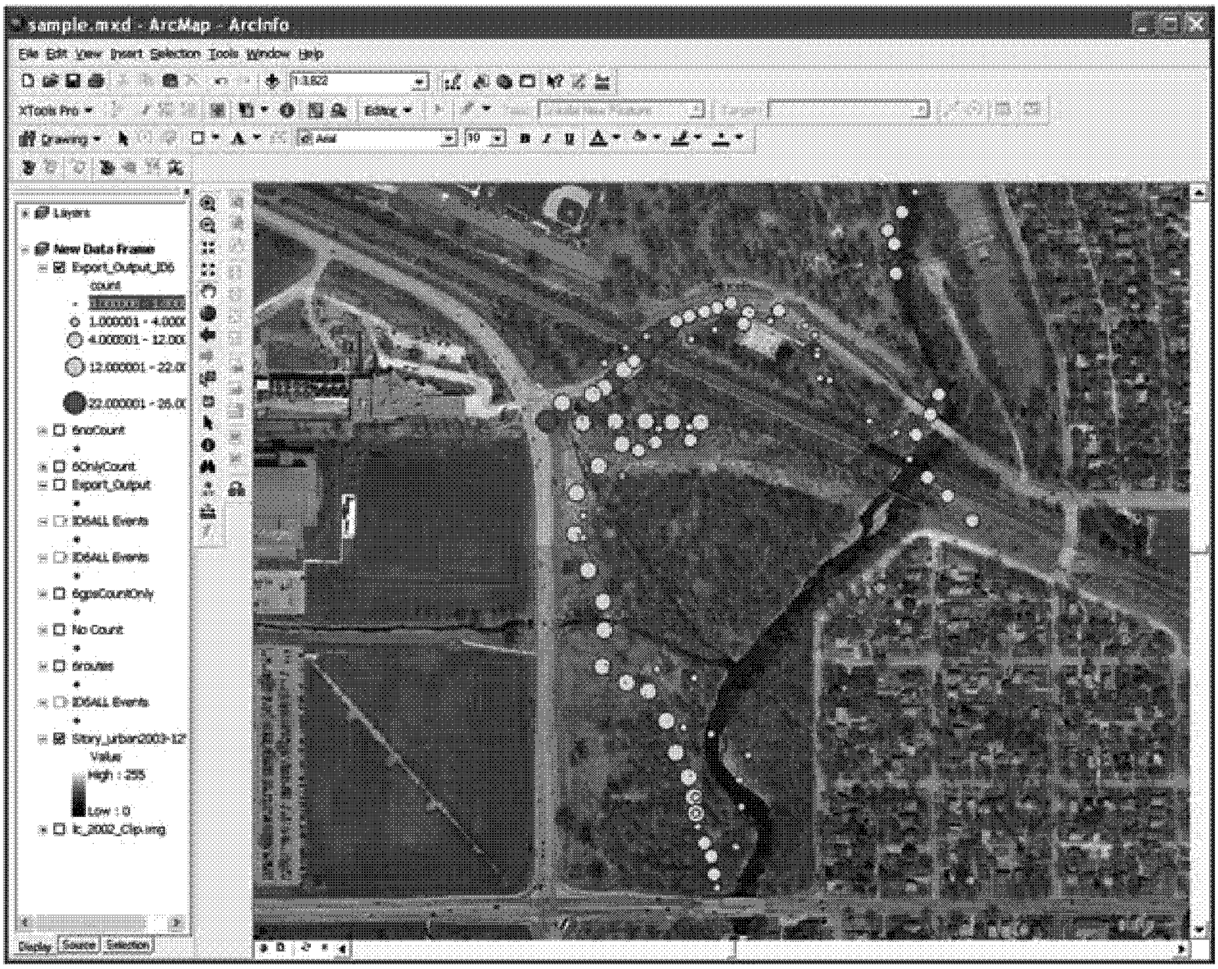

Figure 5: A single participants data for one week symbolized using physical activity counts where red (larger symbols) represents the highest level of activity

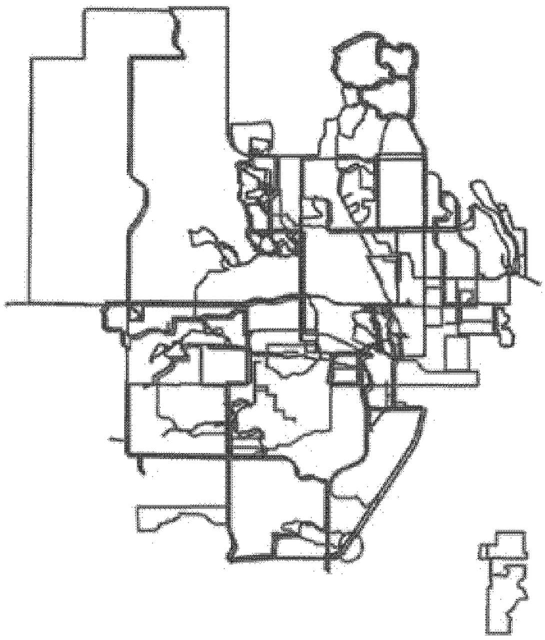

Figure 6: Favorite (blue) and most used (red) walking/biking routes

This allowed for analysis of relationships between vegetation and physical activity. Spatial techniques involving proximity overlap and zonal statistics were use to determine which routes were commonly used, which trails were underutilized, and patterns of vegetation and locations where physical activity values increased or decreased. In addition to providing the join capability, the time-stamp value allowed the data to be queried to only show activity for a specific time of day. Additionally, the use of the Tracking Analyst extension for ArcMap provided a means to visualize the use of routes over a specified period of time through an animation. The weather attributes provided a means to query the data set using a range of temperatures and to locate the most popular locations during these various weather conditions.

6. Spatial Surveys

At the conclusion of the study, all participants were asked to fill out an online textual survey and then stop by the investigators office to complete a spatial survey. In both surveys, participants were asked to identify 1) their most used route and 2) their favorite route. The online survey asked participants to describe these routes in words. In the spatial survey researchers asked the participants to identify the two routes on a digital aerial photo. Facilitation of this spatial survey was handled through the use of ESRI's sample extension ArcSketch. ArcSketch provided a means for quick fluidity in delineating the routes for each participant (Figure 6) using heads-up digitizing. While this information was collected as a last-minute consideration, it has proven to be very useful in developing preliminary analysis models. It also extends the quality of the data collected with the GPS and demonstrates an alternative method for collecting similar data in future studies.

7. Conclusions

This paper presented a framework for joining a variety of data sources together to create a map that can be visualized and analyzed to identify potential relationships between the built environment and physical activity. This process can be replicated and applied to other environmental and health research studies that utilize multiple characteristics of the built environment. Throughout the study, several lessons were learned that should be considered when conducting future studies.

- A paper log file is important to identify where participants did not follow the study protocol.
- It is necessary to clean and check all recorded data to prevent GPS error when the signal is lost.
- The BioTrainer-Pro plastic clip is breakable; therefore an elastic band with an alligator clip should be used as a secondary fastener. The device is prone to fall off when a wearer uses the restroom or changes clothes.
- The wristband (extender included) for the Foretrex GPS may not be long enough for all participants to wear on their wrist over winter clothing. Participants that wore the device under their clothes had weaker signal reception and less-accurate data points.
- Batteries performed poorly during the coldest days of the study. Therefore, research conducted during cold periods should utilize premium quality batteries that can maintain a charge when exposed to freezing temperatures for long periods of time.
- The BioTrainer-Pro includes an LCD display showing the current total count value. It is important to turn off the LCD display when initializing the device to keep participants from "chasing a magic number" in their daily walks/runs.
- It is important that the participants use the same GPS and PAM devices over the entire study. This will reduce the amount of data management required prior to joining the data.

Acknowledgments

Special thanks to graduate students Khalil Ahmad and Zoran Todorovic for their assistance in collecting the field data and managing the transfer of the device data.

References

Centers for Disease Control and Prevention, 2007a, U.S. *Physical Activity Statistics,* http://apps.nccd.cdc.gov/PASurveillance/StateSumV.asp?Year=2005

Centers for Disease Control and Prevention, 2007b, *NHANES Data on the Prevalence of Overweight and Obesity among Adults*: United States, 2003–2004, http://www.cdc.gov/nchs/-products/pubs/pubd/hestats/overweight/overwght_adult_03.htm

Centers for Disease Control and Prevention, 2007c, *Behavioral Risk Factor Surveillance System* (BRFSS) U.S. Obesity Trend 1985-2005, http://www.cdc.gov/nccdphp/dnpa/obesity/trend/maps/index.htm

Obesity and the Built Environment: Improving Public Health through Community Design

Conference Executive Summary, 2004, Washington, D.C., May 24–26.

Rodríguez, D. A., Brown, A., and Troped, P., 2005, Portable global positioning units to complement accelerometry-based physical activity monitors. *Medicine and Science in Sports and Exercise*, 37:11, S572-S581.

Stokols, D., Allen, J., and Bellingham, R. L., 1996, The social ecology of health promotion: implications for research and practice. American Journal of Health Promotion, 10, 247-251.

Transportation Research Board Committee on Physical Activity, Health, Transportation, and Land Use, 2005, Special Report 282: Does the Built Environment Influence Physical Activity? Examining the Evidence (Transportation Research Board and the Institute of Medicine of the National Academy of Sciences).

Welk, G. J., 2002, Use of Accelerometry-Based Activity Monitors for the Assessment of Physical Activity. *In Physical Activity Assessments in Health Related Research*, edited by G.J. Welk (Champaign, IL: Champaign, IL: Human Kinetics Publishers), 125-142.

Welk, G .J., Almeida, J., and Morss, G., 2003, Laboratory calibration and validation of the Biotrainer and Actitrac activity monitors. Medicine and Science in Sports and Exercise, 35, 1057-1064.