Published in final edited form as:

Spat Spatiotemporal Epidemiol. 2013 March ; 4: 15–23. doi:10.1016/j.sste.2012.10.001.

Imputational Modeling of Spatial Context and Social Environmental Predictors of Walking in an Underserved Community: The PATH Trial

Dawn K. Wilson, Ph.D.¹, Caitlyn Ellerbe, B.S.², Andrew B. Lawson, Ph.D.², Cassandra A. Alia, B.A.¹, Duncan C. Meyers, M.A.¹, Sandra M. Coulon, M.A.¹, and Hannah G. Lawman, M.A.¹

¹University of South Carolina, Psychology Department

²Medical University of South Carolina, Division of Biostatistics & Epidemiology

Abstract

Background—This study examined imputational modeling effects of spatial proximity and social factors of walking in African American adults.

Purpose—Models were compared that examined relationships between household proximity to a walking trail and social factors in determining walking status.

Methods—Participants (N=133; 66% female; mean age=55 yrs) were recruited to a police-supported walking and social marketing intervention. Bayesian modeling was used to identify predictors of walking at 12 months.

Results—Sensitivity analysis using different imputation approaches, and spatial contextual effects, were compared. All the imputation methods showed social life and income were significant predictors of walking, however, the complete data approach was the best model indicating Age (1.04, 95% OR: 1.00, 1.08), Social Life (0.83, 95% OR: 0.69, 0.98) and Income > \$10,000 (0.10, 95% OR: 0.01, 0.97) were all predictors of walking.

Conclusions—The complete data approach was the best model of predictors of walking in African Americans.

Keywords

physical activity; trail use; Bayesian spatial model; block group; social context

1.0 Introduction

Previous studies have shown the need to examine spatial clustering effects in social and health related research (King et al. 2004; Lawson 2006). Past studies have demonstrated that both proximity and transportation infrastructure have been positively associated with engaging in physical activity (PA) (Sallis et al. 1993; Gomez et al. 2004; Cohen et al. 2006;

© 2012 Elsevier Ltd. All rights reserved.

Correspondence to: Dawn K. Wilson.

Corresponding Author: Dawn K. Wilson, Ph.D.; wilsondk@mailbox.sc.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Davison and Lawson 2006). Abildso and colleagues (2007) found that neighborhood walkability was associated with perceived proximity to and use of a community trail in adults. Few investigators have, however, evaluated the relation between location of residence and proximity to environmental supports for walking such as community walking trails in underserved communities (low income, high crime rates).

Previous studies regarding neighborhood effects on walking and PA have neglected to consider the role of spatial data in underserved communities. Studies which account for location of residence have the ability to use spatial information not just to determine the proximity or access to PA areas, but also for assessing clustering of behaviors within social networks defined by residence. Consistent with recent work by Fowler and Christakis on the spread of obesity within social networks (Christakis and Fowler 2007), increased PA levels may develop in social network clusters, within which social connectedness among neighbors positively influences PA. Such spatial variation of residential PA levels can also be mapped and examined in relation to individual outcomes (King et al. 2004). This approach can be used to inform the geographic nature of social factors affecting PA and trail use.

The present study expands on previous research by using a hierarchical model-based approach for examining differential imputational modeling approaches to understanding the relationships between household proximity to the walking trail and social factors (social life, safety) in determining walking status in the walking intervention community of the Positive Action for Today's Health (PATH) trial (Wilson et al. 2010; Coulon et al., 2012). Problems with recruitment and retention are often associated with high levels of missingness in underserved populations. While imputation has become an attractive option for handling missing data, these methods have been underutilized and few applied studies exist comparing the impact of different methods of imputation in underserved populations (Schafer and Olsen 1998). Ample research exists for handling levels of missingness seen in prospective controlled studies (<10%), however this study is unique in that it uses sensitivity analysis to compare methods of imputation when levels of missingness are high, as is often seen in underserved community participation research. This study was an exploratory analyses to compare differential imputational models in determining if proximity and social factors would predict walking status in the walking intervention community of the PATH trial.

2.0 Methods

2.1 Participants

The PATH trial was designed to examine a 24-month environmental intervention for increasing PA and walking trail use in rural southeastern counties of South Carolina (Wilson et al. 2010; Coulon et al., 2012). Three communities were matched on census tract information (e.g. poverty, crime rates) and were randomized to either: a combined police-patrolled walking program with social marketing strategies to promote PA (full intervention), a police-patrolled walking program only, or no walking-related intervention. Assessments were conducted within each of the three communities at baseline, 6-, 12-, 18-, and 24-months. The focus of this paper is on the 12-month assessment in the full walking intervention community which took place in Sumter, South Carolina. The social marketing intervention was primarily targeted at the 12 month assessment which was the primary outcome point of the trial.

Participants were actively recruited from a random list of households in the census tract that was provided by the University Survey Laboratory at the University of South Carolina. A total of 828 households were called, with 279 not reached. Of the 549 persons reached, 341 declined discussion of the study and 196 were invited to participate. Of those invited to

participate, 75 enrolled and provided baseline information. The final sample included 133 participants with 56% actively recruited from the randomized phone lists and 44% recruited through volunteer advertisements displayed in churches, schools, and at local businesses.

Inclusion criteria included 1) African-American (three of four grandparents of African heritage), 2) age 18 years or older, 3) no plans to move in the next 2 years, 4) no medical condition that would limit moderate intensity exercise including life-threatening illness, 5) residing in the census area, 6) available to participate in the evaluations over the study period, and 7) controlled blood pressure ($<180/110$) and blood sugar levels (<300 non-fasting, fasting) (Grundy 2005). Enrolled participants completed an informed consent and attended biannual health screenings during which they completed anthropometric and psychosocial measures.

2.2 Demographic and Walking Measures

2.2.1 Demographics—Demographic data included age, gender, household annual income, and marital status. Height and weight were assessed objectively by trained measurement staff with body mass index (BMI) computed based on the standard formula of $\text{weight(kg)}/\text{height(m)}^2$.

2.2.2 Walking Status (Dependent Measure)—Participant walking status was determined using walking attendance data, which were obtained from walker sign-in sheets completed prior to PATH walks. Research staff reviewed sign-in sheets from the start of the program to the end of 12-month follow-up assessment and categorized PATH participants ($n=133$) who walked with the program at any point during this time as a “walker” or, for those who did not walk at all during this time, as a “non-walker”. We elected to dichotomize in this way given because the social marketing approach encouraged new walkers to participate in pride strides and walks which sometimes were just a one-time event. On average individuals who were categorized as walkers completed 12 walks ($SD=17.5$).

2.3 Geospatial Measures

Participant addresses were geocoded using ESRI Street Map North America 2000 9.3 as the reference data. Out of the 133 participants, 125 addresses were geocoded using the ESRI software and 4 were geocoded using batchgeo.com, resulting in a total of 129 geocoded addresses. Geocoding was used to convert each participant’s street address to map coordinates. If an address was unable to be located through use of this software, the unmatched addresses were geocoded using a free website batchgeo.com. Geocoded addresses were then assigned to corresponding block groups so that neighborhood-level crime and random spatial effects could be analyzed. Block groups were chosen as the level of aggregation because they represented the smallest spatial unit containing enough individuals for meaningful analyses, and also allowed the effects of potential miss-assignment to be addressed (see Figure 1). There were no significant differences in recruitment methods for block groups or for the spatial variables used in the modeling.

2.3.1 Distance—Distance from each participant household to the start of the walking trail was estimated using the Network Analyst extension in ESRI ArcGIS 9.3. The community center was designated as the starting point of the trail and community walkers signed in with the walking leader upon arrival to the center. Reference data for the analysis was provided by ESRI Street Map North America 2000 9.3. The Network Analyst extension and the reference data allowed for a dynamically modeled “travelable route”. Such a travelable route is defined by whether street networks (roads and adjacent sidewalks) have turns or route restrictions (road features which limit travel in a given direction, e.g. one-way streets). The resultant route follows a path that adheres to major roads and sidewalks in such a way that

the most efficient means of travel from the participant home to the walking trail can be modeled. The length of the modeled route – in miles – from each home to the walking trail is then used as the measure of distance.

2.3.2 Crime—The police department in the intervention community provided records of crimes occurring in and around the study area for three-month windows from October through December during baseline, 12-month, 18-month, and 24-month assessments. Only armed robbery, shooting, and animal cases were included in the final analysis; these crimes made up 27.6% of the total data collected. The decision to only include armed robbery, shooting and animal cases was based on empirical support indicating the relevance of these crimes in predicting neighborhood walking (Griffin et al. 2008). Furthermore, results from preliminary analyses justified the inclusion of these specific crimes; this is discussed further in section 3.0 Statistical Methods (see below). Records of crimes were provided as an event address recorded by the responding officer and included time and date stamps, and relevant crime codes. Geocoding was performed on unmodified addresses (N=84.8%) when a specific address was provided. If only the street name (e.g. Pennsylvania Ave.) or an attraction (e.g. The White House) was listed an approximate address was manually identified using Google Maps. Because of the high variability in accuracy, locations were smoothed to the block group level with the assumption that wrong assignments occurred in a spatially random and thus unbiased fashion. Crimes in a given category were then summed at the block group level and treated as a neighborhood-level effect for individuals.

2.4 Psychosocial Predictors

2.4.1 Perceptions of Crime—Perceptions of crime were measured using a 6-item subscale from the Neighborhood Environment Walkability Survey (NEWS) (Saelens et al. 2003). Factorial and criterion validity of the survey has been established (Cerin et al. 2009) (Adams et al. 2009). Test-retest reliability of the Crime Safety subscale of the NEWS is 0.80 (Saelens et al. 2003). Scale reliability was $\alpha = 0.67$ in the present study.

2.4.2 Self-Efficacy for PA—Self-efficacy for PA was measured using the Self-Efficacy for Exercise Questionnaire (Garcia and King 1991). The scale includes 16-items which assess respondents' confidence to exercise over the next 6-months despite barriers. Factorial and criterion validity of the questionnaire have been established (Wilcox et al. 2005) and the scale has an internal consistency ranging from 0.83 to 0.93 (Sallis et al. 1988).

2.4.3 Social Life—Social life was assessed using a 9-item self-report measure adapted from Parker and colleagues' 12-item measure of social interaction in the community (Parker et al. 2001). The adapted scale assesses social interaction within the neighborhood. Items ask how many days in the past month respondents have interacted with neighbors for various reasons. The original measure has been associated with the individual-level health (Eng and Parker 1994; Israel et al. 1994). The adapted 9-item scale has demonstrated validity in the Neighborhood Quality of Life Study (Saelens et al. 2002). Scale reliability for the present study was $\alpha = 0.78$.

3.0 Statistical Methods

The major individual outcome was a binary indicator of walking participation at 12 months. To model this outcome a Bayesian hierarchical modeling approach was adopted (Lawson, 2009). The model is depicted as:

$$\begin{aligned}
y_i &\sim \text{Bernoulli}(p_i) \\
\text{logit}(p_i) &= \beta_* X_i + v_i + w_1 + w_2 \\
\beta_* &\sim N(0, \sigma_{\beta_*}^2) \\
v_i &\sim N(0, \sigma_v^2) \\
w_1 &\sim N(0, \sigma_{w_1}^2) \\
w_2 &\sim \text{CAR}(\text{ADJ}, \sigma_{w_2}^2) \\
\sigma_*^2 &\sim \text{Uniform}(a, b)
\end{aligned}$$

The outcome (y_i) was assumed to have a Bernoulli distribution with walking participation probability p_i . The logit of p_i was modeled with a linear predictor consisting of fixed covariate effects and random effects. In this model it was assumed that parameters have prior distributions. At the next level of the hierarchy the fixed effect regression parameters (β_*) were assumed to have non-informative zero mean Gaussian distributions with variances ($\sigma_{\beta_*}^2$). In addition to regression parameters, random effects designed to capture different forms of extra variation (or confounding) in the data were also assumed. In particular an individual level frailty term (v_i) with zero-mean Gaussian prior distribution with variance σ_v^2 was assumed. Block group residence of participants was used to account for unmeasured block group effects via additional random effects. Two such effects were included: an uncorrelated block group effect (w_1) and an effect that addresses residual block group spatial correlation or heterogeneity (w_2). The first effect also had a zero mean Gaussian prior distribution with variance $\sigma_{w_1}^2$. The second effect had a spatially correlated prior distribution (a CAR model) (Lawson 2009). This prior distribution depended on the equal-weight adjacency matrix of the block groups and was also dependent on a variance ($\sigma_{w_2}^2$). At the next level of the hierarchy the standard deviations (σ_*) were given non-informative uniform prior distributions on small positive ranges for the regression parameters, individual frailty/ uncorrelated heterogeneity (UH), and correlated heterogeneity (CH), respectively (Gelman 2006). The models fitted were posterior sampled using MCMC updating on WinBUGS 1.4.3 (Lunn et al. 2009). All models fitted were checked for convergence using the BGR diagnostic after a burn-in of 10,000 iterations. A final thinned posterior sample of size 1,000 was taken.

A large range of candidate models was examined with different combinations of personal and contextual components based on the ecological model used to develop the community based intervention. Final model selection was based on two factors: (1) theoretical justification, and (2) model parsimony as indicated by information criteria and Bayesian modeling selection. A priority was to examine social marketing components of the intervention and thus models were examined which included variables such as social life and self-efficacy, neighborhood support, exercise confidence, distance-based measures, and local crime data at the block group level. Bayesian modeling selection was performed through the use of a Bernoulli indicator variable, which over successive iterations selects the best fitting model from a set of candidate predictors. A predictor was considered beneficial if the posterior mean of the indicator variable was ≥ 0.5 . These results were confirmed by comparing nested models using the deviance information criteria (DIC). In all models a random intercept term was included to allow for possible confounding. All models were fit using Winbugs Version 1.4.3 (Lunn et al (2009).

3.1 Missing Data

Missingness was classified as either complete data, fully missing data at 12-months, or as limited data at 12-months (no more than 20% of data missing). Participant's missingness

was defined using the entire available set of variables to ensure consistency through model building. As such it was possible that an individual with limited missingness may have had a complete record for the final variables of interest. Methods for the handling of missing data vary depending on context (Little and Rubin 2002). Imputation is often used to provide estimates of missing observations in analyses and tends to be used when it is important to make maximal use of observed data. Usually it is important that the mechanism of missingness is fully random. Within Bayesian hierarchical models, missing predictors are considered to be parameters and are estimated as such within the model. The use of the predictive distribution is justified when data are missing at random (MAR) or missing completely at random (MCAR).

The data were judged to be missing at random (exploratory data not shown), however the extensive nature of missingness necessitated that a range of strategies be used for creating missingness categories. Data considered primarily “complete” included only those participants who had complete records on all variables (N=87). Next, data sets were considered where different degrees of missingness are found, and where different strategies for imputation are used. In the first approach all complete cases and limited missing cases were included, where missing values were imputed within the Bayesian hierarchical model (N=102). In the second approach, any missing information was imputed from baseline data when available. Any remaining missingness was imputed within the Bayesian hierarchical model (N=133). In the third imputation approach, all cases were included and jointly modeled the outcome using baseline and 12-month data. Parameters were shared with both models, and all imputation was performed within the Bayesian hierarchical model (N=133). For each of the four modeling strategies, complete, limited imputation, baseline imputation, and joint modeling, a series of four nested models were used that included only fixed effects, an additional individual level frailty effect, an additional block group level uncorrelated spatial heterogeneity effect, and finally an additional block group level correlated spatial heterogeneity effect. Significant predictors from each model were compared to determine whether results were consistent.

4.0 Results

4.1 Demographics

The demographic characteristics are presented in Table 1. Walkers tended to be older, retired, and had slightly higher incomes than non-walkers.

4.2 Comparison of Models

Approximately 34.6% of participants were missing data for at least one variable at 12-months. Within each imputation approach, nested models were fit with the simplest model having fixed effects only and each subsequent model adding, random intercept, uncorrelated heterogeneity at block group level, and correlated heterogeneity in turn. Using deviance information criteria (DIC) to compare the nested models, the model containing all random effects was consistently the most parsimonious. In addition, results from each model indicated that social life and low income (income < \$10,000) were consistently identified as significant predictors. (Appendix Table 2) In light of the consistency of results, we present the results from the complete and limited analysis. Of the imputation methods, unbiasedness of the complete data set relies on assumed missing completely at random (MCAR: where the missingness mechanism is unrelated to the observed data and unobserved parameters). Because this could not be guaranteed in this study an alternative imputation method was adopted. Limited imputation was preferred for its balance between missingness assumptions and retaining a data driven analysis as would be required in the complete data and baseline or joint model imputation respectively.

The empirical correlations displayed in Table 3 support the results observed in the modeling: for the block group crime data, there was some evidence of a positive association between armed robbery and shooting and a negative correlation between crime perception, armed robbery and animal cases. While crime has a weak negative correlation with walking status, there appears a positive and higher correlation between walking status and income as well as age which is mirrored in the modeling results.

Table 4 displays the estimated parameters for both complete and limited imputation for the correlated heterogeneity model. Walking status was predicted positively by age, but negatively by social life and also income (< \$10,000). Other variables tend to be masked by age (crime perception), income (exercise confidence and crime perception), or social life (crime perception). Note that in the limited analysis, age becomes less significant. Figure 2 displays the posterior estimated block group level heterogeneity effects for both complete and limited imputation approaches. In general for either analysis the maps show similar patterns with the uncorrelated heterogeneity mapped estimates showing a relatively random pattern and the correlated estimates showing a clustered area in the north west of the region. This suggests that there is a propensity to walk in that area unexplained by the remainder of the model. In addition to the main significant effects found, a number of other predictors yielded estimates close to 'significant'. Age has a lower bound of 1.0 for the odds interval for the complete data, while armed robbery and animal cases had low odds with an upper limit close to one in both complete and limited imputation cases.

5.0 Discussion and Conclusions

The results of this study demonstrated that across all of the model methods, social life and income were consistent predictors of participants' walking status in the walking intervention community of the PATH trial. There were also marginal effects for age- and safety-related factors (armed robbery, stray animal cases) in predicting walking status across models. Thus, even though there was quite a bit of missing data, the different modeling techniques revealed effects that were fairly consistent regardless of the imputation method. The limited imputation model was based on fewer assumptions and thus provided the least bias results given all the potential imputation models considered.

A growing body of evidence supports the notion that PA is influenced by residential proximity and social factors (Giles-Corti and Donovan 2002; Saelens et al. 2003; Ogilvie et al. 2007; Forsyth et al. 2008). Ogilvie et al. (Ogilvie et al. 2007) reported that walking was influenced by individual's social networks and social support for walking in a review paper. Investigators have also demonstrated that social support has been strongly associated with increased levels of PA (Dowda et al. 2003; Anderson et al. 2006; Christakis and Fowler 2007). Fisher et al. (Fisher et al. 2004) found that community connectedness was associated with increased levels of neighborhood PA, and Nothwehr et al. (Nothwehr et al. 2007) reported that that positive social interactions (presence of strong social networks) were significantly correlated with PA in adults. Several investigators (Sampson et al. 1997; Cohen et al. 2006) have demonstrated that collective efficacy, a construct characterizing community social cohesion and trust, was negatively related to rates of obesity; however, few previous studies have evaluated the role of multilevel modeling that includes proximity and social environmental factors in understanding PA and walking behavior in underserved populations.

This is one of the first studies to demonstrate that a limited imputation analysis was the best fit model for demonstrating that social life predicted walking status in an underserved community-based intervention trial. Interestingly social life showed a negative association with walking, thus perhaps indicating that participants were more likely to join the walking

group if they perceived their immediate neighborhood social life to be low or negative. Few previous studies have examined social environmental factors in a multilevel statistical model to predict walking in underserved ethnic minorities. It has been suggested that effective obesity interventions may be improved through conceptualization of the problem and approach to behavior change on multiple levels of modeling (King et al. 2002). Understanding factors that influence PA beyond the individual level (e.g., the social and environmental level) are thus becoming the increasing focus of on-going research in the field (Michael and Carlson 2009). Mendes de Leon et al. (Mendes de Leon et al. 2009) used a multi-level modeling approach to examine neighborhood social conditions and walking in community-dwelling older adults. The results demonstrated that neighborhood-level disorder, but not social cohesion, was significantly associated with walking independent of individual-level neighborhood perceptions of walking. These associations, when adjusted for race, were no longer significant. Bird et al. (Bird et al. 2010) examined walking in older adults from diverse cultural backgrounds in Australia. Those who met national recommendations for walking attributed walking to environments that were more pleasurable and safe, and facilitated social interactions. Thus, the findings of the present study extend on previous studies by demonstrating that social life, income and safety (marginally) are important comparing differential imputational methods for examining social environmental predictors of trail use in older African Americans living in underserved communities. Most importantly, social interactions seemed to be more important than simply improving perceptions of safety, even though a high crime neighborhood was targeted within the PATH trial (Trumpeter and Wilson, 2012).

There are several limitations to the present study that should be noted. Analyses included only cross-sectional information and only assessed walking in one intervention community. Of the individuals considered walkers a fairly small proportion walked only once. Data from crime indicated that of the crimes included in this analysis (armed robbery, shooting, animal cases), 56.2% occurred in October, 13.2% in November, and 30.6% in December. Literature suggests that there are significant spatio-temporal trends that help to explain when and how often an individual initiates and maintains walking. However, it is beyond the scope of this data to explore this. Instead given the low-retention rate in this community, we focus on those factors that move an individual beyond the considerable barriers to become a walker. On-going and future analyses will report on the overall trial effects and the longitudinal aspects of the project. Furthermore, high levels of barriers to participating in research experienced by underserved adults may have limited the participation rate. However, as little research on underserved populations exists, the present study highlights important relationships in this population.

In summary, the results of this study provide support for using a limited imputation modeling approach for examining residential proximity and social factors on increasing walking in low-income, ethnic minorities. The approach to handling missingness in these underserved populations is a continual concern and this study demonstrated that a partial imputation method is superior to other approaches. Future research should continue to be aware of the data limitations in the longitudinal studies that include social environmental multi-level factors as predictors of health behaviors or outcomes.

Acknowledgments

This article was supported by grants R01 DK067615 and R01 DK067615-03S1A1 funded by the National Institutes of Diabetes, Digestive, and Kidney Diseases to Dawn K. Wilson, Ph.D. Send reprint requests to Dawn K. Wilson, Ph.D. Department of Psychology, Barnwell College, University of South Carolina, Columbia, SC 29208; wilsondk@mailbox.sc.edu

References

- Abildso CG, Zizzi S, Abildso LC, Steele JC, Gordon PM. Built environment and psychosocial factors associated with trail proximity and use. *Am J Health Behav.* 2007; 31:374–383. [PubMed: 17511572]
- Adams MA, Ryan S, Kerr J, Sallis JF, Patrick K, Frank LD, Norman GJ. Validation of the Neighborhood Environment Walkability Scale (NEWS) items using geographic information systems. *J Phys Act Health.* 2009; 6(Suppl 1):S113–S123. [PubMed: 19998857]
- Anderson ES, Wojcik JR, Winett RA, Williams DM. Social-cognitive determinants of physical activity: the influence of social support, self-efficacy, outcome expectations, and self-regulation among participants in a church-based health promotion study. *Health Psychol.* 2006; 25:510–520. [PubMed: 16846326]
- Bird SR, Radermacher H, Sims J, Feldman S, Browning C, Thomas S. Factors affecting walking activity of older people from culturally diverse groups: an Australian experience. *J Sci Med Sport.* 2010; 13:417–423. [PubMed: 19833556]
- Cerin E, Conway TL, Saelens BE, Frank LD, Sallis JF. Cross-validation of the factorial structure of the Neighborhood Environment Walkability Scale (NEWS) and its abbreviated form (NEWS-A). *Int J Behav Nutr Phys Act.* 2009; 6:32. [PubMed: 19508724]
- Christakis NA, Fowler JH. The spread of obesity in a large social network over 32 years. *N Engl J Med.* 2007; 357:370–379. [PubMed: 17652652]
- Cohen DA, Ashwood S, Scott M, Overton A, Evenson KR, Voorhees CC, Bedimo-Rung A, McKenzie TL. Proximity to School and Physical Activity Among Middle School Girls: The Trial of Activity for Adolescent Girls Study. *Journal of Physical Activity & Health.* 2006; 3:S129–S138.
- Cohen DA, Finch BK, Bower A, Sastry N. Collective efficacy and obesity: the potential influence of social factors on health. *Soc Sci Med.* 2006; 62:769–778. [PubMed: 16039767]
- Coulon SM, Wilson DK, Griffin S, St George SM, Kugler K, Trumpeter N, Wandersman A, Forthofer M, Robinson S, Gadson B. Formative Process Evaluation for Implementing a Social Marketing Intervention to Increase Walking in African Americans in the PATH Trial. *American Journal of Public Health.* in press.
- Davison KK, Lawson CT. Do attributes in the physical environment influence children's physical activity? A review of the literature. *Int J Behav Nutr Phys Act.* 2006; 3:19. [PubMed: 16872543]
- Dowda M, Ainsworth BE, Addy CL, Saunders R, Riner W. Correlates of physical activity among U.S. young adults, 18 to 30 years of age, from NHANES III. *Ann Behav Med.* 2003; 26:15–23. [PubMed: 12867350]
- Eng E, Parker E. Measuring community competence in the Mississippi Delta: the interface between program evaluation and empowerment. *Health Educ Q.* 1994; 21:199–220. [PubMed: 8021148]
- Fisher KJ, Li F, Michael Y, Cleveland M. Neighborhood-level influences on physical activity among older adults: a multilevel analysis. *J Aging Phys Act.* 2004; 12:45–63. [PubMed: 15211020]
- Forsyth A, Hearst M, Oakes JM, Schmitz KH. Design and Destinations: Factors Influencing Walking and Total Physical Activity. *Urban Studies* (Sage Publications, Ltd). 2008; 45:1973–1996.
- Garcia AW, King AC. Predicting long-term adherence to aerobic exercise: A comparison of two models. *Journal of Sport & Exercise Psychology.* 1991; 13:394–410.
- Gelman A. Prior Distributions for variance parameters in hierarchical models. *Bayesian Analysis.* 2006:1.
- Giles-Corti B, Donovan RJ. The relative influence of individual, social and physical environment determinants of physical activity. *Soc Sci Med.* 2002; 54:1793–1812. [PubMed: 12113436]
- Gomez JE, Johnson BA, Selva M, Sallis JF. Violent crime and outdoor physical activity among inner-city youth. *Prev Med.* 2004; 39:876–881. [PubMed: 15475019]
- Grundy SM. Diagnosis and management of the metabolic syndrome: an American Heart Association/ National Heart, Lung, and Blood Institute Scientific Statement. *Circulation.* 2005; 112:2735–2752. [PubMed: 16157765]
- Israel BA, Checkoway B, Schulz A, Zimmerman M. Health education and community empowerment: conceptualizing and measuring perceptions of individual, organizational, and community control. *Health Educ Q.* 1994; 21:149–170. [PubMed: 8021145]

- King AC, Stokols D, Talen E, Brassington GS, Killingsworth R. Theoretical approaches to the promotion of physical activity: Forging a transdisciplinary paradigm. *Am J Prev Med.* 2002; 23:15–25. [PubMed: 12133734]
- King, G.; Rosen, O.; Tanner, M. *Ecological Inference: New Methodological Strategies.* New York Cambridge: University Press; 2004.
- Lawson, AB. *Statistical Methods in Spatial Epidemiology.* New York: Wiley; 2006.
- Lawson, AB. *Bayesian Disease Mapping: Hierarchical modeling in spatial epidemiology.* New York: CRC press; 2009.
- Little, RJA.; Rubin, DB. *Statistical Analysis with Missing Data.* 2nd edition. New York: John Wiley; 2002.
- Lunn D, Spiegelhalter D, Thomas A, Best N. The BUGS project: Evolution, critique and future directions. *Stat Med.* 2009; 28:3049–3067. [PubMed: 19630097]
- Mendes de Leon CF, Cagney KA, Bienias JL, Barnes LL, Skarupski KA, Scherr PA, Evans DA. Neighborhood social cohesion and disorder in relation to walking in community-dwelling older adults: a multilevel analysis. *J Aging Health.* 2009; 21:155–171. [PubMed: 19144973]
- Michael YL, Carlson NE. Analysis of Individual Social-ecological Mediators and Moderators and Their Ability to Explain Effect of a Randomized Neighborhood Walking Intervention. *Int J Behav Nutr Phys Act.* 2009; 6:49. [PubMed: 19643024]
- Nothwehr F, Dennis L, Wu H. Measurement of behavioral objectives for weight management. *Health Educ Behav.* 2007; 34:793–809. [PubMed: 16816028]
- Ogilvie D, Foster CE, Rothnie H, Cavill N, Hamilton V, Fitzsimons CF, Mutrie N. Interventions to promote walking: systematic review. *BMJ.* 2007;334–1204. [PubMed: 17303861]
- Parker EA, Lichtenstein RL, Schulz AJ, Israel BA, Schork MA, Steinman KJ, James SA. Disentangling measures of individual perceptions of community social dynamics: results of a community survey. *Health Educ Behav.* 2001; 28:462–486. [PubMed: 11465157]
- Saelens BE, Sallis J, Black J, Chen D. Measuring perceived neighborhood environment factors related to walking/cycling. *Annals of Behavioral Medicine.* 2002; 24:S139.
- Saelens BE, Sallis JF, Black JB, Chen D. Neighborhood-based differences in physical activity: an environment scale evaluation. *Am J Public Health.* 2003; 93:1552–1558. [PubMed: 12948979]
- Saelens BE, Sallis JF, Frank LD. Environmental correlates of walking and cycling: findings from the transportation, urban design, and planning literatures. *Ann Behav Med.* 2003; 25:80–91. [PubMed: 12704009]
- Sallis JF, Nader PR, Broyles SL, Berry CC, Elder JP, McKenzie TL, Nelson JA. Correlates of physical activity at home in Mexican-American and Anglo-American preschool children. *Health Psychology.* 1993; 12:390–398. [PubMed: 8223363]
- Sallis JF, Pinski RB, Grossman RM, Patterson TL. The development of self-efficacy scales for health-related diet and exercise behaviors. *Health Education Research.* 1988; 3:283–292.
- Sampson RJ, Raudenbush SW, Earls F. Neighborhoods and violent crime: a multilevel study of collective efficacy. *Science.* 1997; 277:918–924. [PubMed: 9252316]
- Schafer J, Olsen M. Multiple imputation for multivariate missing-data problems: A data analyst's perspective. *Multivariate Behavioral Research.* 1998; 33:545–571.
- Trumpeter NN, Wilson D. Positive Action for Today's Health (PATH): Sex differences in walking and baseline perceptions of Safety and Access. *Environment and behavior.* in press.
- Wilcox S, Sharpe PA, Hutto B, Granner ML. Psychometric Properties of the Self-Efficacy for Exercise Questionnaire in a Diverse Sample of Men and Women. *Journal Of Physical Activity & Health.* 2005; 2:285–297.
- Wilson DK, Trumpeter NN, St George SM, Coulon SM, Griffin S, Lee Van Horn M, Lawman HG, Wandersman A, Egan B, Forthofer M, et al. An overview of the "Positive Action for Today's Health" (PATH) trial for increasing walking in low income, ethnic minority communities. *Contemp Clin Trials.* 2010; 31:624–633. [PubMed: 20801233]

Highlights

- Differential imputational modeling effects of spatial proximity to a walking trail, police reported crime, and social factors of walking status were compared.
- Bayesian variable selection and Hierarchical Bayesian modeling across all the imputation methods showed that social life and income were consistently significant-predictors of walking status.
- When restricting the data to individuals with complete data, age, social life and income were significant predictors of walking.
- Lack of social life and higher income were predictors of greater trail use in African Americans adults.

Figure 1.
PATH Full Intervention Community

Figure 2.
Differential Imputational Models

Table 1

Population Characteristics at Baseline and 12 Month Follow-up

Characteristic	Walkers N= 30		Non-Walkers N= 103		Missing	
	Baseline	12 Month Follow-up	Baseline	12 Month Follow up	Baseline	12 Month Follow up
Age (years, M [SD])						
	61.0(11.5)	62.3(11.4)	52.6(15.9)	54.2(14.8)	1	31
Female						
	22.0(75.9)	22.0(73.3)	66.0(66.7)	47.0(65.3)	1	31
Income						
< \$10,000	5.0(17.2)	3.0(10.7)	34.0(35.4)	29.0(40.8)	4	34
\$10,000 - \$24,000	7.0(24.1)	12.0(42.9)	35.0(36.5)	21.0(29.6)	4	34
> \$25,000	17.0(58.6)	13.0(46.4)	27.0(28.1)	21.0(29.6)	4	34
Employment Status						
Working	7.0(25.9)	8.0(27.6)	31.0(33.7)	19.0(27.5)	10	35
Retired, Permanently Disabled	14.0(51.9)	16.0(55.2)	30.0(32.6)	26.0(37.7)	10	35
Temporarily laid off, unemployed, homemaker, student	6.0(22.2)	5.0(17.2)	31.0(33.7)	24.0(34.8)	10	35
Marital Status						
Married or in an unmarried couple	14.0(48.3)	14.0(46.7)	29.0(29.3)	23.0(32.9)	1	33
Separated, divorced or widowed	14.0(48.3)	15.0(50.0)	49.0(49.5)	35.0(50.0)	1	33
Never married	1.0(3.4)	1.0(3.3)	21.0(21.2)	12.0(17.1)	1	33
BMI (kg/m², M [SD])						
	31.0(5.2)	32.2(8.2)	31.5(8.6)	32.3(8.6)	1	32

Table 2

Comparison of Model under all Analysis Choices

Imputation Method	Model	Significant Covariates ^a	DIC ^b
Complete (n=87)	Fixed Effect:	Social Life Income:<\$10,000	79.1
	Random Intercept	Social Life Income:<\$10,000	2.5
	UH	Social Life Income:<\$10,000	1.6
	CH	Social Life Income:<\$10,000	-172.2
Limited (n=102)	Fixed Effect	Social Life Income:<\$10,000	2,016.5
	Random Intercept	Social Life Income:<\$10,000	1,921.0
		Social Life Income:<\$10,000	1,913.4
	CH	Social Life Income:<\$10,000	1,741.7
Baseline (n=133)	Fixed Effect	Social Life	2,613.8
	Random Intercept	Social Life Income:<\$10,000	2,492.9
	UH	Social Life	2,470.7
	CH	Social Life	1,730.8
Joint (n=133)	Fixed Effect	Social Life Income:<\$10,000	11,690.1
	Random Intercept	Income:<\$10,000	11,475.3
	UH	Income:<\$10,000	11,477.5
	CH	Income:<\$10,000	10,256.8

^aSignificance determined as 95% Credible interval excluding zero^bDIC: Mean(Deviance) + Var(Deviance)/2

Table 3

Spearman Correlation Coefficients using only complete cases

	Distance	Walking Status	Age	Exercise Confidence	Crime Perception	Social Life	Armed Robbery	Shooting	Animal Case
Walking Status	0.25								
Age (yr)	0.13	0.30							
Exercise Confidence	0.12	0.24	0.14						
Crime Perception	0.16	0.17	0.23	0.22					
Social Life	-0.03	-0.17	-0.07	0.00	0.14				
Armed Robbery	-0.26	-0.18	-0.03	-0.23	-0.23	-0.17			
Shooting	-0.43	-0.11	-0.17	0.04	-0.02	-0.01	0.26		
Animal Case	-0.03	0.02	0.02	-0.16	-0.25	-0.05	0.10	-0.01	
Income	0.19	0.31	0.07	0.13	0.14	-0.03	-0.06	-0.11	-0.15

Table 4

Model Results for Correlated heterogeneity model, reported as MCMC sample mean and OR (95% CI)

Variable	Complete Analysis	Limited Analysis
Intercept	0.74 (0.18,2.18)	0.72 (0.16,1.92)
Distance (mi)	1.12 (0.57, 2.51)	1.19 (0.73,2.37)
Age (yr)	1.04 (1.00,1.08)	1.03 (0.99,1.08)
Exercise Confidence	1.01 (0.99,1.04)	1.01(0.99,1.04)
Crime Perception	0.99 (0.50,2.04)	1.07(0.54,2.21)
Social Life	0.83 (0.69,0.98)	0.81(0.67,0.95)
Armed Robbery	0.60 (0.19,1.32)	0.68(0.24,1.35)
Shooting	0.82 (0.19,2.28)	0.69(0.20,1.59)
Animal Case	0.90 (0.43,1.68)	0.91(0.49,1.63)
< \$10,000	0.10 (0.01,0.97)	0.13(0.01,0.92)
\$10,000–25,000	0.98 (0.19,4.77)	0.82(0.14,3.61)
> \$25,000	1.09 (0.21,5.65)	0.98(0.19,4.99)