HIGH-RESOLUTION DATA ON MOBILITY, THE BUILT ENVIRONMENT, AND PHYSICAL ACTIVITY: Embedding smartphone GPS and consumer wearables within a cohort study to improve environmental epidemiology

HARVARD UNIVERSITY

Graduate School of Arts and Sciences

DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the
Committee on Higher Degrees in Population Health Sciences,
have examined a dissertation, entitled,

"High-Resolution Data on Mobility, the Built Environment, and Physical Activity: Embedding Smartphone GPS and Consumer Wearables within a Cohort Study to Improve Environmental Epidemiology"

presented by

GRETE E. WILT

candidate for the degree of Doctor of Philosophy
and hereby certify that it is worthy of acceptance.

Peter James
Peter James (Mar 21, 2023 13:48 EDT)

Dr. Peter James, Sc.D., Committee Chair, Harvard T.H. Chan School of Public Health

Brent A Coull
Brent A Coull (Mar 21, 2023 14:12 CDT)

Dr. Brent A. Coull, Ph.D., Harvard T.H. Chan School of Public Health

Dr. Steven L. Gortmaker, Ph.D., Harvard T.H. Chan School of Public Health

Jaime Hart

Dr. Jaime E. Hart, Sc.D., Harvard T.H. Chan School of Public Health

Date: 17 March 2023

HIGH-RESOLUTION DATA ON MOBILITY, THE BUILT ENVIRONMENT, AND PHYSICAL ACTIVITY

Embedding smartphone GPS and consumer wearables within a cohort study to improve environmental epidemiology

A dissertation presented

by

GRETE E. WILT

to

the Department of Population Health Sciences

and

the Department of Environmental Health

In partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in the subject of

Population Health Sciences

Harvard University

Cambridge, Massachusetts

March 2023

Dissertation Advisor: Peter James   Grete E. Wilt

High-Resolution Data on Mobility, the Built Environment, and Physical Activity: Embedding smartphone GPS and consumer wearables within a cohort study to improve environmental epidemiology

ABSTRACT

The ubiquitous nature of smartphones and wearable devices create novel opportunities for epidemiological exposure assessment and measurement error correction within the context of the built and natural environments. Since the foundations of the field, environmental epidemiologists have sought to establish methods to quantify and characterize environmental exposures, and to determine appropriate spatial and temporal scale. This dissertation set out to address the following gaps in the literature: 1) quantification of exposure differences between residential and mobility-based greenness; 2) explore momentary associations between greenness and physical activity through smartphone and wearable device data collection; 3) examine associations of walkability and physical activity at the minute level using GPS data; and 4) understand nondifferential misclassification of walkability exposure and implications for regression calibration on the association between residential walkability and physical activity.

The first study (Chapter 2), addressing research gap 1, found residential-based distance buffer estimates of greenness are higher and more variable than mobility-based metrics. These findings contribute to discussions surrounding the choice of an optimal spatial scale for personal greenness exposure assessment.

The second study (Chapter 3) sought to undertake the second research gap of momentary associations by utilizing objective physical activity data at fine temporal and spatial scales to

present novel estimates of the association between mobility-based greenness and step count. Contrary to our hypotheses, higher greenness exposure was associated non-linearly with lower mean steps per minute after adjusting for confounders. We observed statistically significant effect modification by Census region and season.

The third study (Chapter 4) targeted research gap 3 and took a similar approach to the previous chapter. We utilized comprehensive mobility data at fine temporal and spatial scales to present novel estimates on the real time association between walkability and physical activity. We found higher walkability exposure was associated with overall higher mean steps per minute. Associations were non-linear in nature. These findings contribute to discussions surrounding adapting the built environment to increase physical activity, resulting in the potential for improved health outcomes downstream.

Lastly, the fourth study (Chapter 5) set out to assess the impact of measurement error found in residential-based walkability measures. We used mobility-based estimates to correct error-prone residence based estimates and then used these error corrected exposures to correct associations between walkability and self-reported physical activity for the error due to the use of residence-based exposures. This chapter highlights residential-based estimates of walkability slightly underestimate associations between walkability and physical activity. These findings highlight the impact of exposure misclassification on epidemiological studies of the built environment and physical activity. GPS data present a feasible solution to correct residential environmental exposures moving forward.

This dissertation contributes to the literature at the intersection of epidemiology, environmental health, and geography on human movement, the built environment, and physical activity. We build off an established U.S.-based nationwide prospective cohort through an internal mobile health (mHealth) substudy with momentary GPS and wearable data for exposure and outcome and in-depth information on individual and area-level covariates. By combining approaches from the fields of epidemiology and geography this mHealth dissertation explores improved exposure assessment using GPS, real-time mechanisms of association utilizing GPS data and integrating error corrections into large preexisting cohorts.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

DEDICATION

"My wound is geography. It is also my anchorage, my port of call."

- ***Pat Conroy, Prince of Tides***

Sasha, Ranger, and Sam – thank you for always anchoring me on this journey when I felt lost

ACKNOWLEDGMENTS

I am so thankful to so many who have supported me on this journey. My advisor, Dr. Peter James and the rest of my dissertation and oral examination committee: Drs. Francine Laden, Brett Coull, Steve Gortmaker, and Jaime Hart-thank you for helping me grow as a researcher and supporting me along the way.

To my colleagues at the Centers for Disease Control and Prevention and United HealthCare Community and State: thank you. My time at CDC shaped me into the researcher I am and United you provided a space where I felt so valued for who I was. I am endlessly appreciative. Most notably for Andrew Dent, Brian Lewis, Barry Flannagan, Elaine Hallisey, Erica Adams, Rehn LeSuer, Lillie Molavi, Phoebe Chastain, Saad Soroya, and Braya Hicks.

SPACEE lab was an incredible resource for me, I especially want to shout out Dr. Hari Iyer who has provided me with invaluable mentorship during this program.

All my wonderful PHS PhD cohort mates. Especially Unnati Mehta, Caro Park, Keya Joshi, Kat Sadikova, Ella Douglas-Durham, Chih-Fu Wei, Melissa Fiffer, William Borchert, and Futu Chen. I have learned so much from you and your friendship.

I am so thankful to my JP dog walking ladies: Becca Manning, Genevieve Spears, and Ella Douglas-Durham (a crucial cross-over). And their dogs: Molly, Moose, and Finn. You have all brought me so much laughter and support during this time.

To my SPACEE girls – Drs. Charlie Roscoe and Unnati Mehta and soon to be Dr. Cindy Hu. You three have been my mentors, friends, confidants, and greatest supporters these past few years. May everyone be so lucky to experience people like you in their lives.

Mom and Dad – thank you for everything.

Sarah, you are the greatest cousin in the world, thank you for always shining.

To my best friend and sister Ingrid, thank you for always knowing what to say and always being my favorite person to spend the day with. You made this last year in Boston the best.

Sasha and Ranger, I think the only thing that made this PhD bearable was you two.

Sam – we did it, time for our next great adventure. May it involve much less school and zoom and lots of time with the dogs outside.

Chapter 1.

Introduction

Background and Gap Statement

Built and natural environments, which encompass objective and subjective features of the physical environment, exert influence on health behaviors and outcomes. The field of environmental epidemiology has sought to establish associations between these environmental exposures and numerous downstream factors like cancers, cardiovascular disease (CVD), and all-cause mortality. This research is almost exclusively rooted in the built and natural environments around one's residential address. Thus, environmental exposures are contextual—a factor of the economic, social and physical environments (Klepeis et al., 2001), defined through a geography, from administrative units like Census tracts to a radius around a point of interest (likely a residence) called a buffer, rather than the actual personal exposure, suggesting a potentially poor proxy exposure. This link between place, environment, and health is not new.

In 1854, the father of spatial epidemiology, John Snow linked a contaminated water pump on Broad Street to the cholera outbreak, accounting for place and time to attribute exposure. As the field of place, environment, and health developed into environmental epidemiology, contextual exposures remained at the forefront. In 1993, researchers at Harvard linked higher air pollution levels to higher mortality rates across six cities in the United States (Dockery et al., 1993). Air pollution monitors within city limits aided scientists in linking the environmental exposure to individuals living in that geographic context (Dockery et al., 1993; Laden et al., 2006). While the Six Cities Study advanced environmental epidemiology in many regards, (e.g. the use of a large representative cohorts and exposure data that was linked to the individual) in a sense we are always attempting to find our way back to the Broad Street pump.

Not visualized in the infamous map, John Snow determined which households visited the specific water source, linking exposure to the outcome of cholera. In the twenty-first century, environmental epidemiologists still grappled with methods that allow us this insight. We recognize that individuals spend large amounts of time outside of their residential address. Global Positioning System (GPS) data offers a solution. With the advancements of digital health technologies (wearable technology and smartphones), GPS data has become more ubiquitous. This has resulted in improvements in contextual environmental exposures, as we can more accurately measure one's true exposure as they move through time and space. Additionally, in recent years, studies of protective environmental exposures have emerged as ways to promote health behaviors and prevent adverse health effects including cancers, CVD, and all-cause mortality that have been linked to other harmful environmental exposures like air pollution, chemicals, and metals. Two key potentially protective contextual exposures in environmental epidemiology are measures of greenness and walkability.

Greenness is characterized by land covered in vegetation including trees, grass and shrubs. We commonly measure greenness using the normalized difference vegetation index (NDVI). NDVI ranges from -1 to 1 with higher numbers indicating more green vegetation. NDVI values below 0 represent water, ~0 represent rocks and bare soil including concrete, and values ~0.6-0.8 represent temperate and tropical forests. Reviews of greenness and health literature suggest numerous health benefits, including improved mental health, sleep, cognitive function, brain activity, blood pressure and a reduction in all-cause mortality (Dadvand et al., 2014; Fong et al., 2018; James, Kioumourtzoglou, et al., 2017; James et al., 2015; Rojas-Rueda et al., 2019). Greenness is associated with improved mental health and cognitive function through the pathway of social cohesion and stress reduction (Hartig, 2008). The stress reduction theory posits how time in nature influences feelings or emotions through parasympathetic nervous system activation which to reduces stress and autonomic arousal because of our innate connection to the natural world (Ulrich et al., 1991). Regulation and filtration of noise, heat and humidity and air pollution are additional pathways. Physical activity has been explored as a potential pathway to health benefits including lower risk of chronic disease, improved mental health, lower blood pressure and a reduction in all- cause mortality ((Office of the Surgeon General (US), 1996; Office of the Surgeon General (US), 2010; Roscoe et al., 2022). The hypothesized association between greenness and physical activity acts by providing accessible recreation space (Nieuwenhuijsen et al., 2017; Roscoe et al., 2022). A 2021 review by Jimenez et al. found the bulk of research on the association to be cross-sectional in nature using NDVI as the exposure and self-reported physical activity data. Klompmaker et al. found that higher quartiles of residential greenness were associated with increased odds of self-reported outdoor physical activity. Almanza et al. found that greenness was associated with higher odds of moderate to vigorous physical activity, when comparing those in the 90th and 10th percentiles of greenness. They used GPS and accelerometry data for 208 children in California (Almanza et al., 2012). Conversely, Garrett et al. identified inverse associations between greenness and walking physical activity (Garrett et al., 2020). Additionally, In a review of youth health outcomes related to exercising in green spaces, a meta-analysis of fourteen studies across the globe indicated little evidence that exercise in green spaces is more beneficial than physical activity conducted in other locations (Mnich et al., 2019). In a 2016 paper on greenness and all-cause mortality, James et al. highlight that physical activity was not a strong mediator between the exposure and outcome (James et al., 2016).Despite mixed findings, researchers hypothesize physical activity could be a driving mediator in the pathway between greenness and long-term health outcomes.

Higher physical activity is associated with improved health outcomes including reduced levels of depression, anxiety, obesity, cardiovascular disease and mortality in addition to improved birth outcomes (Ambrey, 2016; Chong et al., 2019; Coutts et al., 2013; Cusack et al., 2017; Dai, 2011; Fong et al., 2018; Grigsby-Toussaint et al., 2011; James et al., 2015, 2016; Jimenez et al., 2020; Jones et al., 2009; Lachowycz & Jones, 2011; Maas et al., 2008; Zhang et al., 2018). Much of the greenness and physical activity literature utilizes residential exposure and self-report physical activity survey data. Further investigating the pathway between greenness and physical activity utilizing improved GPS exposure and objective wearable device outcome data allows researchers to ask questions including: “do measures of residential and GPS greenness differ?” and “do individuals partake in higher levels of physical activity in greener areas?”. In environmental epidemiology we aim to minimize error by more accurately measuring exposures and outcomes. Moreover, we aim to understand if greenness and walkable environments cause people to be more active. In the simplest sense, causality requires temporality. The exposure of interest must proceed the outcome. Methods exist in epidemiologic literature to get at the casual nature of associations. However, causality also requires researchers to investigate the pathways that we intuit these associations follow. To answer the questions: “do measures of residential and GPS greenness differ?” and “do individuals partake in higher levels of physical activity in greener areas?” we require high quality GPS and wearable device information. These investigations using GPS data shed light on potential error in exposure measures and mechanisms hypothesized in studies observing protective effects of greenness and numerous behavioral and health outcomes in the literature.

The exposure of walkability typically measures an area’s population, business, and street intersection density. Despite being an individual behavior, physical activity occurs in the context of the built and social environments. These environments exert influence on decision making (National Academies of Sciences et al., 2019). In a 2018 review, Karmeniemi et al. highlights that areas with more features of a built environment (such as housing density, public transit, and population density) were associated with higher levels of physical activity. This was due to higher amenity accessibility and transportation patterns (Kärmeniemi et al., 2018). Several papers have linked access to public transportation to higher physical activity (Giles-Corti et al., 2016; M. Smith et al., 2017). Public transportation provides a crucial opportunity for physical activity as walking often occurs at either end of transit routes. While numerous activities provide an opportunity to meet physical activity guidelines, the Surgeon General's 2015 call to action to increase physical activity, walking was reported as the most common form (Office of the Surgeon General (US), 2015). Research, has found robust associations between walkability and social cohesion, and physical activity (Carr et al., 2010; Frank et al., 2008; Humpel et al., 2002; James, Kioumourtzoglou, et al., 2017; Marquet et al., 2020; McCormack et al., 2017; Orstad et al., 2018; Roscoe et al., 2022; Rundle et al., 2016; Saelens et al., 2003) at both the residential and GPS spatial scale. GPS walkability exposures have been associated with self-reported and objective physical activity previously in the literature (James, Hart, et al., 2017; Marquet et al., 2020, 2022a; Orstad et al., 2018; Roscoe et al., 2022; Rundle et al., 2016). Additional studies have shown positive associations between residential-based walkability exposure and self-reported physical activity (Carr et al., 2010; Frank et al., 2008; Humpel et al., 2002; Saelens et al., 2003). Like with greenness above, by utilizing GPS measures of walkability we have the opportunity to investigate associations between walkability and physical activity. Additionally, we can quantify the error between GPS-based and residential-based measures of exposure. If this error is substantial, techniques like measurement correction will adjust previous estimates of the association between the residential exposure measure and the outcome and provide a more accurate estimate. GPS data can help researchers understand where and how physical activity is occurring within these built and natural environments. From here we can design improved interventions and implement urban planning and design policy to promote features of the environment that promote healthy behaviors like physical activity.

As mentioned above, limitations in both greenness and walkability research remain in exposure assessment and pathway investigation. Improvements to these measures using GPS data that have improved spatio-temporal resolution are essential for determining true extent of exposure. Similar to the Broad Street pump, individuals may be exposed to environments outside their residential locations. From here we can better explore potential pathway hypotheses. Digital health, utilizing GPS and wearable technology pose a solution for improvements to exposure and outcome assessment (Hystad et al., 2022; James et al., 2022). These rich data can address the uncertainty of the true spatial and temporal scale exerting the contextual influences, which is known as the uncertain geographic context problem (Kwan, 2012a, 2012b, 2019; Park & Kwan, 2017). Previously the literature has relied on aggregation and arbitrary spatial scale to derive an exposure to assess associations. However, GPS data captures the true exposure experienced across a specific timescale of interest. The uncertain geographic context problem highlights how the spatial delineation of an exposure or outcome can alter the association observed. For example, examining greenness exposure by aggregating data to a 1200m buffered distance around a residence, approximating walkable distance, is a frequently used metrics of greenness exposure, yet may not provide an accurate reflection of the true exposure experienced in one’s typical human movement patterns (Brokamp et al., 2016; Kwan, 2019). This typical human movement pattern, which we aim to measure, is often referred to as one’s activity space. Activity space isa term used to describe the set of locations with which a person has direct contact during day-to-day activities (Perchoux et al., 2016).

This dissertation addresses the gap in the current literature that relies mainly on residential measures of greenness and walkability contextual exposures. We aimed to do this by assessing how GPS (mobility) measures and residential measures differ, uncovering potential mechanisms through physical activity to long term health outcomes using minute level GPS and accelerometry data and addressing measurement error in residential exposures by correcting with GPS mobility exposures in a full prospective cohort. Our fine scale mobility exposure estimates are derived from a pilot mobile health (mHealth) substudy utilizing smartphone and wearable devices to obtain GPS and objective physical activity data at a 10-minute temporal resolution.

We set out to address the following challenges and literature gaps concerning the uncertain geographic context problem though the utilization of fine scale mHealth data in four distinct aims: 1) quantifying exposure differences between residential and mobility-based greenness; 2 & 3) addressing the uncertain geographic context problem though exploration of associations between greenness, walkability respectively and physical using smartphone GPS data and physical activity data from consumer wearables gathered every 10 minutes; and 4) understanding nondifferential misclassification of residential walkability exposure and implications for regression calibration on the association between residential walkability and physical activity using GPS mobility-based walkability exposure measures.

Gap 1. Estimate an improved mobility-based measure of greenness exposure

Despite growing research interest into the relationship between greenness and human health, measures of greenness exposure continue to be poorly measured in epidemiological studies. We explore differences in greenness measures using both traditional residence-based estimates and novel mobility-based estimates. Limited research exists on the concordance between these exposure measures. Though, this dissertation we aim to provide guidance on best practices for utilizing greenness exposure as a contextual factor.

Due to innovative mobile technologies integrated into the NHS3 mHealth substudy (n=337) of the NHS3 cohort, we measured the greenness individuals are exposed to on a routine basis as they move throughout the day. Aim 1 improves measures of greenness based on exposure derived from GPS data.

Gap 2. Examine minute-level associations of greenness exposure, walkability and physical activity and the impact of geographic bias on findings

A common critique in contextual exposure research is the inability to elucidate the mechanism that links exposure to outcome. Leveraging momentary data from the NHS3 mHealth substudy enables us to investigate potential high resolution pathways. Mobility-based exposure data and wearable accelerometry devices provide fine scale objective spatiotemporal data and allow us to examine the relationship of mobility-based greenness and walkability exposure and mean step count at the 10-minute level using intra-individual repeated measures data (Middelweerd et al., 2017; Prince et al., 2008; Slaght et al., 2017). This minute level data provided the additional opportunity to address the impact of geographic biases that result from the uncertain geographic context problem and daily selective mobility bias (Chaix et al., 2011) through sensitivity analyses limiting the spatial extent of human movement outside their primary activity space (Perchoux et al., 2013), restricting on workplace and examining active transportation (human propelled movement with walk to run velocities) data only.

Gap 3. Identify and correct for walkability exposure misclassification in large cohorts

Residential addresses do not capture true exposure that individuals are exposed to in their daily lives. The uncertain geographic context problem, which we have introduced above, theoretically leads to measurement error of environmental exposures (Kim & Kwan, 2018; Kwan, 2012a). Nondifferential measurement error is ubiquitous in these studies, yet little is done to address this bias, unlike attempts to adjust for confounding (Spiegelman, 2010). Here we investigated the association between exposure to walkability and self-reported physical activity. Utilizing the principles of transportability- the ability to take insights from one population and apply it to another, we quantify measurement error in residential measures of walkability exposure by applying regression calibration to the full study using mobility-based exposure from the NHS3 mHealth substudy (n=337, with smartphone GPS data). We apply mobility-based walkability as the improved exposure estimate to calculate new effect estimates and 95% confidence intervals in the full cohort. Examining the difference in effect estimates will add to the discourse on the uncertain geographic context problem and provide an example of how associations of walkability exposure and physical activity are affected by differences in corrected exposure estimates.

In chapters 2 through 5 we explore the following research questions and evaluate if our findings align with our a priori hypotheses outlined below:

Chapter 2. Quantify differences in greenness exposure comparing activity space estimates and traditional residence-based estimates in the NHS3 mHealth Substudy.

Hypothesis: Normalized Difference Vegetation Index (NDVI) calculated from the residential address-based buffer will be significantly different from the mobility-based NDVI.

Chapter 3. Quantify associations of GPS mobility-based greenness exposure with minute level wearable accelerometry data among participants in NHS3 mHealth Substudy

Hypothesis: Higher fine scale mobility-based greenness exposure will be associated with higher mean step count per ten minutes after adjustment for a priori confounders.

Chapter 4. Quantify associations of GPS mobility-based walkability exposure with minute level wearable accelerometry data among participants in the NHS3 mHealth Substudy

Hypothesis: Higher mean fine scale activity space greenness exposure will be associated with higher mean step count per ten minutes after adjustment for a priori confounders.

Chapter 5. Transport the NHS3 mHealth Substudy mobility-based walkability exposure measurement corrections to determine the association of greenness exposure with physical activity in the full NHS3 cohort.

Hypothesis: Higher mean yearly residence-based walkability exposure will be associated with higher self-reported physical activity controlling for a priori confounders. The estimated association will be corrected for the measurement error correction coefficient that was calculated using mean yearly mobility walkability exposure. The mobility-based walkability exposure estimate will provide higher estimates than the residential-based estimate.

Conclusion

The strength of this dissertation lies in the integration of digital health technology with preexisting prospective cohorts. This unique cohort allows us to target several gaps existing in environmental health research. Smartphone-based GPS enables the collection of improved exposure assessment and provides the opportunity to correct residential estimates. With intra-individual repeated measures GPS data we address geographic biases and examine potential temporal relationships of greenness and walkability exposures and physical activity. Finally, we tackle measurement error utilizing GPS mobility-based estimates from the NHS3 mHealth substudy to correct residential estimates in the full cohort for the association between walkability and physical activity. Overall, in the following chapters we examined how mobility-based exposure estimates of greenness and walkability alter our exposure/outcome associations and impact downstream interventions and population health findings.

CHAPTER 2.

Exposure Differences Between Residential and Smartphone Mobility-Based Greenness in a US Cohort of Nurses

Grete E. Wilt[1], Charlotte Roscoe[1,2], Cindy R. Hu[1], Hari S. Iyer[3], Isabel Holland[1,2], Unnati V. Mehta[1], Brent A. Coull[1,4], Jaime E. Hart[1,2]. Steven Gortmaker[1], Francine Laden[1,2], Peter James[1,6]

1. Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, Massachusetts
2. Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts
3. Section of Cancer Epidemiology and Health Outcomes, Rutgers Cancer Institute of New Jersey, New Brunswick, New Jersey
4. Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, Massachusetts
5. Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health, Boston, Massachusetts
6. Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, Massachusetts

Abstract

Background: Studies of greenness and health often assess exposure using residential circular distance buffers, which do not capture exposure occurring outside of the pre-defined residential environment.

Objectives: We compare greenness measures obtained from traditional residential-based buffers and novel smartphone mobility-based estimates.

Methods: We used data from the US-based Nurses' Health Study 3 mHealth study, which followed 337 participants for four 7-day sampling periods across a year. We used Landsat-derived Normalized Difference Vegetation Index (NDVI) data (30 m x 30 m resolution) for both residential (annual average and seasonal circular buffers of 270m and 1230m) and mobility-based greenness. We calculated mobility-based greenness exposure as seasonal NDVI values at GPS points captured every 10 minutes during the 7-day sampling periods and averaged across the seasonal sampling periods. We compared measures using descriptive statistics, Bland Altman tests, and Generalized Linear Models.

Results: Mean annual NDVI values from residential-based buffers (270m=0.40, SD =0.12; 1230m=0.39, SD=0.12) were higher than those obtained using annual mobility-based NDVI (mean = 0.32, SD=0.11). The Bland Altman agreement bias was 7.8% (95% CI: 6.7%, 8.9%) and 7.3% (95% CI: 6.3%, 8.4%) using the 270 m and 1230 m residential distance buffer, respectively, compared to mobility-based NDVI. Spearman's rank correlations comparing the mobility-based and residential-based NDVI were 0.59 and 0.57 for the 270m and 1230m buffer, respectively. The two residential distance buffers had a Spearman's rank correlation of 0.90. Each 10% increase in both 270m and 1230m residential-based NDVI, was associated with 6.0% increase in mobility-based NDVI (270m 95% CI: 5.3%, 7.1%; 1230m 95% CI: 5.1%, 6.9%).

Discussion: Residential-based distance buffer estimates of greenness are higher and more variable than mobility-based metrics. These findings contribute to discussions surrounding the choice of an optimal spatial scale for personal greenness exposure assessment.

Introduction

A growing body of evidence supports numerous health benefits of greenness exposure. Findings from prospective cohorts have shown that higher greenness is associated with better mental health outcomes, improved sleep patterns and cognitive function, lower blood pressure and lower all-cause mortality (Fong et al., 2018; James et al., 2015, 2016; Jimenez et al., 2021; Kaplan, 1995). Measuring greenness, however, presents a unique and difficult problem as a contextual environmental exposure that varies by place and time (Jimenez et al., 2021; Klepeis et al., 2001; Markevych et al., 2017; Spiegelman, 2010). In epidemiologic studies, greenness is most often measured quantitatively by calculating the mean satellite-derived Normalized Difference Vegetation Index (NDVI) within circular buffers around residential address history (James et al., 2015; Labib, Lindley, et al., 2020; Maas et al., 2008). These circular buffers ranging from 100 m to over 1000 m represent the environment directly surrounding one's residence or accessible within a short walk (James et al., 2016; Villanueva et al., 2014). These measures may fail to capture the exposure of interest, specifically the totality of greenness exposure that one encounters throughout an etiologically relevant time period (Fong et al., 2018; James et al., 2015; Labib, Huck, et al., 2020; Labib, Lindley, et al., 2020).

The discrepancy between measured exposures and the true spatial and temporal boundaries of the exposure exerting the contextual influences, like greenness, is known as the uncertain geographic context problem (Kwan, 2012a, 2012b, 2019; Park & Kwan, 2017). Kwan demonstrates that the spatial delineation of a contextual exposure, rather than the exposure itself, can influence the association observed (Kwan, 2012a). Thus, commonly used greenness exposures may not provide an accurate reflection of the true exposure by each individual, potentially biassing the results (Kwan, 2012b).

GPS data record typical human movement patterns, *mobility*, or a*ctivity space* (Brokamp et al., 2016; Kwan, 2019), a term used to describe the set of locations with which a person has direct contact during day-to-day activities (Perchoux et al., 2016). The widespread use of mobile phone and wearable GPS technology allows researchers to evaluate contextual exposures using mobility measurements and examine associations using a metric of mobility-based greenness exposure (Almanza et al., 2012; Halonen et al., 2020; Marquet et al., 2022a).

The lack of precise activity space data, that can be used to address the uncertain geographic context problem, is a remaining gap in the greenness exposure literature. To address this, our study used fine scale space-time GPS data to improve exposure metrics of greenness. Using mobile technologies integrated into the mHealth substudy (n=337) of the Nurses' Health Study 3 (NHS3) cohort, we measured mobility-based greenness. In this paper, we compare the agreement between residential distance buffer and mobility-based measures of greenness.

Methods

Population

Nurses' Health Study 3 (NHS3)

NHS3 is an ongoing open-enrollment prospective cohort of nurses and nursing students living in the United States or Canada that began in 2010. Study eligibility required participants to be a registered nurse, licensed practical/vocational nurse, or nursing student and to be born on or after January 1, 1965. At the time of selection for the mHealth substudy there were 49,693 participants. Upon entry, participants complete web-based questionnaires on lifestyle and medical characteristics and update their residential address every six months. The response rate for participants who have completed two or more questionnaires is above 80% (Chavarro et al., 2016; Gaskins, Rich-Edwards, Lawson, et al., 2015; Gaskins, Rich-Edwards, Missmer, et al., 2015; Mooney & Garber, 2019).

<u>NHS3 mHealth Substudy</u>

The NHS3 Mobile Health (mHealth) Substudy enrolled 500 NHS3 participants (Figure 2.1). The substudy began enrollment in March 2018 and data collection was completed in February 2020 with participants from 42 of the 48 contiguous states (Figure 2.2). To be eligible for the substudy, participants had to be aged 21 or older on March 12, 2018 and demonstrate adherence to questionnaire completion by providing information on height, weight, physical activity, and sleep in prior questionnaires. Participants with a doctor-diagnosed sleep disorder were not eligible because the study aimed to prospectively examine impacts of various lifestyle risk factors on sleep disturbance and FitBits have reduced accuracy in these populations.

The mHealth participants downloaded a custom smartphone application on their personal smartphones and wore a consumer-wearable fitness tracker (Fitbit™) for seven-day sampling periods every three months for a year from enrollment to capture seasonal variability in behaviors and exposures. Consistent with other mobility studies (Marquet et al., 2022a), we conducted a seven-day protocol to capture behaviors and exposures in a time frame that should include work and nonwork days, despite nurses' potential for shift work and nontraditional work schedules. We acquired GPS location data at ~10-minute intervals for each day throughout the sampling period if the mobile phone application was engaged. Further details of the substudy data collection methodology is detailed in (Fore et al., 2020).

We developed eligibility criteria for inclusion in our analyses. In the primary analyses, we included participants who provided at least eight hours of GPS data on three unique days over the entire study enrollment. Additionally, using Fitbit-derived sleep data, we omitted daily main sleep periods from the dataset as our primary interest was mobility-based greenness (Figure 2.1).

Exposure

We utilized the Normalized Difference Vegetation Index (NDVI) as a measure for greenness exposure. NDVI measures the reflection in the near-infrared (NIR) spectrum minus the reflection in the red range of the spectrum divided by those measures added together, identifying the amount of vegetation corresponding to the minimal difference between the NIR and red reflectance bands. This index ranges from -1 to 1 with higher numbers indicating more photosynthesizing vegetation. NDVI values below 0 represent water, near 0 represent rocks and bare soil including concrete and values near 0.6-0.8 represent temperate and tropical forests (Klompmaker et al., 2018). NDVI was rescaled so all values below 0 were recoded to 0. This practice is implemented so all non-green areas are valued identically (James, Kioumourtzoglou, et al., 2017; James et al., 2016; James, Hart, et al., 2017).

NDVI raster imagery for the residential and mobility greenness exposure metrics were available through the Google Earth Engine (GEE) platform (Gorelick et al., 2017) and processed using Earth Engine Landsat-specific processing methods for Landsat Tier 1 Raw Scene collection for Landsat 8. For mobility measures of greenness exposure, we linked each GPS point to the 30m x 30m NDVI raster grid cell that corresponded with its spatial temporal location.

For this analysis, we used NDVI to measure two distinct exposures to greenness. First, as previously used in greenness research, we calculated focal statistics for residential greenness exposure by proxy of 9 seasonal satellite images of NDVI around the home at both 270m, to approximate visible site distance, and 1230m, to approximate walking distance from residence, buffers (Figure 2.3), (Supplemental Information 2.1). Second, we used information retrieved from the smartphone application to calculate the mobility greenness exposure measure as the average greenness exposure of an individual from 30m NDVI values linked to each GPS point across the study period (Figure 2.3b). Due to the wealth of spatio-temporal geocoded GPS mobility data, we estimated exposure measures as yearly averages of greenness as well as by season (Fall, Winter, Spring Summer).

Covariates

Potential effect modifiers were identified a priori. These included individual participant measures of age (years, continuous), socioeconomic status defined as: education level (advanced degree, binary), and marital status (binary), and area-level measures of neighborhood socioeconomic status (z-score, quartiles), walkability (z-score, quartiles), season (quartiles) and region (census regions). We describe the variables in depth below.

Time invariant variables (age, education level and marital status) were obtained from the full NHS3 cohort study dataset in module 1 predating enrollment in the substudy. Age was reported as a continuous variable. We dichotomized education level to obtaining an advanced degree or not. Marriage was dichotomized into never married and ever married.

We measured Neighborhood Socioeconomic Status (SES) using a composite score of 7 Census tract level variables representing domains that have been previously associated with health outcomes including education, employment, housing, wealth, racial composition, and population composition (*DeVille et al. 2022, in review*). Variables were taken from the 2010 U.S. Census and each variable was z-standardized. We summed the z-scores for each component variable to create a neighborhood SES score. Higher scores indicate higher neighborhood SES. We joined these data based on residential address.

We defined neighborhood walkability as a composite score of intersection density calculated from 2019 Tiger/Line shapefiles of all roads with interstates removed (Bureau, 2020), population density from 2019 ACS population data (*Explore Census Data*, 2020) and 2018 business density data from Infogroup US Historical Business Data (Infogroup, 2020). Variables were z-standardized for each tract in the 2010 U.S. Census. We summed the z-scores for each component variable to create a neighborhood walkability index. Higher scores indicate more walkable areas. We joined these data based on residential address location (Supplemental Information 2.2).

Statistical Methods

We examined differences in residential and mobility-based NDVI measures using two residential greenness measures around the home (270m and 1230m) and a mobility-based greenness measure.

We compared these metrics (illustrated in Figure 2a,b) using histograms that showed the distribution of each exposure, paired t-tests to determine if there was a mean difference in the exposures, Spearman's rank correlation that measured the rank order correlation between exposures, and ANOVA tests that compared mean values across all greenness exposures. We report and compare variability across exposure measures as it relates to epidemiological power. We measured agreement between the measures using Bland Altman agreement tests and plots, a framework that evaluates bias between the mean differences in measures and estimates an agreement interval, within which 95% of the differences of the second measure compared to the first measure fall. Lastly, we regressed each residential greenness exposure on each mobility greenness exposure using generalized linear models to explore the shape of the relationship between the measures (i.e., explore nonlinearity and assess whether agreement was stronger within certain NDVI ranges) using penalized splines. We assessed the appropriateness of the linearity assumption for the relationship between residential and mobility-based measures using log-likelihood tests and Bayesian Inclusion Criteria (BIC).

Sensitivity Analyses

We performed analyses examining the agreement between the mobility-based measure and each residential-based measure of greenness exposure stratified by numerous characteristics including age group, residential neighborhood SES, weekday vs. weekend, geographic census regions, and walkability (Supplemental Information 2.2).

We performed the above agreement analyses and regressions stratified across characteristics including season, geographic region, walkability, and neighborhood SES. We also performed a sensitivity analysis in a restricted analytical cohort limited to participants who provided at least 12 hours of GPS data daily on five unique days in two distinct sample periods (Figure 2.1). To determine if individual's NDVI exposure and degree of agreement differ by work vs. non-work time, we conducted analyses omitting time at work (Figure 2.3c). Time at work was determined by geocoding (transforming a text-based address into a GPS location) workplace addresses at the time of study and restricting GPS points to locations outside of a 160-meter buffer (0.1 mile) of the workplace. The size of this buffer was based on likely hospital dimensions (*Insights from a Healthcare Architect's Journal*, 2019). We also performed sensitivity analyses restricting to walk-only data defined as mobility datapoints with velocities that fell between walking and running (0.8 to 4 m/s) (Cruciani et al., 2018), (Figure 2.3d).

Results

Residential and Mobility-based Greenness in Primary Analytical Cohort

We observed transportability between the full cohort and mHealth substudy (Table 1). The distribution of NDVI varied across residential and mobility-based greenness measures, with residential-based measures having on average slightly higher NDVI values than any of the mobility-based metrics (Supplemental Figures S2.1, S2.2). We measured a mean NDVI of 0.40 (SD = 0.12) for the 270m residential buffer and 0.39 (SD = 0.12) for the 1230m residential buffer. We measured a mean NDVI of 0.32 (SD = 0.11) for 30m mobility averaged NDVI. Annual average mean NDVI values per participant were lowest when calculated using the walking-only mobility dataset (0.28, SD= 0.09). (Table 2.2, Figure 2.4). Spearman's rank correlations between measures of greenness exposure showed highest correlation between the two residential greenness exposures (r=0.90) and lowest correlation between mobility and residential-based greenness exposure (r=0.59 and 0.57 for 270m and 1230m buffers, respectively).

Tests for agreement using Bland Altman methods (a method to plot the different values of two measurements against the mean for each subject and constructing limits of agreement) (Figure 2.5) indicated statistically significant non agreement between mobility and residential measures of greenness. The agreement bias between the mobility-based comprising of all mobility data and residential-based greenness measures was 7.8% (95% CI: 6.7%, 8.9%) using the 270m residential measure and 7.3% (95% CI: 6.3%, 8.4%) using the 1230m residential measure (Supplemental Table S2.1).

Results of linear regression suggest that the mobility-based greenness measures were inversely related to both residential-based greenness measures. For both a 10% increase in 270m and 1230m residential-based NDVI, all-mobility NDVI was associated with 6.0% increase in mobility-based NDVI (270m 95% CI: 5.3%, 7.1%; 1230m 95% CI: 5.1%, 6.9%; Table 2.2, Figure 2.6).

Sensitivity Analyses

The agreement bias between the workplace-omitted mobility-based greenness measure and residential-based greenness measures was 6.9% (95% CI: 5.5%, 8.4%) using the 270m residential measure and 6.6% (95% CI: 5.1%, 8.1%) using the 1230m residential measure. Finally, the bias for walking only mobility-based greenness was 11.5% (95% CI: 10.5%, 12.5%) using the 270m residential measure and 11.0% (95% CI: 9.9%, 12.0%) using the 1230m residential-based measure. In contrast, there was no evidence of significant agreement bias between the two residential-based greenness measures (0.5%, 95% CI -1.0%, 1.0%), (Supplemental Table S2.1).

For each 10% increase in 270m residential NDVI, workplace-omitted mobility-based NDVI increased by 7.3% (95% CI: 6.0%, 8.5%) and for each 10% increase in 1230m residential-based NDVI, workplace-omitted mobility-based NDVI increased by 6.8% (95% CI: 5.5%, 8.2%). Finally, for each 10% increase in 270m residential-based NDVI, walking only mobility-based NDVI increased by 7.8 (95% CI: 6.6%, 8.9%) and for each 10% increase in 1230m residential-based NDVI, walking only mobility-based NDVI increased by 7.7% (95% CI: 6.5%, 8.9%) (Table 2.2, Figure 2.6).

Findings regarding the agreement between residential and mobility-based greenness measures were consistent when the analyses were repeated in the restricted analytical cohort comprised of 208 participants who provided at least 12 hours of GPS data daily on five unique days in two distinct sample periods (Figure 2.1). Like results in the primary analytical cohort, residential-based NDVI measures of greenness were higher in magnitude than the mobility-based measures of greenness (Supplemental Table S2.1). These agreement bias measures were similar to those observed in the primary analytical cohort.

In analyses of the primary analytical cohort stratified by season of the year, agreement bias was lower between mobility-based greenness and 270m residential-based greenness during winter (4.3%, 95% CI 3.1%, 5.6%), while it was higher during summer (12.0%, 95% CI: 10.5%, 13.4%). These findings are consistent with changes in NDVI across season, with winter months typically having the lowest NDVI values reported and summer months recording the highest values of NDVI, potentially due to smaller activity spaces in colder months and lower deciduous vegetation cover resulting in less variation in winter months.

Most notably, there appeared to be differences in the degree of agreement bias present in analyses stratified by walkability of the neighborhood. In low-walkable neighborhoods, the bias in agreement for yearly mobility-based greenness was 9.0% (95% CI 7.7, 10.3) for the 270m residential-based NDVI and 8.3% (95% CI: 7.0, 9.7) for the 1230m residential-based NDVI. In high walkability neighborhoods, evidence for bias in agreement between all mobility-based NDVI and residential-based NDVI was 5.1% (95% CI: 3.0, 7.2) for the 270m residential NDVI and 5.2% (95% CI: 3.1, 7.2) for the 1230m residential-based NDVI. In analyses stratified by median age (<35 years vs. ≥ 35 years) and NSES (binary on median) there were no significant differences across strata.

Discussion

Researchers investigating health impacts of greenspace focus heavily on residential-based NDVI as a metric of greenness exposure (Fong et al., 2018; James et al., 2015; Jimenez et al., 2021). In our study levels of exposure to greenness were higher using residential estimates compared to mobility-based estimates based on GPS data, indicating a bias. We observed bias between mobility-based and yearly averages of 270m and 1230m residential buffer measures of greenness, respectively. Bias was greater in Spring and Summer compared to Fall and Winter. We found some evidence that level of neighborhood walkability modified the extent of the agreement bias between mobility and residential-based greenness measures as the bias was about 10% lower between the mobility measure and each residential based measure in high compared to low walkability neighborhoods. Walkability and greenness are strongly negatively correlated and participants living in highly walkable areas may both live and spend time in those walkable areas, whereas those residing in less walkable areas may travel to walkable areas for work and leisure.

Greenness exposure estimated by residential-based NDVI is potentially higher than mobility-based estimates due to greening of private property. Locations where populations gather like city centers and business districts tend to contain less vegetation (Nardone et al., 2021; Y. Zhang et al., 2017). These findings suggest that to increase population wide exposure to greenness, we must start with greening of population centers. Walkability and greenness are inversely correlated, suggesting that areas that are suitable for active transportation do not have high vegetation, thus we discovered active transport mobility had lower mean NDVI than the all-mobility metric and relative rankings differed, suggesting a potential source of bias in analyses of residential-based NDVI (James, Hart, et al., 2017).

There are limitations to this study. Not all participants contributed the same amount of person time, leading to better accuracy for individuals who provided more data and poorer accuracy for participants who provided less. However, our results regarding the comparison of mobility and residential based greenness measures were similar in a sensitivity analysis limited to participants who met strict inclusion criteria for wear time. While our study population is representative of the overall NHS3 population, participants were predominantly middle to high SES white females (Bao et al., 2016; Fore et al., 2020). Our study inclusion criteria of ownership of an iPhone likely contributed to overrepresentation of white individuals as it has been previously reported that iPhone ownership is more prevalent in white vs. other racial/ethnic groups where android phone ownership is more common. (Smith, 2020.). Thus, our findings may not be generalizable to individuals of other racial/ethnic groups and men if these demographics are related to mobility patterns.

Strengths of our study include a relatively large nationwide sample with rich covariate data. Our sampling method provided high spatial and temporal resolution time activity data with up to 28 days per person spaced out over a year for seasonal variability. We utilized high resolution spatial data on greenness, which we temporally matched to GPS data. Our time intensive longitudinal dataset provided the ability to omit sleep data, examine travel mode, and omit work addresses. Lastly, we incorporated multiple residential-based exposure buffers in our comparisons so in the future we can assess associations with health and behavioral outcomes to evaluate prediction.

While this study highlights differences in exposure measures of greenness, the methodology and principles we used also have application to other contextual built and natural environmental exposures such as neighborhood walkability, air pollution, and noise. Improving measures of these contextual exposures though mobility data is an important next step for environmental epidemiology.

Implications

By studying how estimates of greenness differ across measures, we gain important context on what existing measures are capturing, how to interpret these measures and correct them. Exposure to greenness during periods of active transportation (walking or running) was lower than greenness exposure measured via all mobility data. Use of residential-based measures resulted in higher values of greenness exposure. Thus, interventions promoting greening of active transportation pathways and other highly walkable areas in communities should be prioritized as areas of high walkability are frequently used public areas and the interaction of greenness and physical activity may reduce harms of urban walkability (noise and air pollution foremost), while promoting health benefits of active transport. As we look to green our cities, the answer may not lie solely in maintaining parkland but connecting neighborhoods and areas via green corridors.

Table 2.1. Characteristics of Nurses' Health Study 3 (N = 49,693) and the mHealth Substudy including recruited population and observations (N = 500, n = 701,696), primary analytic dataset population and observations (N = 337, n = 639,364), and secondary analytic dataset population and observations (N = 208, n = 498,521)

Variable		**NHS3 Cohort (N=49,693)**		**mHealth Population (N=337)**	
	Categories	N	% or Mean (SD)	N	% or Mean (SD)
Age	Continuous, years	49,516	36.33 (7.29)	330	36.01 (7.3)
Race	White	43,026	88.2	317	94.1
	Black	1,797	3.7	8	1.8
	Asian	1,529	3.1	2	0.1
	Mixed Race	1,058	2.2	4	1.2
	Other	1,385	2.8	6	1.8
Ethnicity	Hispanic	2,538	5.2	14	4.2
Married	Yes	27,852	57.1	207	61.4
	No	20,943	42.9	130	38.6
Advanced Degree	Yes	41,027	84.1	249	73.9
	No	7,768	15.9	88	26.1
Employment	Yes	40,808	93.0	319	96.6
	No	3,084	7.0	11	3.4
Seasonality	Fall	--	--	167,871	26.3
	Winter	--	--	127,860	19.9
	Spring	--	--	136,117	21.3
	Summer	--	--	207,496	32.4

Table 2.2. Distribution of Normalized Difference Vegetation Index (NDVI) by residential and GPS mobility-based measures and generalized linear models depicting association between residential and mobility-based greenness exposures using NDVI.

Exposure Measure	N	Mean Yearly (SD)	Min, Max Yearly	IQR Yearly	Residential NDVI 270m GLM		Residential NDVI 1230m GLM	
					Beta	95% CI	Beta	95% CI
270m Residential Buffer	333	0.396 (0.116)	0.069, 0.612	0.158	--	--	--	--
1230m Residential Buffer	334	0.391 (0.117)	0.005, 0.636	0.148	--	--	--	--
30m Mobility Average	335	0.318 (0.110)	0.007, 0.571	0.157	0.618	0.526, 0.710	0.599	0.509, 0.693
30m non-work Mobility Average	181	0.324 (0.108)	0.035, 0.565	0.143	0.729	0.604, 0.854	0.683	0.548, 0.818
30m Walk Only Mobility Average	336	0.287 (0.088)	0.013, 0.517	0.116	0.778	0.663, 0.893	0.77	0.654, 0.887

Figure 2.1: Study participant flow diagram for Nurses' Health Study 3 mHealth Substudy and restriction criteria for primary analytic dataset and secondary analytic dataset for cohort population (N) and GPS mobility observations (n).

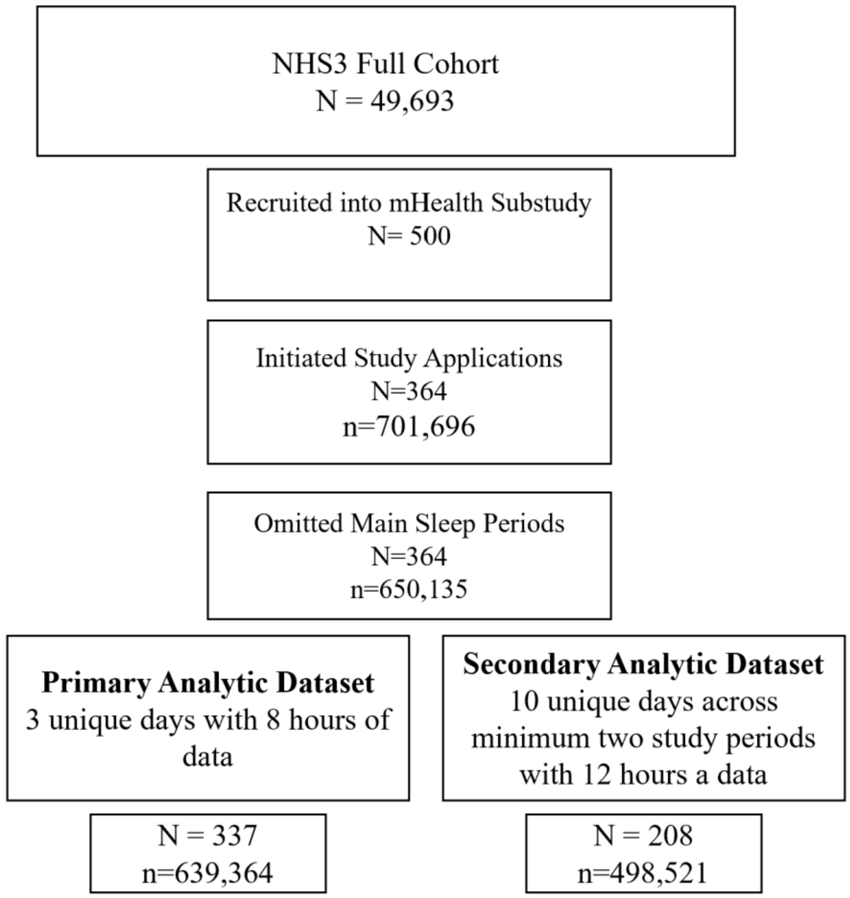


Figure 2.2. NHS3 mHealth Substudy participants residential locations across the contiguous United States, 2018

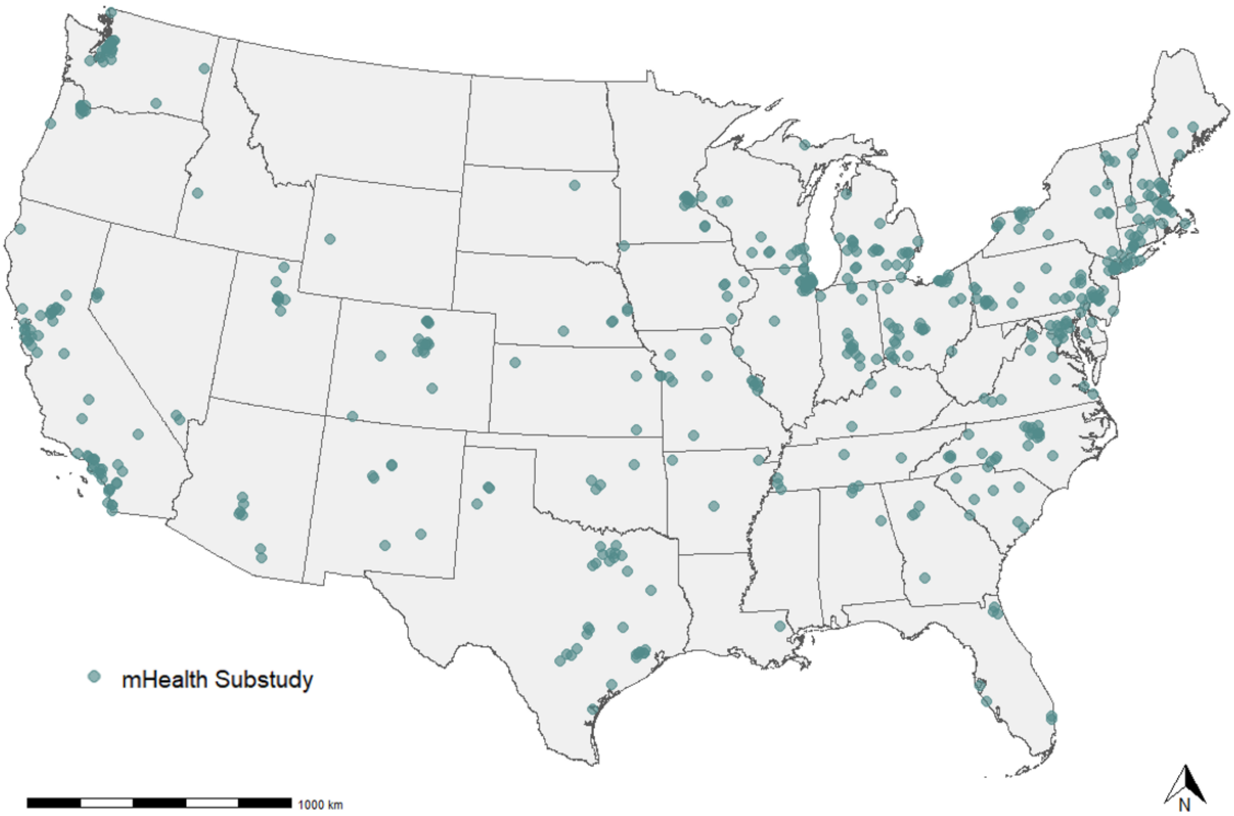


Figure 2.3: Four panel exposure map comparison of a) residential buffers, b) all GPS mobility data, c) non-work GPS mobility data and d) active transport (walk to run velocities) GPS mobility data (note: not actual participant data)

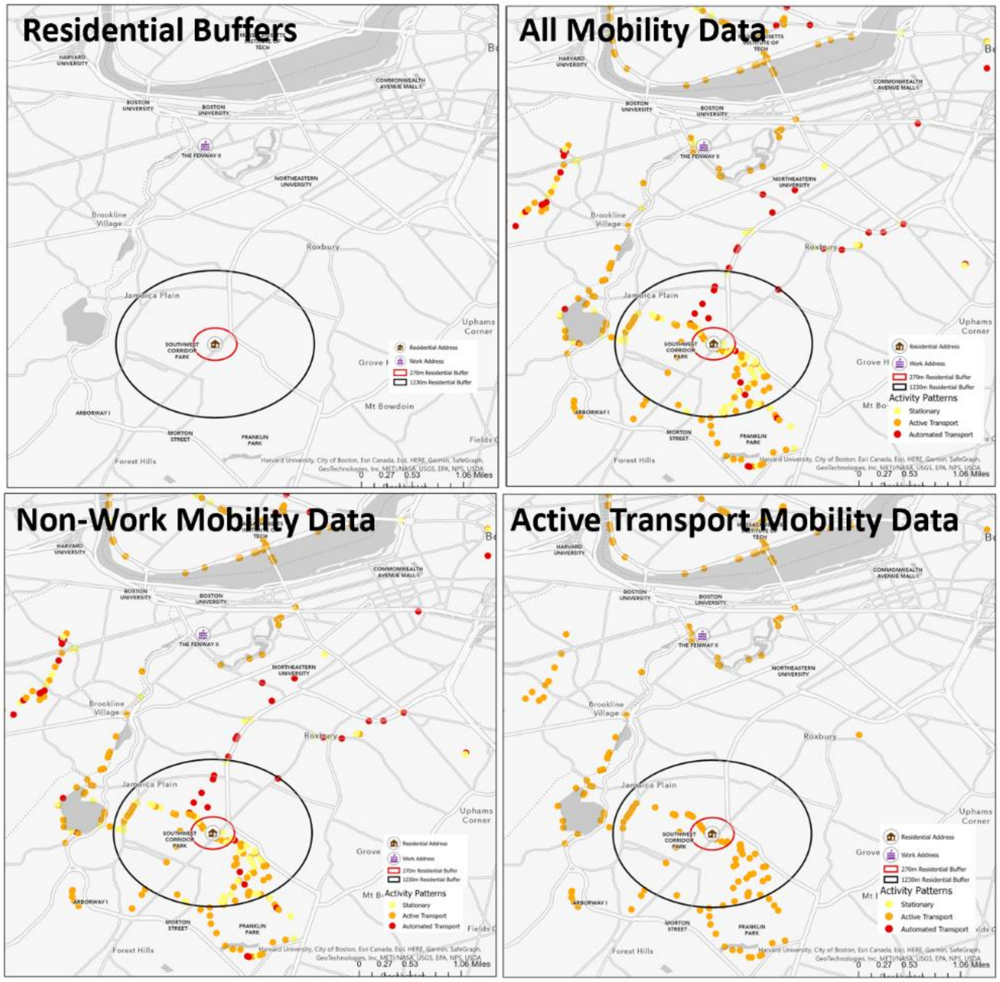


Figure 2.4: Yearly Normalized Difference Vegetation Index (NDVI) histogram distributions for 270m and 1230m residential measures and averaged 30m mobility exposure measures

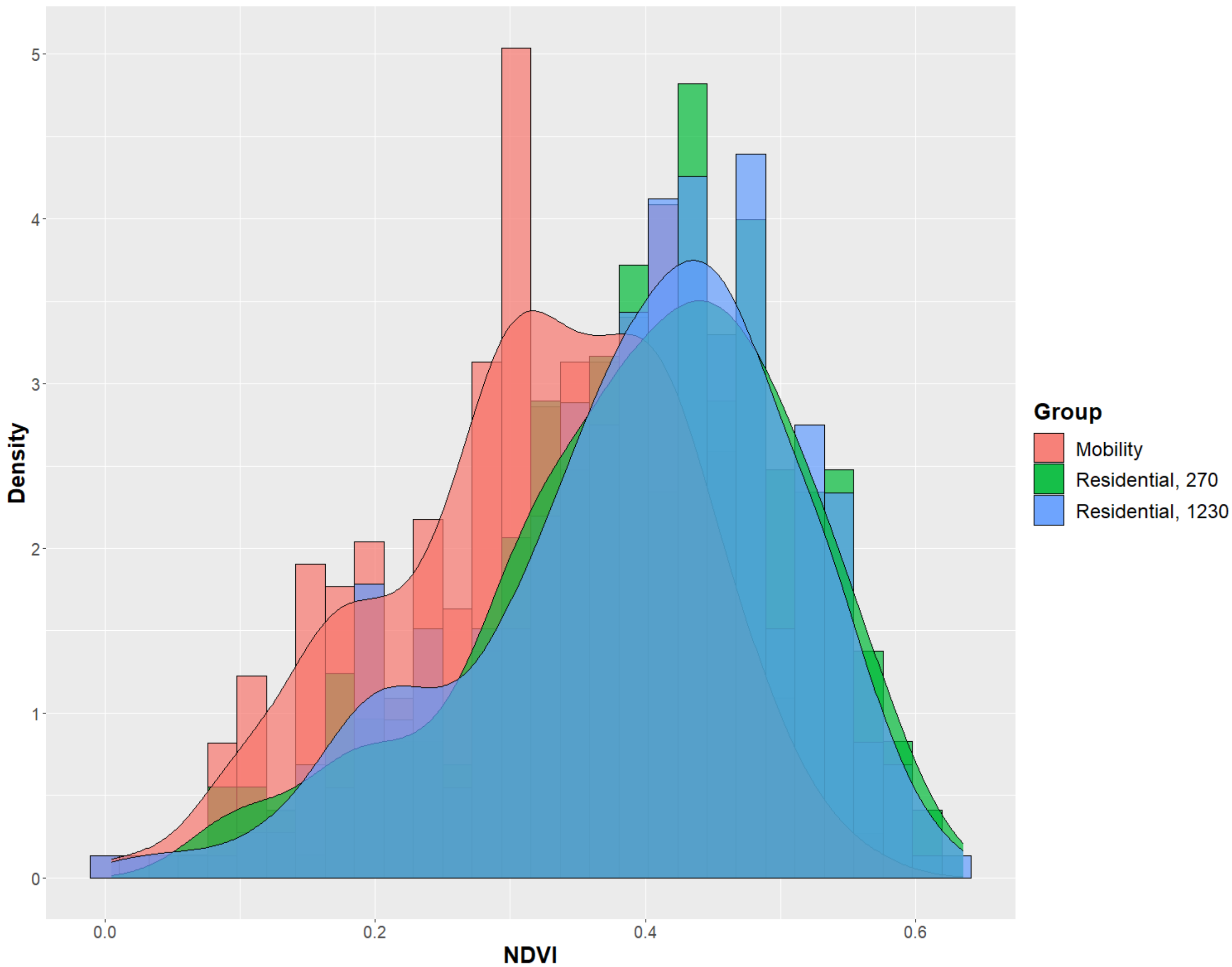


Figure 2.5: Bland Altman plots comparing mobility and residential Normalized Difference Vegetation Index measures. A) 270m residential NDVI vs. mobility-based NDVI, B) 1230m residential NDVI vs. mobility-based NDVI, C) 270m residential NDVI vs. 1230m residential NDVI, Purple band indicates 95% confidence level with red and green bands as lower and upper limit of agreement confidence bands.

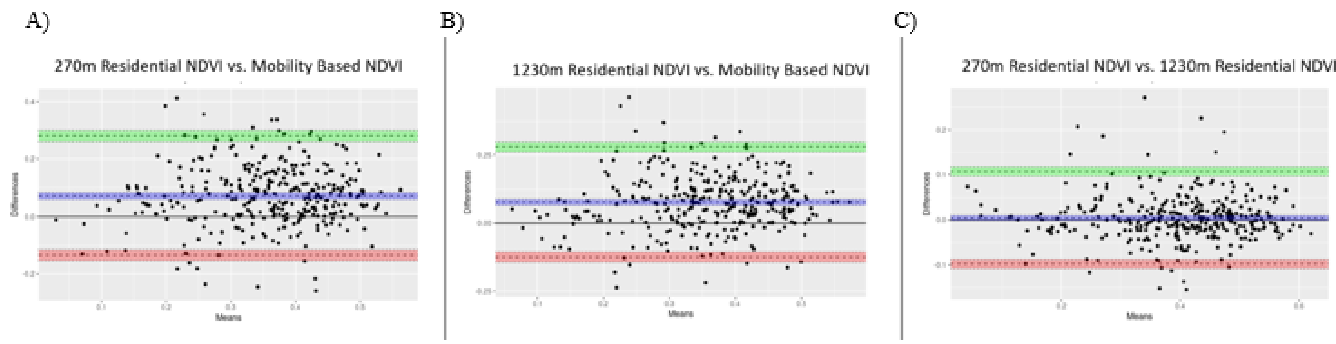
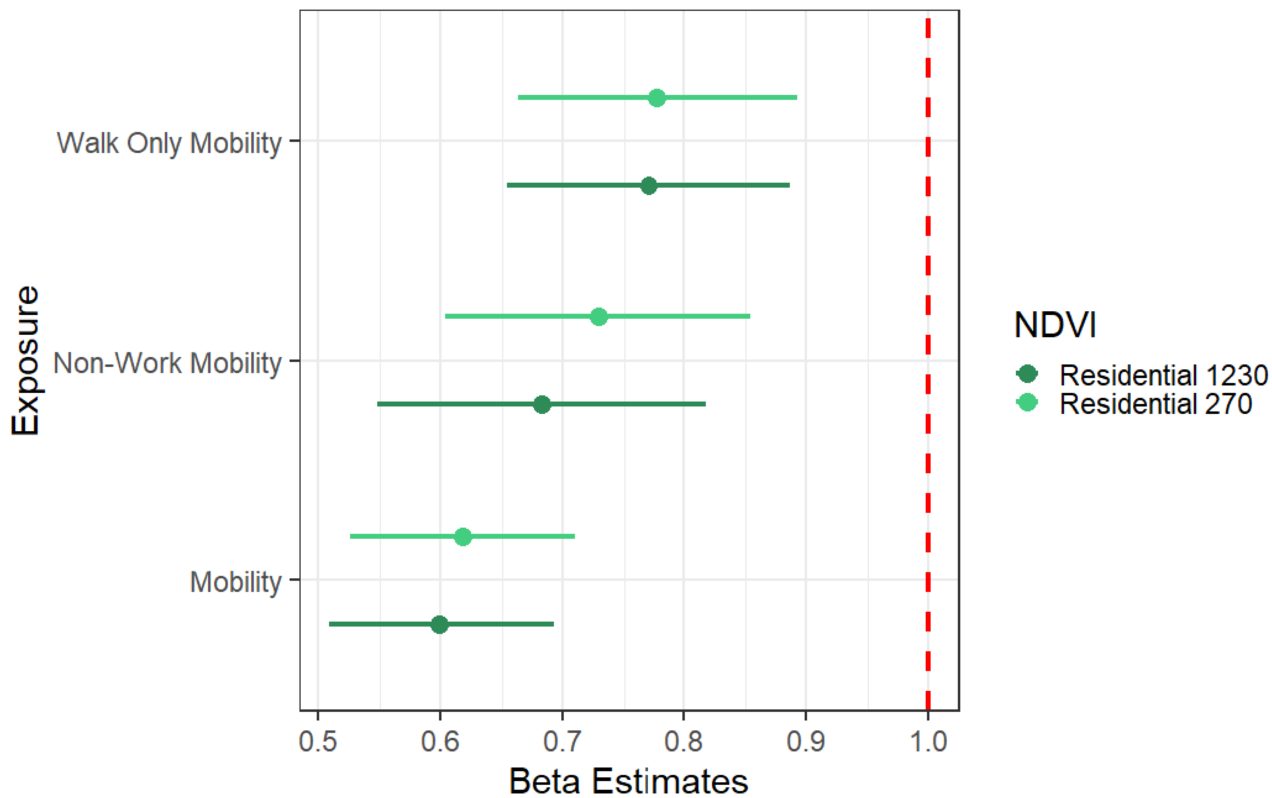


Figure 2.6 Generalized Linear Models Depicting Association Between Residential (270 and 1230m buffers) and GPS Mobility Based Greenness Exposures Using NDVI



Chapter 3: Minute Level Smartphone Derived Exposure to Greenness and Consumer Wearable Derived Physical Activity in a Cohort of US Women

Grete E. Wilt[1], Charlotte Roscoe[1,2,3], Cindy R. Hu[1], Brent A. Coull[1,4], Jaime E. Hart[1,2]. Steven Gortmaker[5], Francine Laden[1,2], Peter James[1,6]

1. Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, United States
2. Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, United States
3. Division of Population Sciences, Dana Farber Cancer Institute, Boston, Massachusetts, United States
4. Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, United States
5. Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, United States
6. Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, Massachusetts, United States

Abstract

Background: Inconsistent results have been found in the literature on associations of greenness, or vegetation quantity, and physical activity. However, few studies have assessed associations between mobility-based greenness and physical activity from mobile health data from smartphone and wearable devices with fine spatial and temporal resolution.

Methods: We assessed mobility-based greenness exposure and wearable accelerometer data from participants in the US-based prospective Nurses' Health Study 3 cohort Mobile Health (mHealth) Substudy (2018-2020). We recruited 500 female participants with instructions to wear devices over four 7-day sampling periods equally spaced throughout the year. After restriction criteria there were 337 participants (mean age 36 years) with n=639,364 unique observations. Normalized Difference Vegetation Index (NDVI) data were derived from 30 m x 30 m Landsat-8 imagery and spatially joined to GPS points recorded every 10 minutes. Fitbit proprietary algorithms provided physical activity summarized as mean number of steps per minute, which we averaged during the 10-minute period following a GPS-based greenness exposure assessment. We utilized Generalized Additive Mixed Models to examine associations (every 10 minutes) between greenness and physical activity adjusting for neighborhood and individual socioeconomic status, Census region, season, neighborhood walkability, daily mean temperature and precipitation. We assessed effect modification through stratification and interaction models and conducted sensitivity analyses.

Results: Mean 10-minute step count averaged 7.0 steps (SD 14.9) and greenness (NDVI) averaged 0.3 (SD 0.2). Contrary to our hypotheses, higher greenness exposure was associated non-linearly with lower mean steps per minute after adjusting for confounders. We observed statistically significant effect modification by Census region and season.

Discussion: We utilized objective physical activity data at fine temporal and spatial scales to present novel estimates of the association between mobility-based greenness and step count. We found higher levels of greenness were inversely associated with steps per minute.

Introduction

The explosion of research on nature and health in environmental epidemiology led to numerous studies investigating the association between exposure to greenness, or vegetation quantity, and physical activity, as well as chronic disease outcomes (Fong et al., 2018; James et al., 2015, 2016; Jimenez et al., 2021; Kaplan, 1995). Green environments have been hypothesized to be associated with higher levels of physical activity and to provide additional benefits compared to physical activity in non-green environments due to increased opportunities for physical activity and psychological restoration (Almanza et al., 2012; Coombes et al., 2010; Dewulf et al., 2016; Hillsdon et al., 2006; Kajosaari & Pasanen, 2021; Markevych et al., 2017; Mnich et al., 2019; Wheeler et al., 2010). However, previous studies examining the association of greenness with physical activity have reported inconsistent results (Klompmaker et al., 2018; Roscoe et al., 2022). Also, most of these studies used residential-based measures of exposure and self-reported measures of physical activity, making it difficult to infer true associations due to potential for measurement error (James et al., 2015; Jimenez et al., 2021).

Greenness exposure is often quantified by measuring surrounding residential greenness via satellite-derived greenness (Normalized Difference Vegetation Index) or greenspaces such as parks and gardens within a specific distance of the residential address (Fong et al., 2018; James et al., 2015; Jimenez et al., 2021). Residential exposures do not quantify exposure occurring outside of these selected distances, nor do they capture how much time an individual spends in nature, and residence-based analyses cannot be used to explore if individuals obtain their physical activity in green environments. Additionally, the appropriate scale of residential exposures is challenging to discern. Researchers remain uncertain of the true spatial and temporal boundaries exerting contextual influences (James 2014). This potential source of bias is known as the Uncertain Geographic Context Problem, which remains a critical limitation of prior research studies (Chaix et al., 2012, 2013; Kwan, 2012a, 2012b, 2019; Park & Kwan, 2017) evaluating greenness as an exposure. Due to the contextual nature of environmental exposures, there is not a set spatial boundary of influence. Measures of *activity space* (Brokamp et al., 2016; Kwan, 2019) – a term used to describe the set of locations with which a person has direct contact during day-to-day activities (Perchoux et al., 2016) – present a solution to the Uncertain Geographic Context Problem. A growing number of studies have collected objective measures of mobility-based greenness exposure and physical activity (Almanza et al., 2012; James, Hart, et al., 2017; Marquet et al., 2020, 2022a). Widespread use of mobile phone and wearable global positioning systems (GPS) technology (Markevych et al., 2017) have allowed researchers to evaluate contextual exposures using mobility-based measurements to quantify mobility-based greenness.

Using data collected from the Nurses' Health Study 3 (NHS3) Mobile Health (mHealth) Substudy participants, the aim of this intra-individual and repeated measures GPS study was to quantify associations of 10-minute level mobility-based greenness exposure with aggregated 10-minute level physical activity captured by a wearable device. Our aim was to determine associations between greenness exposure and physical activity using this rich source of objective data. We hypothesized that higher mobility-based greenness exposure was associated with higher mean steps-per-minute averaged over a 10-minute period, after adjustment for potential confounders.

Methods

Population

Nurses' Health Study 3 (NHS3)

NHS3 began in 2010 and is an ongoing open-enrollment prospective cohort of nurses and nursing students living in the US or Canada. Participants are required to be a registered nurse, licensed practical/vocational nurse, or nursing student and to be born on or after January 1, 1965 for eligibility into the study. At the time of selection for the mHealth Substudy there were 49,693 participants enrolled in NHS3. Once enrolled, participants provide updated residential history and complete web-based questionnaires on lifestyle and medical characteristics every six months. For participants who have completed two or more questionnaires, the response rate is above 80% (Chavarro et al., 2016; Gaskins, Rich-Edwards, Lawson, et al., 2015; Gaskins, Rich-Edwards, Missmer, et al., 2015; Mooney & Garber, 2019).

NHS3 Mobile Health (mHealth) Substudy

The NHS3 mHealth Substudy began enrollment in March 2018 and data collection was completed in February 2020 with 500 enrolled participants (Figure 3.1) residing in 42 of the 48 contiguous states during the data collection period (Figure 3.2). The mHealth Substudy required participants to be aged 21 or older on March 12, 2018 and demonstrate adherence to questionnaire completion by providing information on height, weight, physical activity, and sleep in prior NHS3 questionnaires for enrollment. As the study aimed to prospectively examine impacts of various lifestyle risk factors on sleep disturbance and Fitbit wearables have reduced accuracy in these populations, participants with a doctor-diagnosed sleep disorder were not eligible.

Fore et al. provide a detailed description of data collection methodology (Fore et al., 2020). In brief, mHealth participants wore a consumer-wearable fitness tracker (Fitbit™ Charge HR, Fitbit™ Charge 2 and Fitbit™ Charge 3) and downloaded a custom smartphone application on their personal smartphones for seven-day sampling periods every three months for a year from enrollment. This allowed us to capture seasonal variability in behaviors and exposures. Consistent with other mobility studies (Marquet et al., 2022b), we conducted a 7-day protocol. This time frame should capture behaviors and exposures across work and nonwork days. A mobile phone application acquired GPS location data at ~10-minute intervals throughout the 7-day sampling period. We omitted daily main sleep periods from the dataset under the assumption that physical activity does not occur during sleep periods using Fitbit™-derived sleep data. We included participants who provided at least eight hours of GPS data on at least three unique days in primary analyses (Figure 3.1, Figure 3.3).

Exposure

We used the Normalized Difference Vegetation Index (NDVI) as a measure of vegetation exposure, which was linked to GPS data to create a mobility-based greenness exposure. The NDVI is the most widely used satellite-derived indicator of the quantity of photosynthesizing vegetation and has been previously used as a marker for exposure to greenness in epidemiological studies (Fong et al., 2018; James et al., 2015). NDVI ranges from -1 to 1 with higher numbers indicating more green vegetation. NDVI values below 0 represent water, ~0 represent rocks and bare soil including concrete, and values ~0.6-0.8 represent temperate and tropical forests (Klompmaker et al., 2018). NDVI was rescaled so all values below zero were recoded to zero, so that all non-green areas were valued identically (James, Kioumourtzoglou, et al., 2017; James et al., 2016; James, Hart, et al., 2017).We used Google Earth Engine Landsat specific processing methods to produce seasonal, cloud-free, Landsat 8 raster images (Appendix A). We used Google Earth Engine Landsat specific processing methods to produce seasonal, cloud-free, Landsat 8 raster images (Supplemental Information 2.1). We linked these seasonal 30 m x 30 m NDVI raster images to season-matched GPS mobility data.

Outcome

We used accelerometry data from Fitbit™ wearable devices (Fore et al., 2020) to summarize physical activity in mean steps-per-minute, which we averaged for 10-minute interval after each GPS-greenness location. Mean steps-per-minute is preferable to raw step counts, as averages fluctuate less with fine scale missingness in GPS data (Armstrong et al., 2019; Yuenyongchaiwat, 2016).

Covariates

We identified potential confounders *a priori*. These included individual participant measures of age (years; continuous), socioeconomic status defined as: education level (masters in nursing or higher; binary), and marital status (never [never married]]/ever [married, widowed, divorced]; binary). Area-level measures included neighborhood socioeconomic status (z-score; quartiles), walkability (z-score; quartiles), mean daily temperature (Celsius; quartiles), daily precipitation (millimeters; binary), season and Census region (Northeast, Midwest, South, West).

We obtained age, education level and marital status from the full NHS3 cohort study dataset from participants initial questionnaire return (Module 1). Module 1 predated enrollment in the Substudy.

We used a composite score of 7 census tract level variables from the 2010 Census to estimate neighborhood Socioeconomic Status (nSES). Variables represented domains that have been previously associated with health outcomes, including education, employment, housing, wealth, racial composition, and population density (DeVille, n.d.). Z-scores were summed for each variable to create a nSES score. Higher scores indicated higher nSES (i.e. less socioeconomic deprivation). We joined quartiles of nSES score using the location of each 10-minute GPS point to create a mobility-based nSES.

We defined neighborhood walkability, a measure of population and business density, for each Census tract in the US as a composite 3-item score. This included z-scored intersection density calculated from 2019 Tiger/Line shapefiles of all roads with interstates removed (Bureau, 2020), population density, from 2019 ACS population data (*Explore Census Data*, 2020), and business density, from 2018 Infogroup US Historical Business Data (Infogroup, 2020). We summed the z-scores for each component variable (3-items) to create a neighborhood walkability index. Higher scores indicated more walkable areas. We joined quartiles of walkability score using the location of each 10-minute GPS point to create a mobility-based walkability.

We obtained daily mean temperature and precipitation data at 800m spatial resolution for the study period (2018-2020) from Parameter-elevation Regression on Independent Slopes Model (PRISM) (Luzio et al., 2008). PRISM variables were joined on date and paired GPS coordinates of each 10-minute repeated measure for mobility-based measures of temperature and precipitation. We classified daily mean temperature into quartiles and dichotomized precipitation to any precipitation/no precipitation.

We defined the Census region of each GPS point as one of 4 census regions (Northeast, Midwest, South, West), and derived season (Spring (March-May), Summer (June-August) Fall (September-November), Winter (December-Febuary)) from the date (month) associated with each GPS point.

Statistical Methods

Due to the intensive longitudinal nature of the dataset, we explored the possibility of nonlinear associations between mobility-based greenness exposure and physical activity using Generalized Additive Mixed Models (GAMM). We accounted for repeated measures within the same participant using a random intercept for participant. We fit NDVI using natural cubic splines with three knots using the mgcv package in R 4.1 to account for possible non-linearity. We adjusted models for the *a priori* selected confounders listed above. We specified an autoregressive correlation structure due to the repeated-measure, longitudinal nature of the data.

Effect Measure Modification

We assessed the presence of effect measure modification through models stratified on quartiles of walkability and nSES, median age (<35 years vs. ≥ 35 years), race (white vs. non-white participants), median temperature (<15.11C vs. ≥ 15.11C), precipitation (<0.01mm, >=0.01mm), weekday vs. weekend, region (Northeast, Midwest, South, West) and season (Spring, Summer, Fall, Winter). To test the potential statistical significance ($p<0.05$) of effect modification we included multiplicative interaction terms.

Sensitivity Analyses

To address epidemiologic and geographic biases, we conducted four sensitivity analyses to test the robustness of our analyses. Figure 3.3 provides a visual representation of the smartphone mobility data from participants used in this analysis and how we restricted these data for the sensitivity analyses described in detail below.

The first sensitivity analysis was designed to minimize selective daily mobility bias (Figure 3.3a) (Plue et al., 2020). In mobility studies with intensive longitudinal data, this bias functions as a confounder. The phenomenon, where it is difficult to discern whether an individual is passively exposed to a space or actively seeks it, is referred to as a 'selective (daily) mobility bias'. As researchers' understanding of this bias is relatively new, it is understudied. To assess the impact of selective daily mobility bias, we restricted activity space to GPS locations within a standard deviation ellipse — subject-specific standard deviation of the x-coordinates and y-coordinates from the mean center of that subject's points, to eliminate locations outside of an individual's normal range.

We focused our second sensitivity analysis on associations during time outside of work (Figure3.3b). We omitted time at work by geocoding workplace addresses at the time of study and restricting GPS location data to locations outside of a 160-meter radial buffer (0.1 mile). The size of this buffer was derived from hospital dimensions (the typical workplace of our study participants) as the majority of hospital sizes are thought to fall within this buffer size (*Insights from a Healthcare Architect's Journal*, 2019).

In our third analysis, we omitted datapoints that may include sedentary behaviors or driving (Figure 3.3c). We used timestamps in addition to GPS locations to estimate velocity, and restricted analyses to velocities that fell between walking and running (0.8 to 4 m/s) to obtain datapoints of active transport or recreating (Cruciani et al., 2018).

Lastly, we restricted our cohort to 208 participants who provided at least 12 hours of GPS location data daily on 5 unique days in two distinct sample periods (restricted analytical dataset) (Figure 3.1). This stringent criterion maximizes the amount of data per individual across time, to support the primary analysis findings with a robust intra-individual sample.

Results

Descriptive

Participants in the primary analytical cohort of the NHS3 mHealth Substudy resided in 42 out of 48 states across the contiguous US (Figure 3.2). After selecting participants who provided at least 8 hours of GPS data daily on 3 unique days and omitting main sleep periods, the primary analytic cohort included 337 participants with 639,364 observations (Figure 3.1). Each participant had on average 96.2 observations per day (SD 44.1) or approximately 16 hours per day and a total of 1,878 observations (SD 847.2) or approximately 313 hours during the 1-year study period (Table 3.1). Averaged across seasons, greenness exposure was 0.31 (SD 0.2) and participants took 7.0 (SD 14.9) steps per minute (Table 3.2). On average, we observed small variations by season for both the exposure and outcome with the spring months having the highest mean greenness exposure and highest average step count per minute (Table 3.2). Participants residing in the South had the smallest annual change in NDVI (Table 3.2). Similar seasonal variations were observed among the restricted dataset of 208 participants with 498,521 observations who provided at least 12 hours of GPS data daily on 5 unique days in two distinct sample periods (Supplemental Table 3.1).

Generalized Additive Mixed Models

We observed a statistically significant non-linear association between mobility NDVI and mean steps per minute (Figure 3.4). There were three distinct relationships with inflection points at 0.2 and 0.6 NDVI. Between NDVI values of 0 to 0.20, higher values of NDVI were very weakly associated with greater mean steps per minute (0.8 step more per 0.1 difference in NDVI). In contrast, between NDVI values of 0.2 to 0.6, higher values of NDVI were associated with fewer mean steps per minutes 1.0 fewer mean steps per minute per 0.1 difference in NDVI). Lastly at NDVI values above 0.6, higher values of NDVI were weakly associated with increased mean steps per minute with 0.5 step more per 0.1 increase in NDVI.

Stratified Analyses

We observed no evidence of effect modification by median age, race, neighborhood SES, neighborhood walkability, mean daily temperature and daily precipitation presence. Statistically significant effect modification by both season and region were observed. Seasonal stratified analyses revealed inverse associations in the Fall, Spring and Summer with no association in the Winter (Figure 3.5). Regional differences were observed across the strata, with the Northeast and Southern regions following the pattern of the main analysis (Figure 3.6). An inverse association was observed in participants residing in the Midwest and no association was observed in the West until NDVI was greater than 0.6, whereupon increasing values of NDVI were inversely associated with steps per minute (Figure 3.6).

Sensitivity Analyses

In sensitivity analyses attempting to restrict bias due to selective mobility, non-work location and restriction of cohort to those with more data, we did not identify any statistically significant differences from the primary analysis (Supplemental Figures 3.1, 3.2, 3.3). When we restricted our analyses to active transportation velocities that fell within walking and running (Figure 3.3c) as a transportation mode, we observed no association between NDVI value and steps per minute (Figure 3.7).

Discussion

Overall, we found a small negative association of 10-minute level mobility-based greenness with objectively measured mean step count per minute across the most frequent NDVI exposure range (0.2 up to 0.60). Our results expanded upon previous work on the association between greenness and physical activity at the residential level. Klompmaker et al. saw a positive relationship between residential NDVI and self-reported physical activity (Klompmaker et al., 2018) in a Dutch national health surveys, and Marquet et al. observed a positive association looking at weekly activity spaces and step counts among working adults in the US (Marquet et al., 2022b). However, our findings were the 10-minute scale and attempted to assess the momentary association between greenness and physical activity, whereas previous studies examined greenness exposure over a longer timescale. Our results were consistent with those of Persson et al. (2019), in which individuals moving to greener environments had a decrease in their physical activity. Furthermore, when we restricted our analysis to walking or running physical activity data only, we did not observe an association between smartphone mobility-based greenness and steps-per-minute. This suggests that green environments may be associated with sedentary behavior but when an individual conducts physical activity, their speed does not alter across levels of NDVI. This finding supports conclusions by James et al. (James, Hart, et al., 2017) who suggest walkability rather than greenness as a predominant driver of accelerometry based physical activity (Baobeid et al., 2021).

We observed evidence of effect modification across region and season. Regional differences drove associations with mild nonlinear positive associations between increases in NDVI and mean steps-per-minute observed in lowest and highest levels of NDVI in the South, and consistent negative associations observed in the Midwest. The South has the smallest seasonal change in NDVI, suggesting the positive association could be due to maintaining a green environment throughout the year.

Our results driven by fine-scale spatial and temporal data suggest that more research is needed to understand physical activity as a mechanism underlying how exposure to greenness is associated with improved health outcomes across various spatial and temporal scales, due to inconsistent results in the literature.

Our study has limitations. First, NHS3 is a cohort of predominantly upper-middle class white women nurses and as such these findings may have limited generalizability outside this population. Diverse cohorts should assess effect modification across race/ethnicity and SES to further confirm our findings. Secondly, step count as a proxy for physical activity remains another limitation, as it does not capture physical activity from weight-lifting, cycling, gardening, or swimming. However, most of the US and NHS3 participants record walking as the primary source of physical activity (CDC, 2013).

Our study also had a number of strengths. First, we were able to utilize a time-variant mobility greenness measure at 30m resolution, which enabled us to identify the quantitative value of greenness at a precise moment better addressing the exposure of interest. The intensive longitudinal spatial and temporal data allowed us to quantify momentary greenness exposure and physical activity at the minute-level and conduct several analyses examining seasonal trends and potential confounders or effect modifiers of the association. Second, utilizing an objective physical activity metric instead of self-reported physical activity reduced the likelihood of recall bias in our study compared to studies that used self-reported measures of physical activity. Lastly, as the mHealth Substudy was nested within the larger NHS3 cohort, we obtained high quality data from participants and covariate data prior to collection of exposure or outcome, reducing the likelihood of misclassification.

As environmental data becomes easier to access in mass quantities, it is essential that we prioritize real time exposure data. Environmental epidemiology too often ignores consequences of the uncertain geographic context problem and defining the extent of the exposure in question (Hooper et al., 2013; Kwan, 2012a; Spiegelman, 2010). By linking fine scale spatial and temporal greenness and physical activity data, we attempt to address critical gaps in the literature and look holistically at contextual environmental exposures beyond the residential environment.

In conclusion, we did not observe higher levels of physical activity in greener locations in this intensive longitudinal spatial temporal analysis. Rather, the association was nonlinear in nature and across most frequent exposure distributions, greener locations were observed to be associated with fewer steps-per-minute.

Table 3.1. Characteristics of the 2018-2020 Nurses' Health Study 3 mHealth Substudy including the primary analytic dataset population and observations (N = 337, n = 639,364), and secondary analytic dataset population and observations (N = 208, n = 498,521)

Variable	**Primary Analytic Dataset Population (n=337)**	**Secondary Analytic Dataset Population (n=208)**
	% (N)/ Mean (SD)	% (N)/ Mean (SD)
Age		
Continuous	36.0 (7.3)	26.0 (7.0)
Race		
White	94.1% (317)	92.3% (192)
Black	1.8% (8)	2.9% (6)
Asian	0.1% (2)	1.0% (2)
Mixed Race	1.2% (4)	1.0% (2)
Other	1.8% (6)	2.9% (6)
Ethnicity		
Hispanic	4.2% (14)	3.9% (8)
Married		
Yes	61.4% (207)	61.1% (207)
No	38.6% (130)	38.9 (81)
Advanced Degree		
Yes	73.9% (249)	75.0% (156)
No	26.1% (88)	25.0% (52)
Employment		
Yes	96.7% (319)	97.6% (203)
No	3.3% (11)	2.4% (5)
	Primary Analytic Dataset Mobility Observations (n=639,364)	**Secondary Analytic Dataset Mobility Observations (n=498,521)**
Walkability		
Mean	-0.02 (2.5)	-0.1 (2.4)
Neighborhood SES		
Mean	1.6, (3.3)	1.7 (3.3)
Temperature		
Mean	15.1 (10.0)	15.0 (10.0)
Precipitation		
Mean	3.4 (9.0)	3.3 (8.9)
Greenness		
Mean	0.3 (0.2)	0.3 (0.2)
Seasonality		
Fall	26.3% (167,871)	27.0% (134,660)
Winter	19.9% (127,860)	20.8% (103,797)
Spring	21.3% (136,117)	21.1% (105,248)
Summer	32.4% (207,496)	31.1% (154,816)

Table 3.2. Participant Greenness and Physical Activity Distributions Across Seasons in the 2018-2020 Nurses' Health Study 3 mHealth Substudy primary analytic dataset (N=327)

	N	Steps/Min Mean (SD)	**Steps/Min Min, Max**	**Greenness Mean (SD)**	**Greenness Min Max**
Total Participants	337	7.04 (14.93)	0.00, 263.78	0.31 (0.21)	0.00, 0.84
Fall	277	6.76 (14.62)	0.00, 181.00	0.27 (0.20)	0.00, 0.82
Winter	252	6.60 (14.25)	0.00, 183.13	0.21 (0.15)	0.00, 0.73
Spring	202	7.43 (15.44)	0.00, 219.00	0.37 (0.20)	0.00, 0.84
Summer	283	7.27 (15.19)	0.00, 263.78	0.37 (0.21)	0.00, 0.84

Figure 3.1. Study participant flow diagram for the Nurses' Health Study 3 mHealth Substudy and restriction criteria for primary analytic dataset and secondary analytic dataset for cohort population (N) and GPS mobility observations (n).

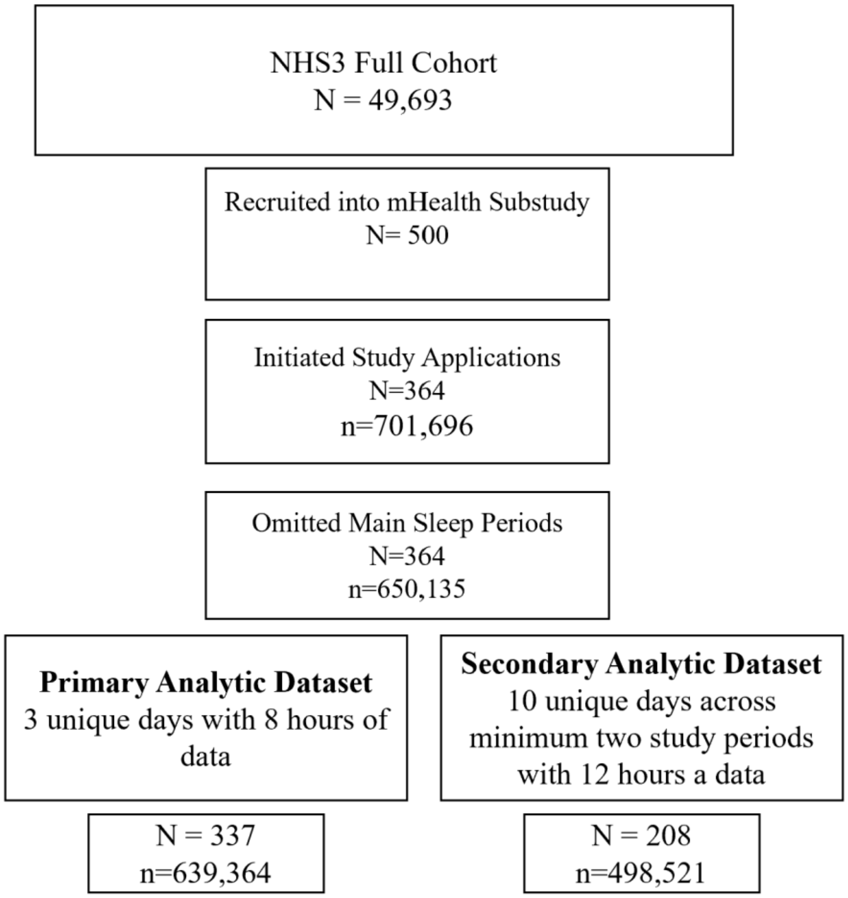


Figure 3.2. The Nurses' Health Study 3 mHealth Substudy participants residential locations across the contiguous US, 2018.

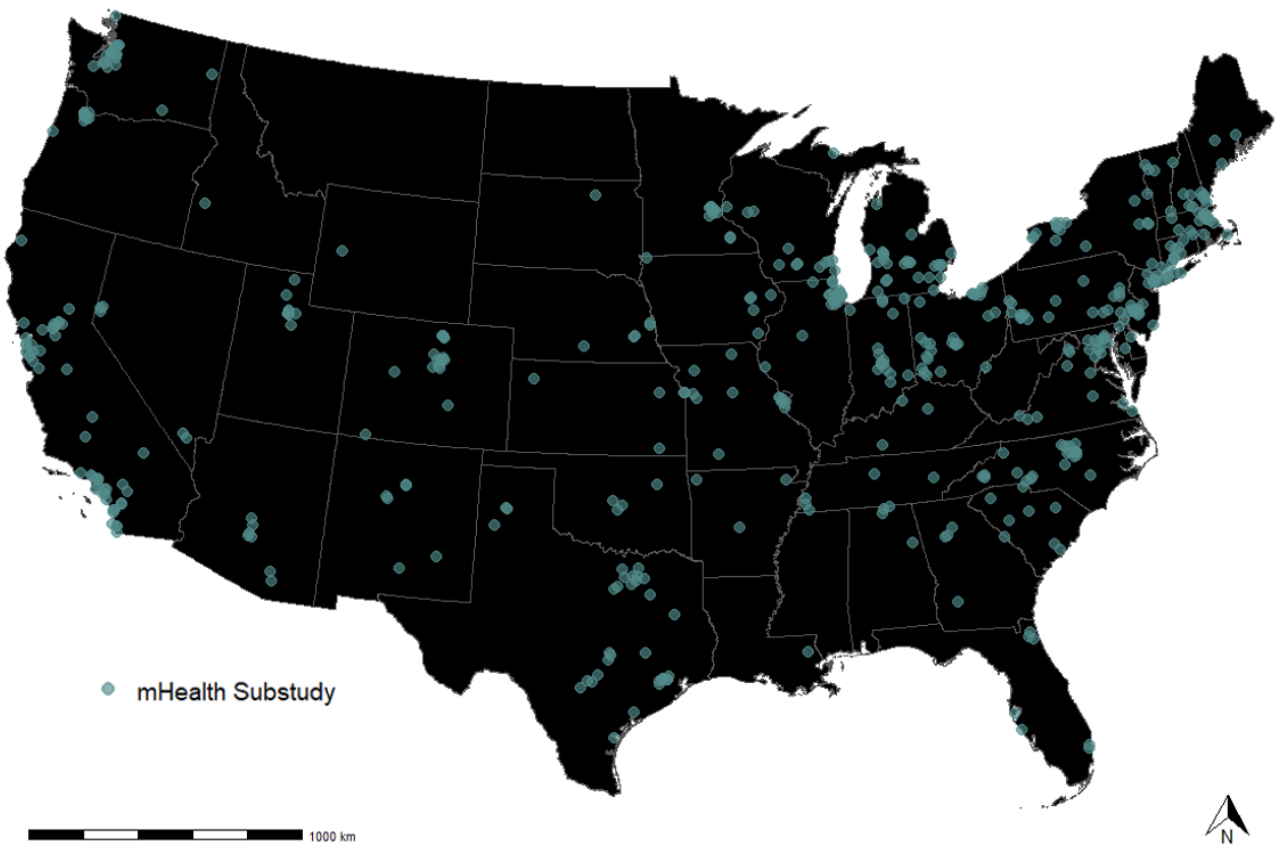


Figure 3.3: Three panel exposure map*: a) GPS mobility data over traditional residential buffers of 270m and 1230m and a selective daily mobility bias restriction criterion, b) workplace omitted GPS mobility over traditional residential buffers and c) active transport (walk to run velocities) GPS mobility metrics of exposure over traditional residential buffers. *This figure does not represent participant data. Data were obtained via the author's personal data collection.

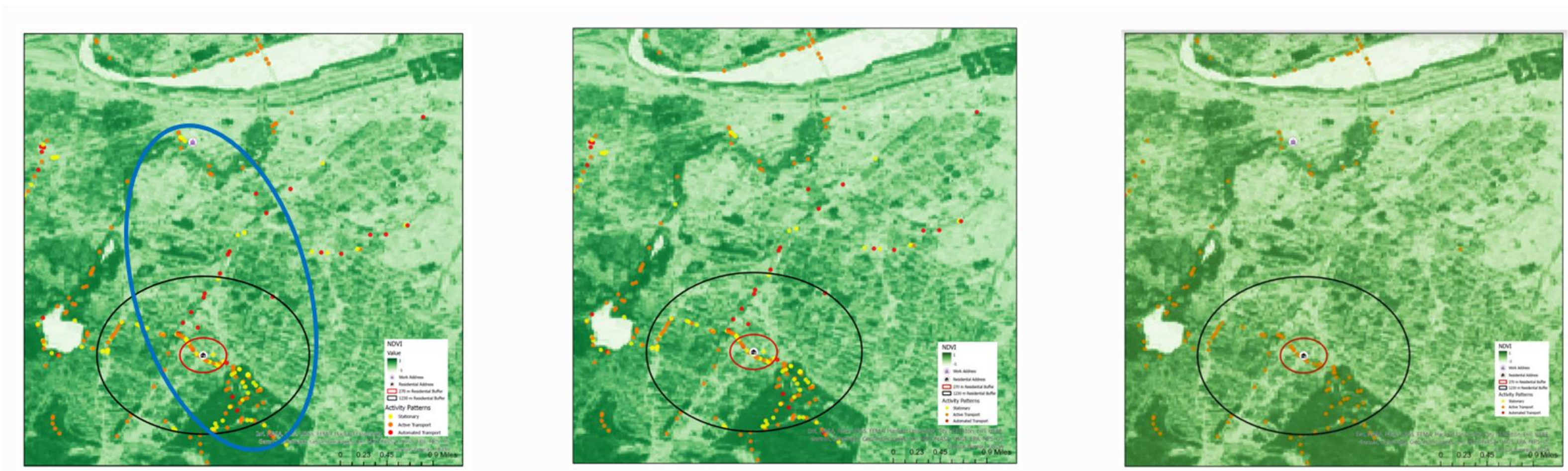
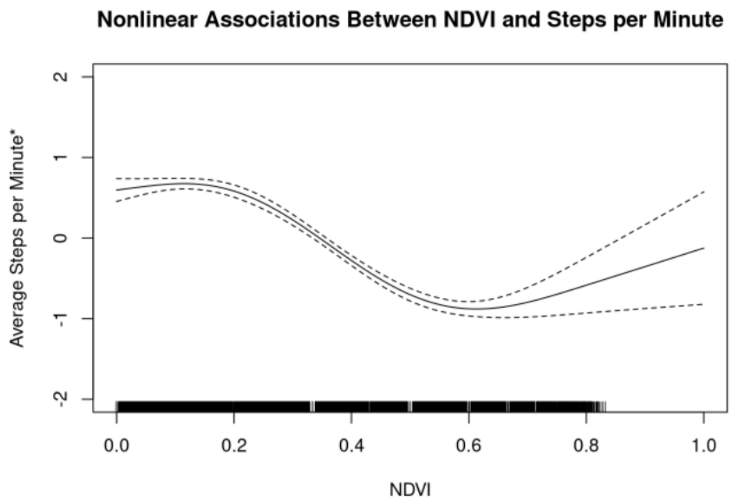
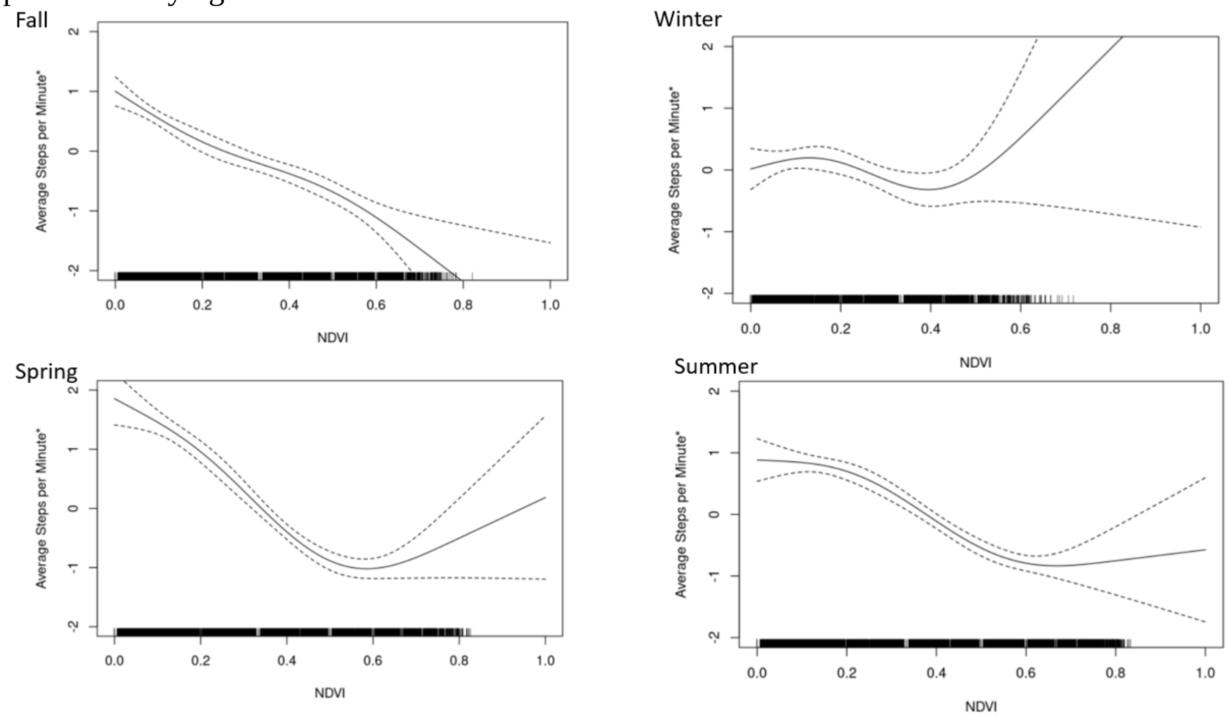


Figure 3.4. Associations[a] between NDVI[b] and average steps per minute across a 10-minute period



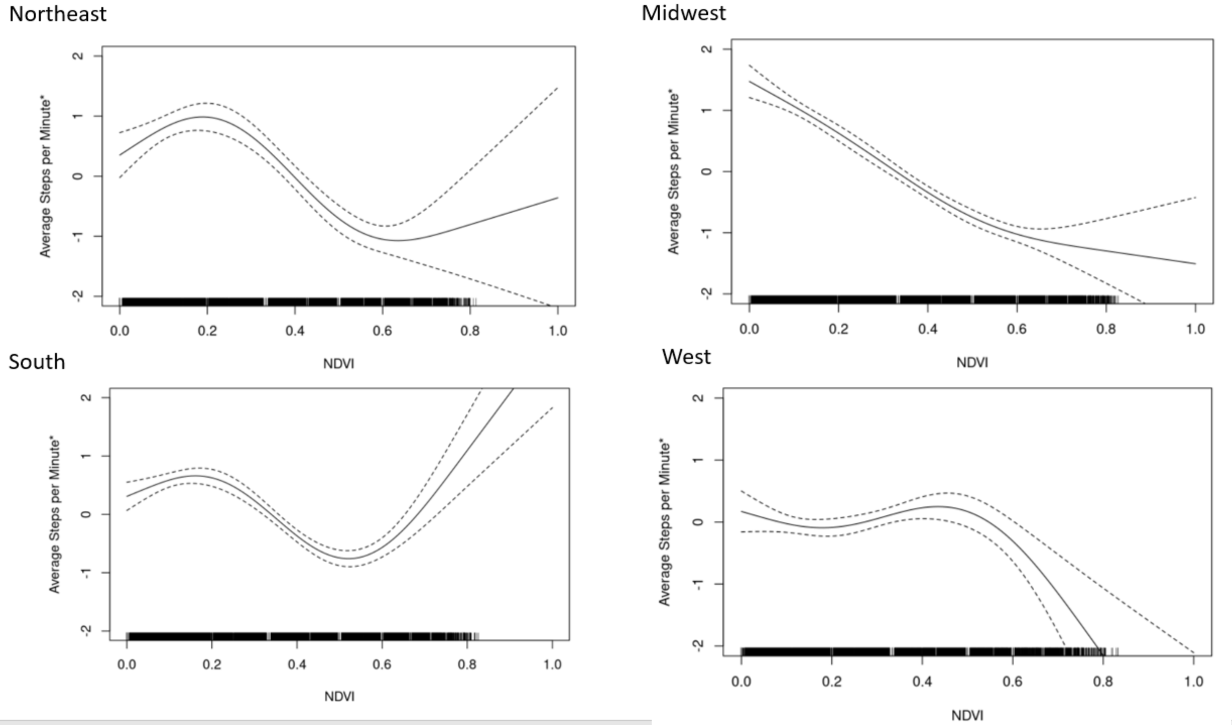
[a] Controlling for age (years; continuous), socioeconomic status defined as: education level (masters in nursing or higher; binary), and marital status (never [never married]]/ever [married, widowed, divorced]; binary), and area-level measures of neighborhood socioeconomic status (z-score; quartiles), walkability (z-scores; quartiles), mean daily temperature (Celsius; quartiles), daily precipitation (millimeters; binary), season and Census region in the 2018-2020 Nurses' Health Study mHealth Substudy.
[b] NDVI values below 0 represent water, ~0 represent rocks and bare soil including concrete, and values ~0.6-0.8 represent temperate and tropical forests.
[*] Average steps per minute across each ten-minute collection period.

Figure 3.5. Associations[a] between NDVI[b] and and average steps per minute across a 10-minute period stratfiying on season



[a] Controlling for age (years; continuous), socioeconomic status defined as: education level (masters in nursing or higher; binary), and marital status (never [never married]]/ever [married, widowed, divorced]; binary), and area-level measures of neighborhood socioeconomic status (z-score; quartiles), walkability (z-scores; quartiles), mean daily temperature (Celsius; quartiles), daily precipitation (millimeters; binary), and Census region in the 2018-2020 Nurses' Health Study mHealth Substudy.
[b] NDVI values below 0 represent water, ~0 represent rocks and bare soil including concrete, and values ~0.6-0.8 represent temperate and tropical forests
*Average steps per minute across each ten-minute collection period.

Figure 3.6. Associations[a] between NDVI[b] and and average steps per minute across a 10-minute period stratifying on region

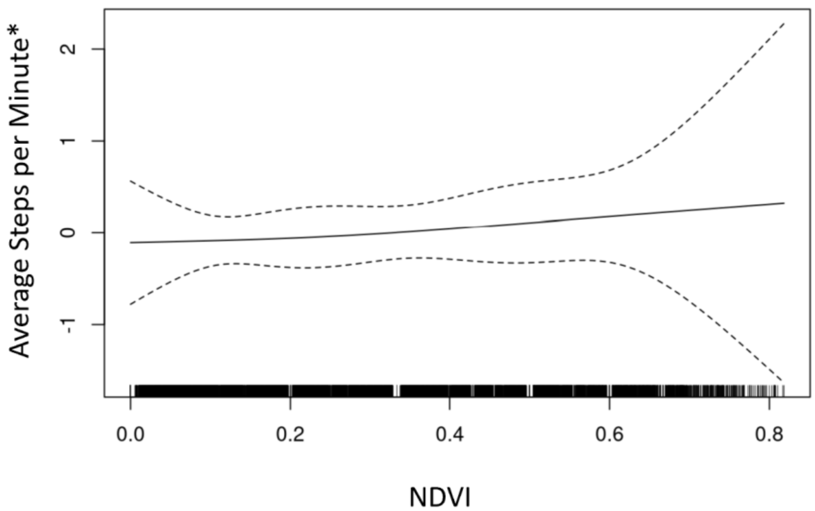


[a] Controlling for age (years; continuous), socioeconomic status defined as: education level (masters in nursing or higher; binary), and marital status (never [never married]]/ever [married, widowed, divorced]; binary), and area-level measures of neighborhood socioeconomic status (z-score; quartiles), walkability (z-scores; quartiles), mean daily temperature (Celsius; quartiles), daily precipitation (millimeters; binary), and season in the 2018-2020 Nurses' Health Study mHealth Substudy.

[b] NDVI values below 0 represent water, ~0 represent rocks and bare soil including concrete, and values ~0.6-0.8 represent temperate and tropical forests

*Average steps per minute across each ten-minute collection period.

Figure 3.7. Associations[a] between NDVI[b] and and average steps per minute across a 10-minute period, restricting on active transportation (walk to run velocity) GPS mobillity data



[a] Controlling for age (years; continuous), socioeconomic status defined as: education level (masters in nursing or higher; binary), and marital status (never [never married]]/ever [married, widowed, divorced]; binary), and area-level measures of neighborhood socioeconomic status (z-score; quartiles), walkability (z-scores; quartiles), mean daily temperature (Celsius; quartiles), daily precipitation (millimeters; binary), season, and Census Region in the 2018-2020 Nurses' Health Study mHealth Substudy.
[b] NDVI values below 0 represent water, ~0 represent rocks and bare soil including concrete, and values ~0.6-0.8 represent temperate and tropical forests
*Average steps per minute across each ten-minute collection period.

Chapter 4: Minute Level Smartphone Derived Exposure to Walkability and Consumer Wearable Derived Physical Activity in a US Cohort of Women

Target Journal:
Target First Draft Date:

Grete E. Wilt[1], Charlotte Roscoe[1,2,3], Cindy R. Hu[1], Brent A. Coull[1,4], Jaime E. Hart[1,2], Steven Gortmaker[5], Francine Laden[1,2], Peter James[1,6]

1. Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, United States
2. Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, United States
3. Division of Population Sciences, Dana Farber Cancer Institute, Boston, Massachusetts, United States
4. Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, United States
5. Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, United States
6. Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, Massachusetts, United States

Introduction: Walkable environments, referring to the density of people and businesses, have been linked to increased physical activity and improved chronic health outcomes. However, most studies measure exposure based on residential address alone, instead of considering exposures at all locations where people spend time.

Methods: We quantified associations of minute level GPS-based walkability exposure with accelerometry-measured activity among participants in Nurses' Health Study 3 Mobile Health (mHealth) Substudy from 2018-2020 in the US. The mHealth Substudy consisted of 337 participants who undertook 7-day sampling periods four equidistant times across a year, to capture potential seasonal variability. The walkability index was calculated from summed z-score of intersection density, business density and population density at the Census tract level and was linked to GPS points every 10 minutes. Fitbit accelerometry data summarized physical activity in mean steps per minute for each 10-minute period. Generalized Additive Mixed Models with penalized splines were utilized to explore associations. We adjusted for socioeconomic status, weather, greenness, and individual factors as a priori confounders.

Results: Across the sample the mean walkability z-score was 0.2 (SD 3.1) with individuals taking on average 7.0 steps per minute (SD 14.9). As the majority of walkability scores fell between -1.7 to 7.0, we focused primarily on this relationship. Between walkability values of -1.7 to 7.0, a 1-point increase in walkability score was associated with an increase of 1.9 (95% CI; 1.0, 2.8) steps per minute.

Discussion: We utilized comprehensive mobility data at fine temporal and spatial scales to present novel estimates on the real time association between walkability and physical activity. These findings contribute to discussions surrounding adapting the built environment to improve human health.

Introduction

The majority of adults report walking as their primary form of physical activity (CDC, 2013; Ussery, 2017). Physical activity has numerous health benefits including longer lifespans, and reductions in heart disease, stroke, type 2 diabetes, depression, and specific cancers including colon and breast cancers (James, Hart, et al., 2017). Despite this, in 2017 only 54% of adults in the US met the guideline of 150 minutes of moderate aerobic physical activity per week. (CDC, 2019). The definition of moderate activity includes brisk walks, water aerobics, and bike rides on flat ground (CDC, 2022). Emerging research indicates physical activity patterns are associated with factors of the built and social environments. Features of the built environment such as building density, population density, and well-connected streets increase efficiency of reaching destinations and may create opportunities for physical activity (Carr et al., 2010; Frank et al., 2008; Humpel et al., 2002; Saelens et al., 2003). Environments friendly to physical activity are associated with higher physical activity levels among residents (Carr et al., 2010; Frank et al., 2008; Humpel et al., 2002; Saelens et al., 2003). However, associations have predominantly been observed in observational or cross sectional studies with static residence-based measures of exposure and self-reported physical activity (James, Hart, et al., 2017; Marquet et al., 2020; McCormack et al., 2017; Orstad et al., 2018; Roscoe et al., 2022). Other studies explored walkability and associations with objective physical activity (James, Hart, et al., 2017; Rundle et al., 2016) highlighting the need to expand mobility-based built environment research. Studies are consistent in reporting positive associations between walkability and physical activity.

Traditional residential exposures of walkability have failed to consider typical human movement patterns or a*ctivity space*; (Brokamp et al., 2016; Kwan, 2019) a term used to describe the set of locations with which a person has direct contact during day-to-day activities (Chaix et al., 2012, 2013; Kwan, 2012a, 2012b, 2019; Park & Kwan, 2017; Perchoux et al., 2016). To assess if an individual’s physical activity increases in more walkable areas we used smartphone and wearable technology similar to those used in studies by James and Rundle (James, Hart, et al., 2017; Rundle et al., 2016).

The aim of this study was to quantify associations of daily smartphone GPS mobility-based walkability exposure with minute level wearable accelerometry data among participants in the Nurses’ Health Study 3 (NHS3) mobile Health (mHealth) Substudy. We hypothesized that higher 10-minute smartphone mobility-based walkability exposure would be associated with higher 10-minute level mean steps, even after consideration of potential confounders including individual and neighborhood-level socioeconomic status, greenness, seasonality, and region.

Methods

Population

Nurses' Health Study 3 (NHS3)

NHS3 is an ongoing open-enrollment prospective cohort of nurses and nursing students living in the US or Canada that began in 2010. Study eligibility required participants to be a registered nurse, licensed practical/vocational nurse, or nursing student and to be born on or after January 1, 1965. At the time of selection for the Mobile Health mHealth Substudy there were 49,693 participants enrolled in NHS3. Participants complete web-based questionnaires on lifestyle and medical characteristics and update their residential address every six months. The response rate for participants who have completed two or more questionnaires is above 80% (Chavarro et al., 2016; Gaskins, Rich-Edwards, Lawson, et al., 2015; Gaskins, Rich-Edwards, Missmer, et al., 2015; Mooney & Garber, 2019).

NHS3 mHealth Substudy

The NHS3 (mHealth) Substudy enrolled 500 NHS3 participants (Figure 4.1). The substudy began enrollment in March 2018 and data collection was completed in February 2020 with participants residing in 42 of the 48 contiguous states during the data collection period (Figure 4.2). To be eligible for the mHealth Substudy, participants had to be aged 21 or older on March 12, 2018 and demonstrate adherence to questionnaire completion by providing information on height, weight, physical activity, and sleep in prior NHS3 questionnaires. Participants with a doctor-diagnosed sleep disorder were not eligible because the study aimed to prospectively examine impacts of various lifestyle risk factors on sleep disturbance and Fitbit wearables have reduced accuracy in these populations (Fore et al., 2020).

Data collection methodology is described in detail elsewhere (Fore et al., 2020). In brief, mHealth participants downloaded a custom smartphone application on their personal smartphones and wore a consumer-wearable fitness tracker (Fitbit™ Charge HR, Fitbit™ Charge 2 or Fitbit™ Charge 3) for seven-day sampling periods every three months for a year from enrollment. Consistent with other mobility studies (Marquet et al., 2022b), we conducted a 7-day protocol to capture behaviors and exposures in a time frame that should include work and nonwork days. We acquired GPS location data at ~10-minute intervals throughout the 7-day sampling period through a mobile phone application. We developed eligibility criteria for inclusion in our analyses. In the primary analyses, we included participants who provided at least eight hours of GPS data on three unique days over the entire study enrollment. Additionally, using Fitbit™-derived sleep data, we omitted daily main sleep periods from the dataset under the assumption that physical activity does not occur during sleep periods, and restricted to participants who provided at least 8 hours of GPS data daily, during time they were awake, across 3 unique days of the total study period (Figure 4.1).

Exposure

We defined neighborhood walkability as a composite score of intersection density calculated from 2019 Tiger/Line shapefiles of all roads with interstates removed (Bureau, 2020), population density from 2019 ACS population data (*Explore Census Data*, 2020) and 2018 business density data from Infogroup US Historical Business Data (Infogroup, 2020). Variables were z-standardized at the tract for the entire 2010 U.S. Census. We summed the z-scores for each component variable to create a neighborhood walkability index. Higher scores indicate more walkable areas—referring to density of people, streets, and business activity (Figure 4.3). As such parks are considered not very walkable. We joined these data based on date and location of each 10-minute aggregated measure for a mobility-based value of neighborhood walkability data.

Outcome

Accelerometry data from Fitbit wearable devices (Fore et al., 2020) were used to summarize physical activity in mean steps per minute for the approximate 10-minute interval after each pair of GPS coordinates. This metric is valid (Diaz et al., 2016) and preferable to raw step counts because averages will not fluctuate if fine scale missingness occurs and has been used as a measure of physical activity in previous studies (Armstrong et al., 2019; Yuenyongchaiwat, 2016).

Covariates

Potential confounders were identified a priori. These included individual participant measures of age (years; continuous), socioeconomic status defined as: education level (masters in nursing or higher; binary), and marital status (never [never married]]/ever [married, widowed, divorced]; binary) obtained from the initial questionnaire administered to the full NHS3 cohort. Area-level measures of neighborhood socioeconomic status (z-score; quartiles), greenness (NDVI; quartiles), mean daily temperature (Celsius; quartiles), daily precipitation (millimeters; binary), season and Census region were obtained during the course of the Substudy.

We measured neighborhood Socioeconomic Status (nSES) using a composite score of 7 Census tract level variables representing domains that have been previously associated with health outcomes, including education, employment, housing, wealth, racial composition, and population density (DeVille, n.d.). Variables were carried forward from the prior U.S. Census (2010) and each variable was z-standardized. We summed the z-scores for each variable to create a nSES score. Higher scores indicated higher nSES (i.e. less socioeconomic deprivation). We joined quartiles of nSES score using the location of each 10-minute GPS point to create a mobility-based nSES score at the same temporal resolution as the greenness exposure assessment.

We utilized the Normalized Difference Vegetation Index (NDVI) at a 30m spatial resolution as a proxy measure for greenness exposure, spatially joined to the GPS locations. NDVI measures the reflection in the near infrared (NIR) spectrum minus the reflection in the red range of the spectrum divided by those measures added together, identifying the amount of vegetation corresponding to the minimal difference between the NIR and red reflectance bands. This index ranges from -1 to 1 with higher numbers indicating more green vegetation. NDVI values below 0 represent water, near 0 represent rocks and bare soil including concrete and values near 0.6-0.8 represent temperate and tropical forests (Klompmaker et al., 2018). We joined these data based on date and location of each 10-minute repeated measure for a mobility-based value of greenness expressed as quartiles for analysis purposes.

We obtained daily mean temperature and precipitation data at 800m spatial resolution for the study period (2018-2020) from Parameter-elevation Regression on Independent Slopes Model (PRISM) (Luzio et al., 2008). We joined PRISM variables by date and GPS coordinates to create mobility-based measures of temperature and precipitation. We classified daily mean temperature into quartiles and dichotomized precipitation to any precipitation/no precipitation.

We defined the Census region of each GPS point as one of 4 census regions (Northeast, Midwest, South, West), and derived season (Spring, Summer Fall, Winter) from the date (month) of each GPS point.

Statistical Methods

We explored nonlinearity and accounted for repeated measures through Generalized Additive Mixed Models (GAMM). Natural cubic splines were fit with three knots using the mgcv package in R 4.1 to account for non-linearity. We adjusted models for the a priori confounders listed above. We specified an autoregressive correlation structure due to the longitudinal nature of the data (Table 4.1).

Effect Measure Modification

We assessed the presence of effect measure modification through models stratified on quartiles of neighborhood SES and NDVI, median age (<35 years vs. >= 35 years), race (white vs. non-white participants), median temperature (<15.11C vs. >= 15.11C), precipitation (<0.01mm, >=0.01mm), region (Northeast, Midwest, South, West) and season (Fall, Winter, Spring, Summer). We examined p for interaction and determined significance at $p<0.05$.

Sensitivity Analyses

We conducted four sensitivity analyses to test the robustness of our analyses to potential sources of bias. Figure 4.4 provides a visual representation of the fine scale smartphone mobility data from participants used in this analysis and how we restricted these data for further sensitivity analyses described in detail below.

The first sensitivity analysis was designed to minimize selective daily mobility bias (Figure 4.4a) (Plue et al., 2020). This bias may function as a confounder in mobility studies with intensive longitudinal data. This phenomenon, in which it is difficult to discern whether an individual is passively exposed to a space or actively seeks it, is referred to as a ‘selective (daily) mobility bias’. Researchers’ understanding of this bias is relatively new, and as a result it is understudied. To assess the impact of this bias on our results, we restricted activity space to GPS locations within a standard deviation ellipse — subject-specific standard deviation of the x-coordinates and y-coordinates from the mean center of that subject’s points, to eliminate locations outside of an individual’s normal range.

The second sensitivity analysis focused on associations during leisure time (Figure 4.4b). We omitted time at work by geocoding workplace addresses at the time of study and restricting GPS location data to locations outside of a 160-meter radial buffer (0.1 mile). The size of this buffer was derived from hospital dimension (the typical workplace of our study participants) as the majority of hospital sizes are thought to fall within this buffer size (*Insights from a Healthcare Architect's Journal*, 2019).

We omitted datapoints that may include sedentary behaviors or driving for the third sensitivity analysis (Figure 4.4c). We used timestamps in addition to a GPS location to estimate velocity between each mobility datapoint, and restricted analyses to velocities that fell between walking and running (0.8 to 4 m/s) to obtain datapoints of active transport or recreating (Cruciani et al., 2018).

Lastly, we performed a sensitivity analysis which restricted our cohort to 208 participants with at least 12 hours of GPS location data daily on 5 unique days in two distinct sample periods (restricted analytical dataset) (Figure 4.1) as Zenk et al. suggest stringent cut points to gain clarity on full activity space profile (Zenk et al., 2018). This stringent criterion maximizes the amount of data per individual, to support the primary analysis findings with a robust intra-individual sample.

Results

Descriptive

Participants in the primary analytical cohort resided in 42 out of 48 states across the contiguous US (Figure 4.2). After selecting participants who provided at least 8 hours of GPS data daily on 3 unique days and omitting main sleep periods, the primary analytic cohort included 337 participants with 639,364 observations (Figure 4.1). On average, each participant had 96.7 observations per day (SD 44.5) and a total of 1878 observations (SD 847.2) during the 1-year study period.

Averaged across study periods, mobility-based neighborhood walkability exposure was 0.2 (SD 3.1) and participants took on average 7.0 steps per minute (SD 14.9). We observed small variations for both exposure and outcome across seasons. The highest average step count per minute occurred in the spring (Table 2.2). Similar seasonal variations were observed among the 208 participants with 498,521 observations who provided at least 12 hours of GPS data daily on 5 unique days in two distinct sample periods (Figure 4.1, Supplemental Table 4.1).

Generalized Additive Mixed Models

We observed a statistically significant non-linear association between 10-minute level smartphone mobility-based walkability and aggregated 10-minute mean steps per minute (Figure 4.5). There were 3 distinct relationships with inflection points of walkability z-scores at 7 and 28. As the majority of walkability scores fell between -1.7 to 7.0, we focused primarily on this relationship. Between walkability values of -1.7 to 7.0, a 1-point increase in walkability score was associated with an increase of 1.9 (95% CI; 1.0, 2.8) steps per minute.

Stratified Analyses

Statistically significant effect modification by nSES, region and season were observed, using a $p<0.05$ for interaction (supplemental figures 4.1 - 4.3.). We observed strongest associations between walkability and steps-per-minute in the lowest nSES quartile: between walkability values of -1.7 to 7.0, a 1-point increase in walkability score was associated with an increase of 32.2 (95% CI; 5.8, 58.7) steps-per-minute in the lowest nSES group. Patterns in the Midwest and

Southern regions followed those in the main analysis. In the Northeast, however, there was an inverse association of walkability with step counts.

Lastly, step counts were positively associated with walkability in the winter, spring, and summer, but there was no association in the fall. We observed no evidence of effect modification by age, NDVI, median temperature or precipitation (Supplemental Figures S4.1, S4.2, S4.3).

Sensitivity analyses

When we restricted our analyses to velocities that fell within walking and running (Figure 4.4c) as a transportation mode, we observed no association between walkability z-score and steps per minute (Figure 4.6). Additional sensitivity analyses are explored in the supplement (Supplemental Figures S4.4, S4.5, S4.6).

Discussion

We observed a statistically significant non-linear association between mobility-based walkability and mean steps per minute. Much of the walkability exposure data are clustered around a score of 0-1 where we see a strong positive association between smartphone GPS based exposure to neighborhood walkability and objectively measured physical activity. Stratified analyses further confirmed these findings and suggested that these associations were strongest in areas with the lowest neighborhood socioeconomic status. These associations differed greatly by region with the Midwest and South mimicking the main findings. An inverse relationship observed in the West region, with higher walkability being associated with fewer steps per minute. Sensitivity analyses to understand non-work walkability and physical activity associations and the effects of selective daily mobility bias further confirmed an overall positive association between walkability and steps-per-minute. When restricting data to walk only datapoints, we observed a null association. This suggests while inactivity may take place more frequently in low walkability areas, when active, steps per minute do not differ across increasing walkability scores. This can be interpreted as more movement occurs in higher walkable areas but an increase of steps per minute does not occur after a participant is already in motion, regardless of location i.e. they do not walk faster or run in the most walkable areas.

These findings suggest that this population of female, predominately white nurses and nursing students record more steps-per-minute in more walkable areas. Our results are based on fine scale spatial and temporal data suggest that walkable areas encourage movement. The association between walkability and physical activity may be an important pathway toward improved health outcomes, as physical activity is often cited as a likely mediator for numerous proposed health benefits in relation to exposure to the built environment that are downstream.

Our study has limitations. First, NHS3 is a cohort of predominantly upper-middle class white women nurses and as such these findings may have limited generalizability outside this population. Despite non-traditional work schedules, we anticipated the majority of nurses would have some work-free days in a 7-day period; however, weekly self-report of work-time would provide greater insights into their routine. Diverse cohorts should assess effect modification across race and SES to further confirm these findings of minute level associations between greenness and physical activity. Second, step count as a proxy for physical activity does not capture physical activity from weight-lifting, cycling, gardening, or swimming. However, most of the US and NHS3 participants record walking as the primary source of physical activity (CDC, 2013).

Our study also has a number of strengths. First, we were able to utilize a time-specific exposure measure that captured walkability experienced in the moment of the outcome. This enabled us to identify the quantitative value of walkability at a precise moment, thus better addressing the exposure of interest. The intensive longitudinal spatial and temporal data allowed us to quantify momentary walkability exposure and physical activity at the minute-level and conduct several analyses examining seasonal trends and potential confounders or effect modifiers of the association. Second, utilizing an objective physical activity metric instead of self-reported physical activity reduced the likelihood of recall bias in our study. Third, we obtained high quality data from participants and had covariate data prior to exposure or outcome due to the nested design of the mHealth Substudy within the larger NHS3 study, reducing the likelihood of misclassification.

As built environmental data becomes easier to access in mass quantities, it is essential that we prioritize fine scale exposure collection. Environmental epidemiology too often ignores consequences of the uncertain geographic context problem, defined as the error from contextual environmental exposures existing without a set spatial boundary of influence, nd hence defining the extent of the exposure in question (Kwan, 2012a; Spiegelman, 2010). By linking fine-scale spatial and temporal greenness data to momentary physical activity information, we attempt to answer critical gaps in the research and look holistically at contextual environmental exposures beyond the residential environment. In conclusion, we observed increases in physical activity as mobility-based walkability measures increase. Further research on the joint effect of built environmental factors such as noise, air pollution, and light at night in tandem with walkable locations should be investigated as a potential promoter of physical activity.

Tables and Figures

Table 4.1. Characteristics of the 2018-2020 Nurses' Health Study 3 mHealth Substudy including the primary analytic dataset population and observations using participants with minimum 3 days of 8 hours of observations (N = 337, n = 639,364), and secondary analytic dataset population and observations using participants with minimum 10 days of 12 hours of observations (N = 208, n = 498,521)

Variable	**Primary Analytic Dataset Population (n=337)**	**Secondary Analytic Dataset Population (n=208)**
	% (N)/ Mean (SD)	% (N)/ Mean (SD)
Age		
Continuous	36.0 (7.3)	26.0 (7.0)
Race		
White	94.1% (317)	92.3% (192)
Black	1.8% (8)	2.9% (6)
Asian	0.1% (2)	1.0% (2)
Mixed Race	1.2% (4)	1.0% (2)
Other	1.8% (6)	2.9% (6)
Ethnicity		
Hispanic	4.2% (14)	3.9% (8)
Married		
Yes	61.4% (207)	61.1% (207)
No	38.6% (130)	38.9 (81)
Advanced Degree		
Yes	73.9% (249)	75.0% (156)
No	26.1% (88)	25.0% (52)
Employment		
Yes	96.7% (319)	97.6% (203)
No	3.3% (11)	2.4% (5)
	Primary Analytic Dataset Mobility Observations (n=639,364)	**Secondary Analytic Dataset Mobility Observations (n=498,521)**
Walkability		
Mean	-0.02 (2.5)	-0.1 (2.4)
Neighborhood SES		
Mean	1.6, (3.3)	1.7 (3.3)
Temperature		
Mean	15.1 (10.0)	15.0 (10.0)
Precipitation		
Mean	3.4 (9.0)	3.3 (8.9)
Greenness		
Mean	0.3 (0.2)	0.3 (0.2)
Seasonality		
Fall	26.3% (167,871)	27.0% (134,660)
Winter	19.9% (127,860)	20.8% (103,797)
Spring	21.3% (136,117)	21.1% (105,248)
Summer	32.4% (207,496)	31.1% (154,816)

Table 4.2. Participant walkability and physical activity distributions across seasons in 2018-2020 Nurses' Health Study 3 mHealth Substudy primary analytic dataset (N=337)

	N	Steps/Min Mean (SD)	Steps/Min Min, Max	Walkability Mean (SD)	Walkability Min Max
Total Participants	337	7.04 (14.93)	0.00, 263.78	0.24 (3.07)	-1.66, 52.85
Fall	281	6.76 (14.62)	0.00, 181.00	0.24 (3.02)	-1.66, 52.85
Winter	255	6.60 (14.25)	0.00, 183.13	0.24 (3.24)	-1.66, 52.85
Spring	205	7.43 (15.44)	0.00, 219.00	0.28 (2.79)	-1.66, 33.27
Summer	289	7.27 (15.19)	0.00, 263.78	0.25 (3.18)	-1.66, 44.60

Figure 4.1: Study Participant flow diagram for the nurses’ Health Study 3 mHealth Substudy and restriction criteria for primary analytic dataset and secondary analytic dataset for cohort population (N) and GPS mobility observations (n).

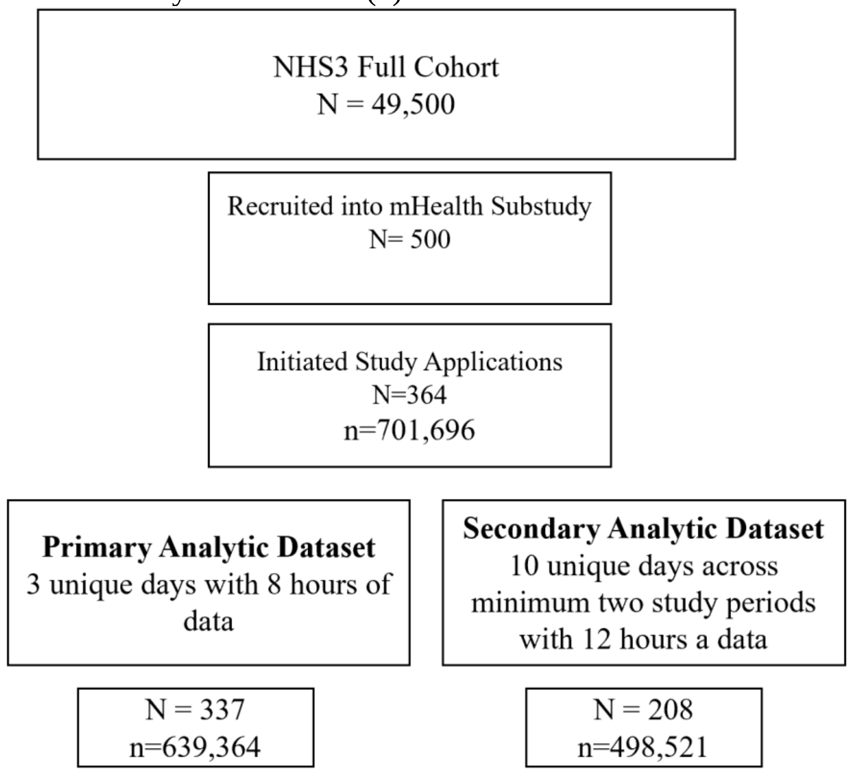


Figure 4.2. The Nurses' Health Study 3 mHealth Substudy participants residential locations across the contiguous US in 2018

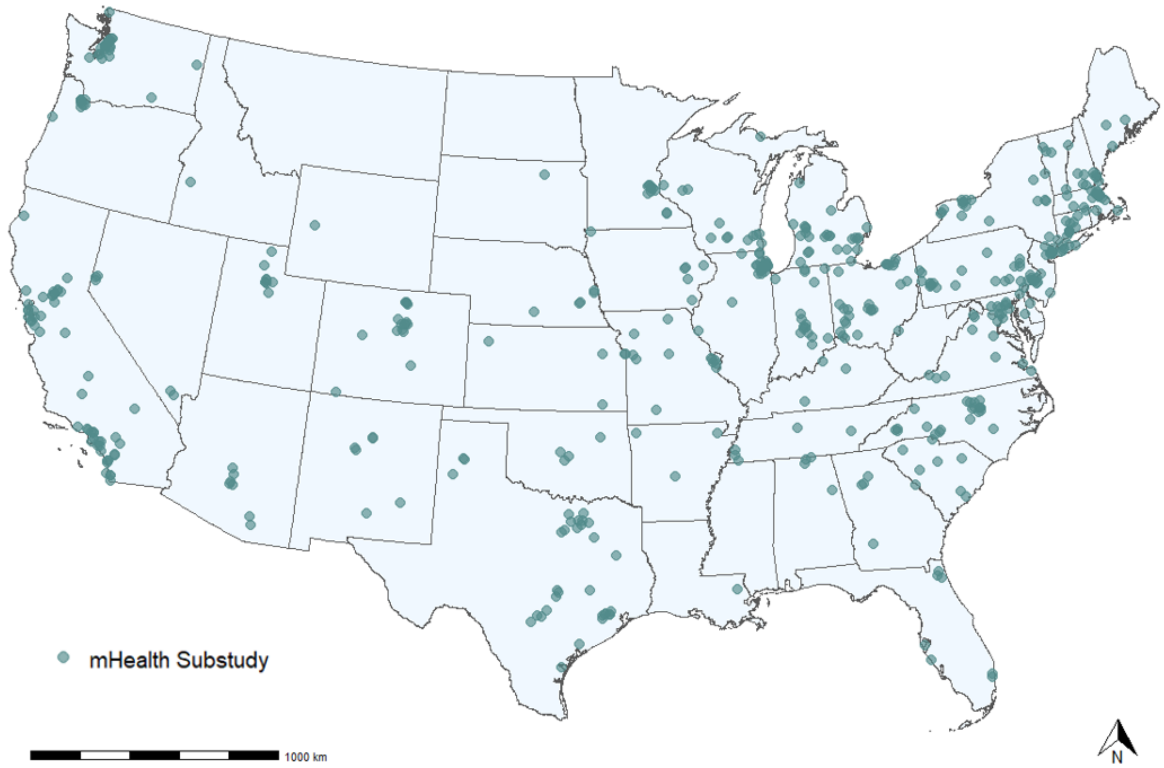


Figure 4.3. A) Boston, MA and B) Atlanta, GA. Population density is A) 6,517 people/km and B) 1465 people/km. Business density is A) 376 businesses/km and B) 108 business/km. Intersection density is A) 239 intersections/km and B)104 intersections/km. Walkability z-scores are 5 and 0.8 respectively.

A B

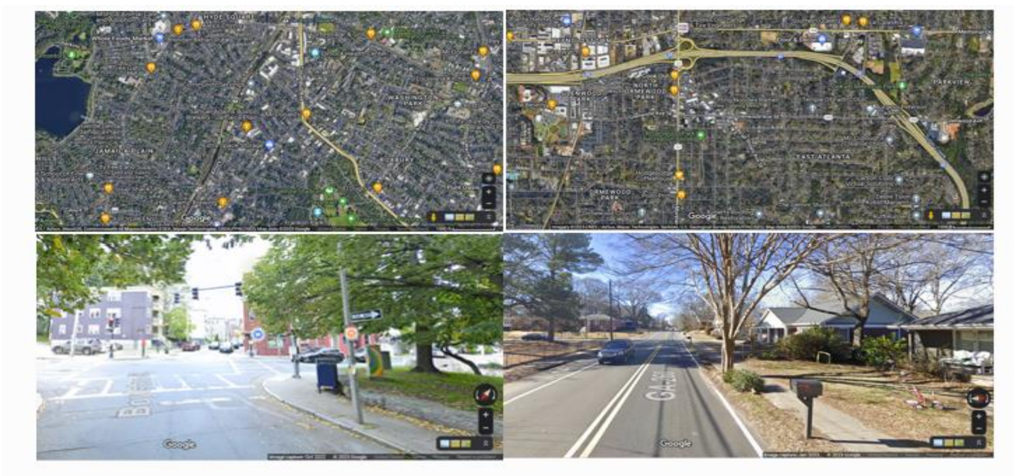


Figure 4.3: Three panel walkability exposure map*: a) GPS mobility data and selective daily mobility bias, b) workplace omitted GPS mobility and c) active transport (walk to run velocity) GPS mobility metrics. * This figure does not represent participant data. Data were obtained via the authors personal data collection.

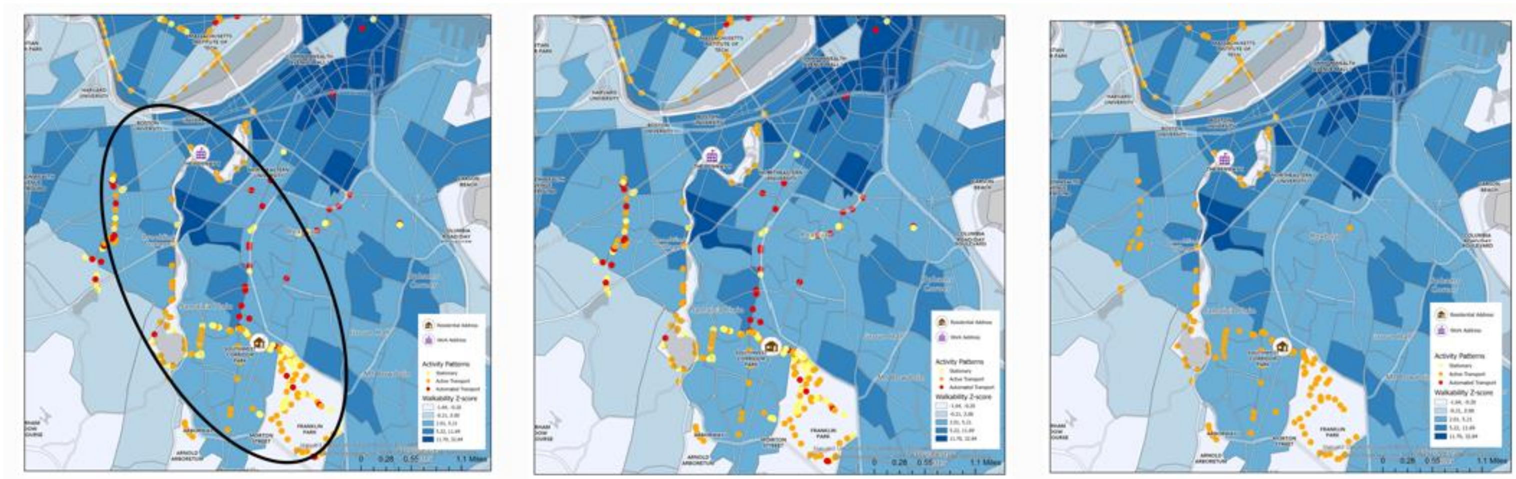


Figure 4.5. Nonlinear associations between walkability Z-Score and average steps per minute across a 10-minute period controlling for age (years; continuous), socioeconomic status defined as: education level (masters in nursing or higher; binary), and marital status (never [never married]]/ever [married, widowed, divorced]; binary), and area-level measures of neighborhood socioeconomic status (z-score; quartiles), greenness (NDVI; quartiles), mean daily temperature (Celsius; quartiles), daily precipitation (millimeters; binary), season and Census region in the 2018-2020 Nurses' Health Study mHealth Substudy.

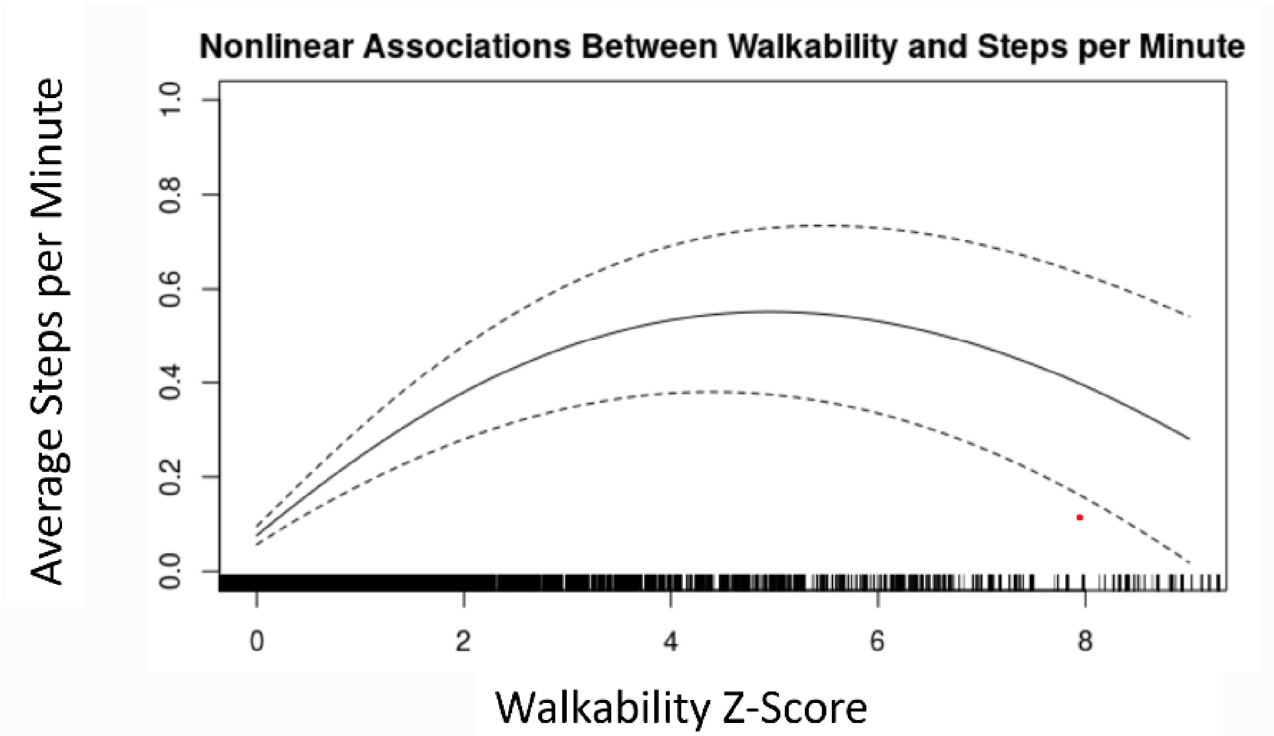
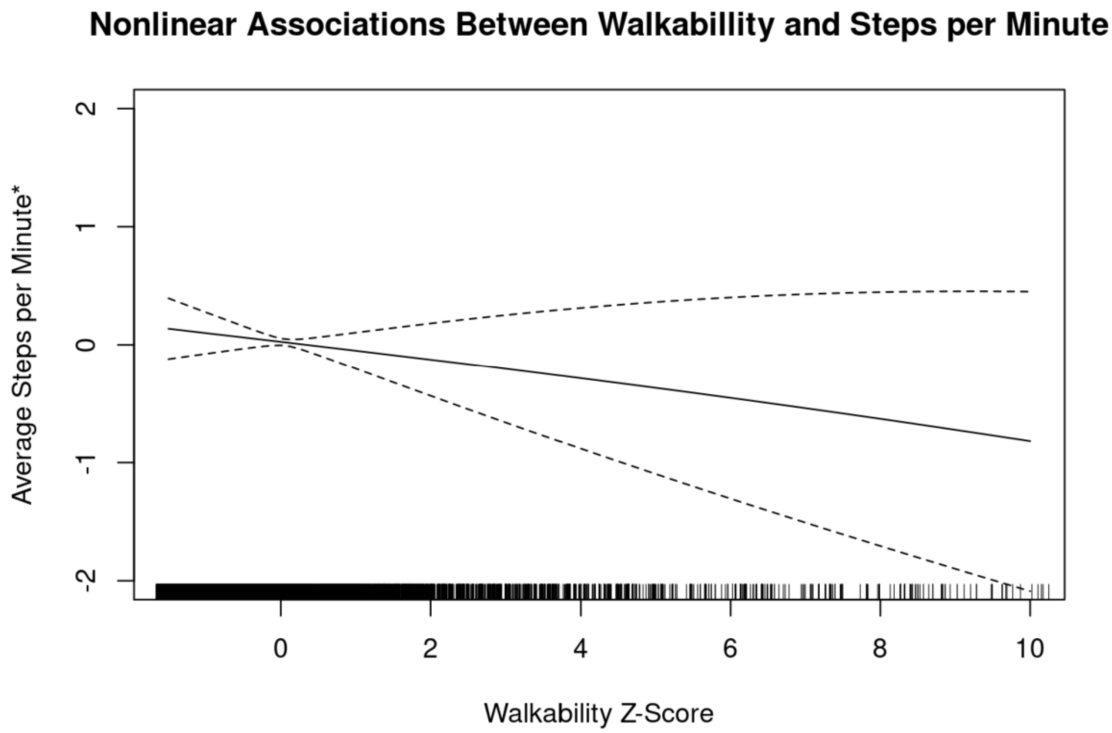


Figure 4.6. Associations between walkability and steps per minute restricted to active transportation (walk to run velocities) GPS mobility data, controlling for age (years; continuous), socioeconomic status defined as: education level (masters in nursing or higher; binary), and marital status (never [never married]]/ever [married, widowed, divorced]; binary), and area-level measures of neighborhood socioeconomic status (z-score; quartiles), greenness (NDVI; quartiles), mean daily temperature (Celsius; quartiles), daily precipitation (millimeters; binary), season and Census region in the 2018-2020 Nurses' Health Study mHealth Substudy.



Chapter 5.

Measurement Error Correction Using Smartphone Mobility Derived Association Between Walkability and Physical Activity in a US Cohort of Nurses

Grete E. Wilt[1], Charlotte Roscoe[1,2], Cindy R. Hu[1], Brent Coull[3], Jaime E. Hart[1,2]. Steven Gortmaker[4], Francine Laden[1,2], Peter James[1,5]

1. Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, Massachusetts
2. Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts
3. Department of Biostatics, Harvard T.H. Chan School of Public Health, Boston, Massachusetts
4. Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health, Boston, Massachusetts
5. Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, Massachusetts

Abstract

Background: Walkability exposures, often defined as population density, business density, and street connectivity that create opportunities for walking, often rely on residential-based measures that ignore exposure away from home, potentially leading to exposure misclassification.

Aim: We aimed to use smartphone GPS mobility-based walkability estimates from a validation study to correct for possible measurement error in associations between residence-based walkability exposure and self-reported physical activity in a US-based cohort.

Methods: We developed a Census tract-level walkability score utilizing 2010 US Census tracts and summing z-scores of 2018 TIGER/Line intersection density; 2018 Infogroup business density; and 2015-2019 5-year American Community Survey population density estimates. Our validation population was a subset of Nurses' Health Study 3 (NHS3) (n=337) who provided smartphone GPS data every 10-minutes across four 7-day sampling periods. Each GPS location was spatially joined with census tract walkability data and averaged across all sampling periods

to create our ‘gold standard’ walkability exposure. In the full NHS3 cohort (n=23,983), we spatially joined residential addresses to census tract walkability as our error-prone measure. Physical activity was derived from questionnaires in the full NHS3. We used standard regression calibration for linear models to produce error-corrected estimates and 95% confidence intervals for associations of walkability with physical activity in the full NHS3, adjusting for confounders.

Results: Participants (n=28,650) reported on average 273.2 minutes of walking, jogging, and running per week (SD = 257.0). In the subsample, participants had a mean residential-based walkability exposure of 0.0 (SD=2.6), and a mean mobility-based walkability exposure of 0.3 (SD=2.4). Each 1-unit SD difference in uncorrected residential-based walkability z-score was associated with 15.0-minute difference in time spent walking, jogging, and running per week (95% CI; 11.7, 18.4), whereas each 1-unit SD difference in measurement error-corrected residential-based walkability was associated with a 23.6-minute difference in time spent walking, jogging, and running per week (95% CI: 17.3, 29.8).

Conclusions: Our study indicates that residential-based estimates of walkability slightly underestimate associations between walkability and physical activity. Findings highlight the impact of exposure misclassification on epidemiological studies of the built environment and physical activity.

Introduction

Features of the built environment, including building density, population density and well-connected streets, have been associated with physical activity (Carr et al., 2010; Frank et al., 2008; Humpel et al., 2002; Saelens et al., 2003). These built environment features—commonly referred to as "walkability"—may affect the efficiency of reaching destinations and create opportunities for routine physical activity(National Academies of Sciences et al., 2019). The majority of studies on built environment exposures and physical activity focus on the residential address to define exposure (McCormack et al., 2017; Orstad et al., 2018; Roscoe et al., 2022), yet research suggests that individuals may spend less than 50% of their awake time at home (Zenk et al., 2019). Therefore, residence-based estimates of exposure likely contain substantial measurement error (Bhopal, 2016; Blair et al., 2007; Hart et al., 2015a; Kwan, 2012a, 2012b; Salvo et al., 2019; Suzuki et al., 2017; Zhao et al., 2018) In recent years, researchers have suggested using global positioning systems (GPS) data to generate activity spaces, a term used to describe the set of locations with which a person has direct contact during day-to-day activities (Perchoux et al. 2016), to identify environmental exposures beyond the home.

In this study we addressed the impact of measurement error due to the use of residence-based exposures on the relationship between the built environment and physical activity. We aimed to quantify associations of prospective residential-based walkability exposure with self-reported physical activity among participants in the Nurses' Health Study 3 (NHS3) cohort, while correcting for measurement error using data on mobility-based walkability exposure from a substudy of NHS3 participants (Fore et al., 2020). We hypothesized that greater error-corrected walkability exposure would be associated with higher self-reported physical activity minutes per week, even after consideration of potential confounders and error-correction would lead to stronger effect estimates. This study targets several gaps existing in environmental health research by accounting for high spatio-temporal resolution mobility-based exposure measures to correct for residential-based measures in associations with physical activity.

Methods

Population

Nurses' Health Study 3 (NHS3)

NHS3 is an ongoing open-enrollment prospective cohort of nurses and nursing students living in the US or Canada that began in 2010. Study eligibility requires participants to be a registered nurse, licensed practical/vocational nurse, or nursing student, and to be born on or after January 1, 1965. As of substudy collection (detailed below), N=34,477 participants had joined the cohort and completed the second questionnaire. Participants complete web-based questionnaires on lifestyle and medical characteristics on an individualized timeline and update their residential address every six months. (Chavarro et al., 2016; Gaskins, Rich-Edwards, Lawson, et al., 2015; Gaskins, Rich-Edwards, Missmer, et al., 2015; Mooney & Garber, 2019).

NHS3 mHealth Substudy

The NHS3 mobile Health (mHealth) Substudy was a pilot study that enrolled a subset of NHS3 participants (n = 500) (Figure 5.1, Table 5.1). We have previously published details of the substudy in Fore et al. (2020). Substudy began enrollment in March 2018. We completed data collection in February 2020. Participants had to be 21 years old, and have completed questions on height and weight, physical activity, and sleep on questionnaires prior to enrollment. Furthermore, participants were required to own an iPhone and live in the contiguous US to meet all requirements of substudy eligibility.

Once enrolled, participants downloaded a custom smartphone application on their personal smartphones and wore a consumer-wearable fitness tracker (Fitbit™) for seven-day sampling periods every three months for a year to capture seasonal variability in behaviors. We conducted a seven-day protocol to capture behaviors and exposures in a time frame that should include work and nonwork days, consistent with other mobility studies (James, Hart, et al., 2017; Marquet et al., 2020; Zenk et al., 2018),. We acquired smartphone location services GPS data at approximately 10-minute intervals for each day throughout the sampling period.

Figure 1 details the dataset's restriction criteria. There were 34,477 participants enrolled in the full NHS3 cohort through the second questionnaire. Of these individuals, 86.3% filled out the module 2 physical activity questionnaire (N=29,674). Every enrollee of mHealth validation cohort (N=337) had the physical activity portion of questionnaire 2 completed (Figure 5.1).

Walkability Exposure

We defined neighborhood walkability at the census tract level using the composite z-score of three variables: intersection density, population density, and business density. Three-way intersections were calculated from 2019 Tiger/Line shapefiles of all roads with interstates removed (Bureau, 2020). First road networks by county were merged into a national dataset. Next, using ESRI ArcPro coincident endpoints of lines were merged using the unsplit lines tool. Finally, on the unsplit lines feature, the ESRI ArcPro intersect tool was run to identify three-way intersections. From here we exported a shapefile for further analysis and calculated density using the area of the census tracts as a denominator. Population density was calculated from 2015-2019 ACS population data estimates (*Explore Census Data*, 2020) and business density was derived from 2018 Infogroup US Historical Business Data (Infogroup, 2020). Variables were z-standardized at 2010 US Census tract geographies. We summed the z-scores for each component variable to create a neighborhood walkability index. Higher scores indicate more walkable tracts (Figure 4.3).

GPS Mobility-Based Walkability

A common correction for measurement error uses data on both a gold standard or proxy gold standard and standard exposure and a measure assumed to have more error in the same population, or validation data. We used data from subjects in the NHS3 substudy (N=337) who provided smartphone GPS data every 10 minutes for at least 8 hours across at least three of the four 7-day sampling periods that captured seasonal variability (n=639,364 observations) as our validation subset. We spatially joined the GPS mobility data to the Census tract-level walkability index. Then we calculated the average walkability of an individual from walkability values based on their GPS locations across all study periods to create our 'gold standard' walkability exposure metric.

Residential-based Walkability

For the full NHS3 cohort, we used geocoded NHS3 residential addresses at the time of the physical activity questionnaire (Questionnaire 2) and linked the Census tract-level walkability index to the residential address as our error-prone measure of walkability exposure. For the validation study, walkability at the residential address at the time of substudy enrollment was also appended.

Outcome

The primary outcome was self-reported physical activity derived from validated physical activity questionnaires asked in NHS3 study Questionnaire 2. Questionnaire-based outcomes of physical activity included all active transport physical activity (walk, jog, and run combined) and secondary outcomes of walking only, low intensity physical activity only, vigorous physical activity only, moderate to vigorous physical activity, and jogging/running only. Response options were none; 0-0.5 hour/week; 1-2 hours/week; 3-4hours/week; 5-6 hours/week; 7-10 hours/week; and ≥11 hours/week. Using the mean time in the ranges above we calculated typical minutes per week spent participating in the following activities: all active transport physical activity (walk, jog, and run combined), walking only, low intensity physical activity only, vigorous physical activity only, moderate to vigorous physical activity, and jogging/running only. Following protocol from other NHS physical activity studies if at least one survey question from the physical activity questionnaire was filled out, we set remaining blank answers to zero (Bao et al., 2016).

Statistical Methods

We compared demographic factors and geographic location between the full NHS3 cohort and mHealth substudy to ensure transportability. We compared residential and GPS mobility-based estimates of walkability among the mHealth substudy participants using histograms that assessed the distribution of each exposure measurement, and Spearman's rank correlation to quantify the rank order correlation. We assessed disagreement between the two exposure measures in the substudy using Bland Altman agreement tests and plots, a simple method that evaluates bias between the mean differences in measures and estimates an agreement interval, within which 95% of the differences of the second measure compared to the first measure fall.

We used linear models to determine the association of residential-based measure of walkability with average minutes spent walking and running per week and calculated beta estimates and 95% confidence intervals. Utilizing the data from the mHealth substudy, we conducted standard regression calibration methods from Spiegelman et al. (Spiegelman, 2010; Thurston et al., 2005). We produced corrected estimates and 95% confidence intervals for associations of walkability with physical activity in the full NHS3, adjusting for confounders. To perform these analyses, we deployed the mecor package in R 4.04.

Confounders

Potential confounders identified *a priori* include individual participant measures of age (years, continuous), socioeconomic status defined as: education level (advanced degree, binary), and marital status (binary), and area-level measures of neighborhood socioeconomic status (z-score, quartiles), greenness (normalized difference vegetation index (NDVI), continuous), mean yearly temperature (Celsius, quartiles), and total yearly precipitation (dichotomized to <0.01mm, ≥0.01mm). We describe the confounders in detail below.

We obtain spatially invariant confounders (age, education level and marital status) from NHS3 baseline questionnaire predating enrollment in the substudy.

We evaluate Neighborhood Socioeconomic Status (nSES) with a composite score of seven Census tract level variables representing domains that have been previously associated with health outcomes including education, employment, housing, wealth, racial composition, and population composition. (*DeVille et al. 2022*). Variables were taken from the 2010 U.S. Census and were z-standardized and summed to create a nSES score linked to residential address at start of Questionnaire 2. Higher scores indicate higher nSES.

We develop focal statistics to measure annual NDVI at 270m residential buffers, which proxies visible distance (James et al., 2016). NDVI measures the reflection in the near infrared (NIR) spectrum minus the reflection in the red range of the spectrum divided by those measures added together, identifying the amount of vegetation corresponding to the minimal difference between the NIR and red reflectance bands. This index ranges from -1 to 1 with higher numbers indicating more green vegetation. NDVI values below 0 represent water, near 0 represent rocks and bare soil including concrete, and values near 0.6-0.8 represent temperate and tropical forests (Klompmaker et al., 2018). We link mean and maximum annual NDVI for the Questionnaire 2 completion year to the geocoded residential address of the participant.

We obtain monthly and annual mean temperature and total yearly precipitation data for study period (2010-present) at 800m spatial resolution from PRISM and spatial temporally join these data to the residential address points and year for Questionnaire 2 in NHS3.

Sensitivity analyses

We investigate associations with several secondary outcomes: walking only, low intensity physical activity only, vigorous physical activity only, moderate and vigorous physical activity, and jogging/running only.

Results

As of April 2022 (the time of the last NHS3 data extraction), 34,477 NHS3 participants had completed Questionnaire 2 and 86.2% percent of these participants filled out at least one physical activity question. After restriction criteria (Figure 5.1), there were 28,650 participants with physical activity data in NHS3 Questionnaire 2. Additionally, we observed no notable differences between demographic characteristics (age, and individual and neighborhood socioeconomic status) or geographic distribution in the full NHS3 cohort and mHealth substudy (Table 5.1, Figure 5.1). The substudy was slightly more white and more likely to be married than the full cohort. Overall, this speaks to the general transportability of the mHealth validation substudy to the broader cohort. Geographic transportability was observed, with the substudy providing participants from 42 out of 48 contiguous states (Figure 5.2)

Agreement between Residential-Based and Mobility-Based Walkability Exposure Measures in Substudy

In the substudy, residential-based walkability z-scores were lower than mobility-based measures (Figure 5.3). We measured a mean z-score of 0.0 (SD 2.6) for residential-based walkability and a mean z-score of 0.4(SD 2.4) for mobility-based walkability. The IQR for residential-based walkability was 1.8 and the IQR for mobility-based walkability was 1.5 (Table 5.2).

Tests for Bland Altman agreement indicated disagreement between residential-based and mobility-based measures of walkability. The Bland Altman agreement bias between the mHealth substudy participant's residential walkability and mobility averaged walkability was -0.2 (95% CI: -0.4, 0.1), meaning that the mean of participants' residential-based measures of walkability were lower by a z-score value of 0.2 compared to their mean mobility-based measure of walkability (Figure 5.4).

For analyses in the full NHS3 cohort (Figure 5.5), unadjusted models showed that each 1 SD increase (all SD scaled using residential walkability z-score SD 2.6) in uncorrected residential-based walkability z-score was associated with 19.4-minute difference (95% CI: 16.8, 22.1) in time spent preforming all physical activity (walking, running, and jogging). After accounting for individual participant measures of age, socioeconomic status, education level, and marital status, and residential measures of neighborhood socioeconomic status, greenness, mean yearly temperature, and total yearly precipitation, the association was modestly attenuated; each 1 SD difference in uncorrected residential walkability z-score was associated with 15.0-minute difference (95% CI; 11.7, 18.4) in time spent walking and running per week. In measurement error-corrected adjusted models each 1 SD difference in walkability z-score was associated with 23.6-minute difference (95% CI: 17.3, 29.8) in time spent walking and running per week.

Associations with other outcomes were similar, except for analyses of vigorous physical activity. In fully adjusted uncorrected estimates, each 1 SD difference in walkability z-score was associated with 1.9 fewer minutes in time spent conducting vigorous physical activity per week (-3., -0.7), while the corrected estimate was -3.0 (-4.9, -1.0) minutes. Across all outcomes, the error corrected estimates were corrected away from the null than uncorrected estimates (Figure 5.5, Supplemental Table 5.2).

Discussion

Our results suggest that traditional residential-based estimates of walkability exposure are slightly lower than mobility-based estimates of exposure, and that residential exposures may underestimate the magnitude of the true association between neighborhood walkability and physical activity. In addition, we observed that findings were consistent across all outcomes except vigorous physical activity only. We would not expect walkability to impact vigorous physical activity (Giles-Corti et al., 2016; National Academies of Sciences et al., 2019). This conclusion supports previous findings in the mHealth substudy highlighting that identified walkable areas are associated with movement, but no relationship was found between walkable environments and higher velocity physical activity. Additionally in a 2021 review on walkability and its relations with health, Boabeid et al. (Baobeid et al., 2021) highlight walking as the primary form of active transportation as opposed to more vigorous physical activity like running. These findings are some of the first to highlight exposure misclassification in epidemiological studies of the built environment with physical activity and how that misclassification impacts effect estimates. Previous work on measurement error in environmental epidemiology has focused heavily on air pollution (Feng et al., 2023; Hart et al., 2015b; Wei et al., 2022), and specifically fine particulate matter ($PM_{2.5}$). Similar to our findings, this work has suggested that effect estimates increase slightly after error correction.

Physical activity, while an individual behavior occurs in the context of the built and social environments that influence an individual's decisions about physical activity (National Academies of Sciences et al., 2019). In Karmeniemi et al.'s 2018 review features of the built environment were found to be associated with higher levels of physical activity due to amenity accessibility and transportation patterns (Kärmeniemi et al., 2018). Access to transportation provides a crucial opportunity for physical activity as walking often occurs at either end (Giles-Corti et al., 2016; M. Smith et al., 2017). Similarly the Surgeon General's call to action to increase physical activity reported walking at the most common form of physical activity in the US (Office of the Surgeon General (US), 2015). Our results support these findings, suggesting the association with walkability and physical activity is strongest for active transport physical activity patterns like walking. No studies to our knowledge utilize GPS measures of walkability to correct residential estimates of walkability exposure in the association with physical activity. GPS walkability exposures have been associated with self-reported and objective physical activity previously in the literature (James, Hart, et al., 2017; Marquet et al., 2020, 2022a; Orstad et al., 2018; Roscoe et al., 2022; Rundle et al., 2016). Similar to our study, studies have shown positive associations between residential-based walkability exposure and self-reported physical activity (Carr et al., 2010; Frank et al., 2008; Humpel et al., 2002; Saelens et al., 2003). Lastly, in a validation study in Seattle, Washington housing density and transportation density were found as key built environmental drivers of the association between walkability and objective physical activity (Mooney et al., 2020). As outlined above there is robust evidence on the association between walkability and physical activity examining the exposure at both residential and mobility. Our study builds on this previous literature to compare estimates using both residential and mobility walkability exposures to explore the potential for bias in residential estimates.

This study has some limitations. NHS3 is a cohort of predominantly upper-middle class, white, female nurses and as such these findings may have limited generalizability outside this population. Diverse cohorts should assess effect modification across race and SES to further confirm these findings. There is some concern about the transportability between NHS3 and the mHealth substudy, however all our internal checks confirmed transportability principals were generally met as the full NHS3 cohort and mHealth substudy had similar demographic distributions though the substudy was slightly more white and likely to be married. The physical activity outcome depends on self-report status for the entire year in question and as such there may be recall bias, though this is expected to be non-differential with respect to the exposure. Next, the census tract scale of walkability additionally opens concerns for non-differential exposure error, however we anticipate this error to be minor as Census tracts are relatively small and homogenous with respect to the built environment. Future studies could develop a sub-tract walkability index to omit this potential for error.

This study does, however, have some substantial strengths. First, we were able to use a walkability measure utilizing three robust datasets to quantify walkability across the contiguous United States. Due to the fine activity space data of the validation dataset, we obtained an estimation of yearly time-weighted mobility-based exposure to compare to residential estimates. Lastly validating walkability measures using the same internal population lends a key strength of transportability to the validation study.

Our findings highlight the importance of considering measurement error with respect to the built environment and physical activity. Residence-based exposure to walkability may underestimate the true association with physical activity. In multivariable models, even non-differential measurement error can be biased away from the null. Here we saw a difference in the effect estimate when correcting the exposure of interest across key outcome definitions, including all physical activity and walk only physical activity. We must validate our physical activity questionnaires compared to gold standard activity data to ensure reduction of measurement error in the outcome as well. Other exposures and outcomes (sleep, heart rate, temperature etc.) could be used within a similar framework. With the advances of mobile health technology, measurement error in environmental epidemiology studies can be corrected for-accounting for exposure ascertainment outside of the home environment.

Table 1. Characteristics of women in the Nurses' Health Study 3 (NHS3, N = 49,693) cohort and the Mobile Health (mHealth) Substudy population (N = 337)

Variable		**NHS3 Cohort (N=49,693)**		**mHealth Population (N=337)**	
	Categories	N	% or Mean (SD)	N	% or Mean (SD)
Age					
	Continuous, years	49,516	36.33 (7.29)	330	36.01 (7.3)
Race					
	White	43,026	88.2	317	94.1
	Black	1,797	3.7	8	1.8
	Asian	1,529	3.1	2	0.1
	Mixed Race	1,058	2.2	4	1.2
	Other	1,385	2.8	6	1.8
Ethnicity	Hispanic	2,538	5.2	14	4.2
Married					
	Yes	27,852	57.1	207	61.4
	No	20,943	42.9	130	38.6
Advanced Degree					
	Yes	41,027	84.1	249	73.9
	No	7,768	15.9	88	26.1
Employment					
	Yes	40,808	93.0	319	96.6
	No	3,084	7.0	11	3.4

Table 2. Uncorrected Generalized Linear Models Depicting Association Between Residential Walkability and Weekly Physical Activity Minutes and Regression Calibration GPS Mobility Corrected Generalized Linear Models Depicting Association Between Residential Walkability and Weekly Physical Activity Minutes. GPS locations were used to calculate mobility-based walkability estimates for corrected exposure. Associations expressed in difference in minutes of weekly physical activity by 1 SD change in walkability z-score (SD 2.64). Adjusted models controlled for individual participant measures of age (years, continuous), socioeconomic status defined as: education level (advanced degree, binary), and marital status (binary), and area-level measures of neighborhood socioeconomic status (z-score, quartiles), greenness (normalized difference vegetation index (NDVI), continuous), mean yearly temperature (Celsius, quartiles), and total yearly precipitation (dichotomized to <0.01mm, ≥0.01mm).

Exposure Measure		**Uncorrected**		**Mobility Corrected**	
	Model	Beta	95% CI	Beta	95% CI
All Physical Activity	Unadjusted	19.43	16.79, 22.07	30.54	25.11, 36.01
	Adjusted	15.02	11.67, 18.35	23.55	17.27, 29.81
Walk Only	Unadjusted	11.14	9.27, 13.04	17.53	13.94, 21.12
	Adjusted	10.67	8.26, 13.07	16.74	12.25, 21.23
Low Intensity	Unadjusted	5.76	4.88, 6.63	9.06	7.34, 10.77
	Adjusted	4.22	3.12, 5.33	6.63	4.65, 8.61
Vigorous	Unadjusted	-3.78	-4.73, -2.82	-5.94	-7.60, -4.30
	Adjusted	-1.90	-3.12, -0.69	-2.96	-4.94, -1.03
Moderate and Vigorous	Unadjusted	3.54	1.45, 5.62	5.57	2.22, 8.92
	Adjusted	1.58	-1.08, 4.25	2.48	-1.74, 6.68

Figure 5.1 Sutdy participant flow diagram for Nurses' Health Study 3 and the mHealth Substudy

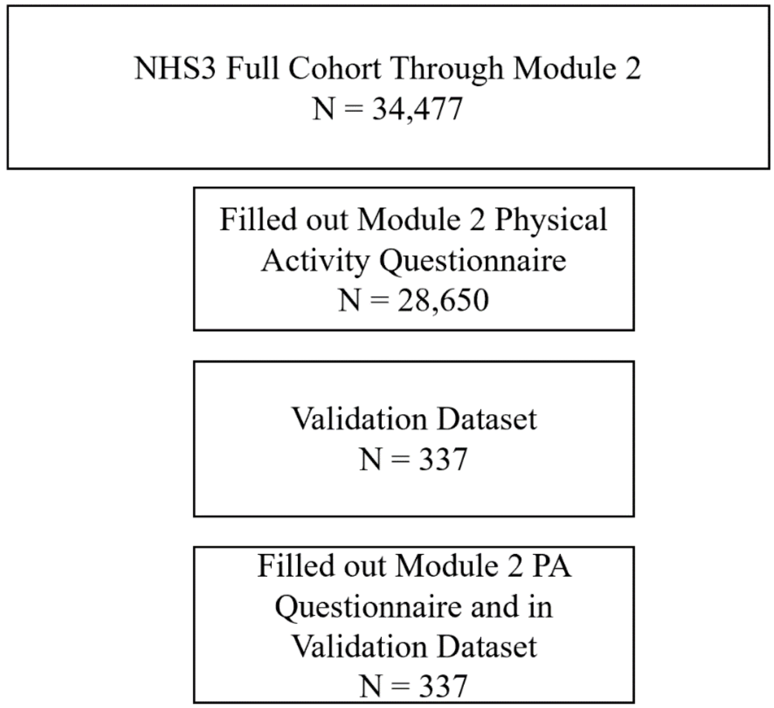


Figure 5.2. Nurses Health Study 3 (N=49,693) and mHealth Substudy (N=337) participants residential locations across the contiguous United States, 2010-present

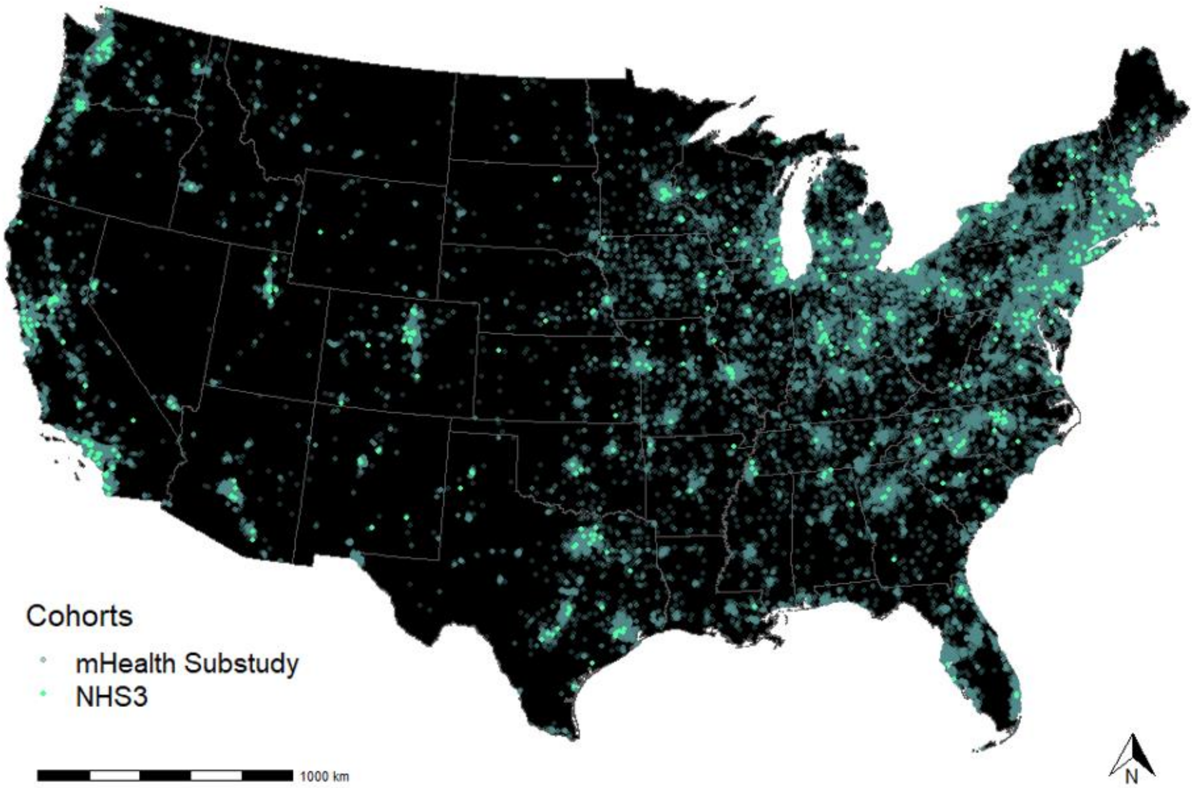


Figure 5.3. Distribution of walkability Z-Scores for two measures of Census tract walkability: residential and mobillity-based for the Nurses' Health Study 3 mHealth Substudy (N=337). The vertical black dotted line represents the mean of all walkability Z-scores.

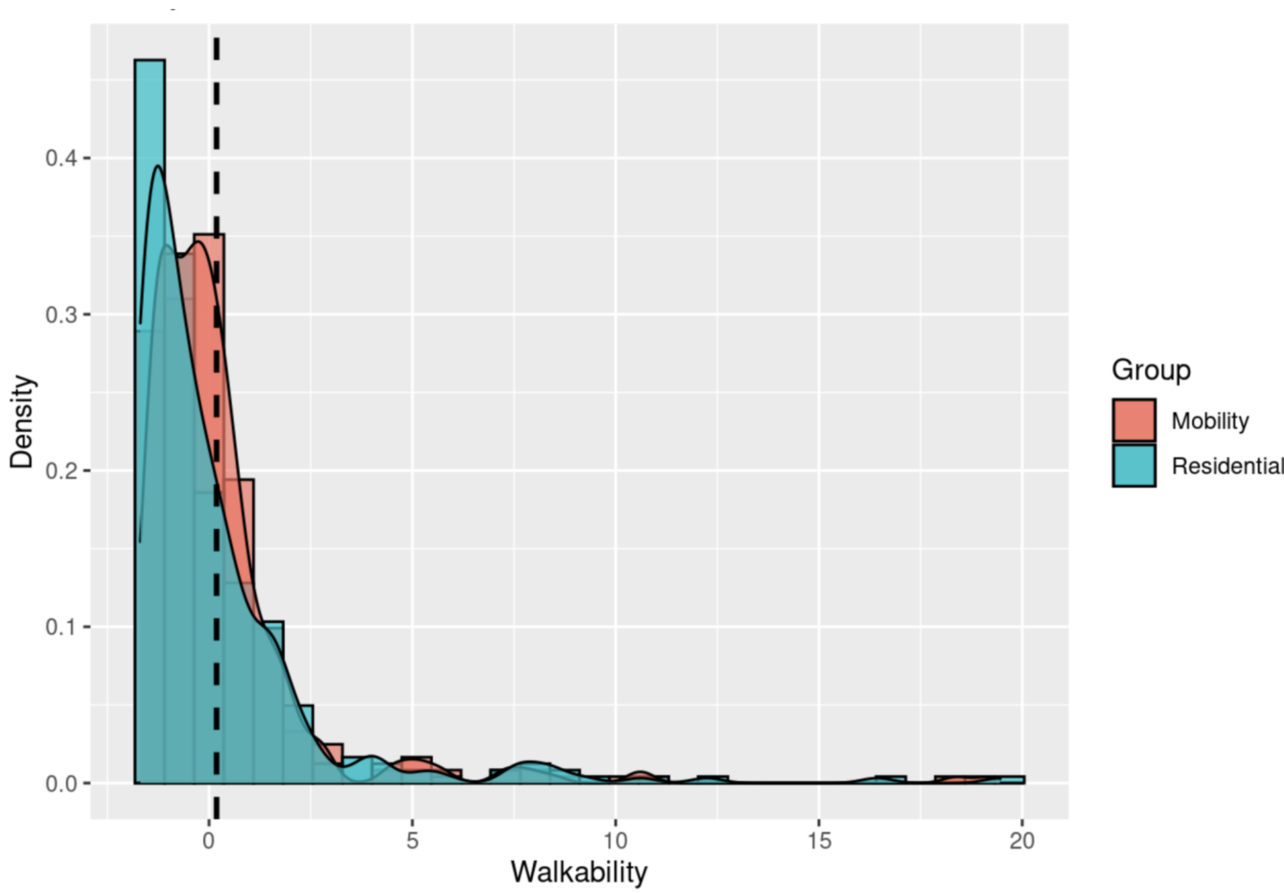


Figure 5.4. Bland Altman plots comparing mobility and residential walkability z-score measures. Purple band indicates 95% confidence level with red and green bands as lower and upper limit of agreement confidence bands for the Nurses' Health Study 3 mHealth Substudy (N=337).

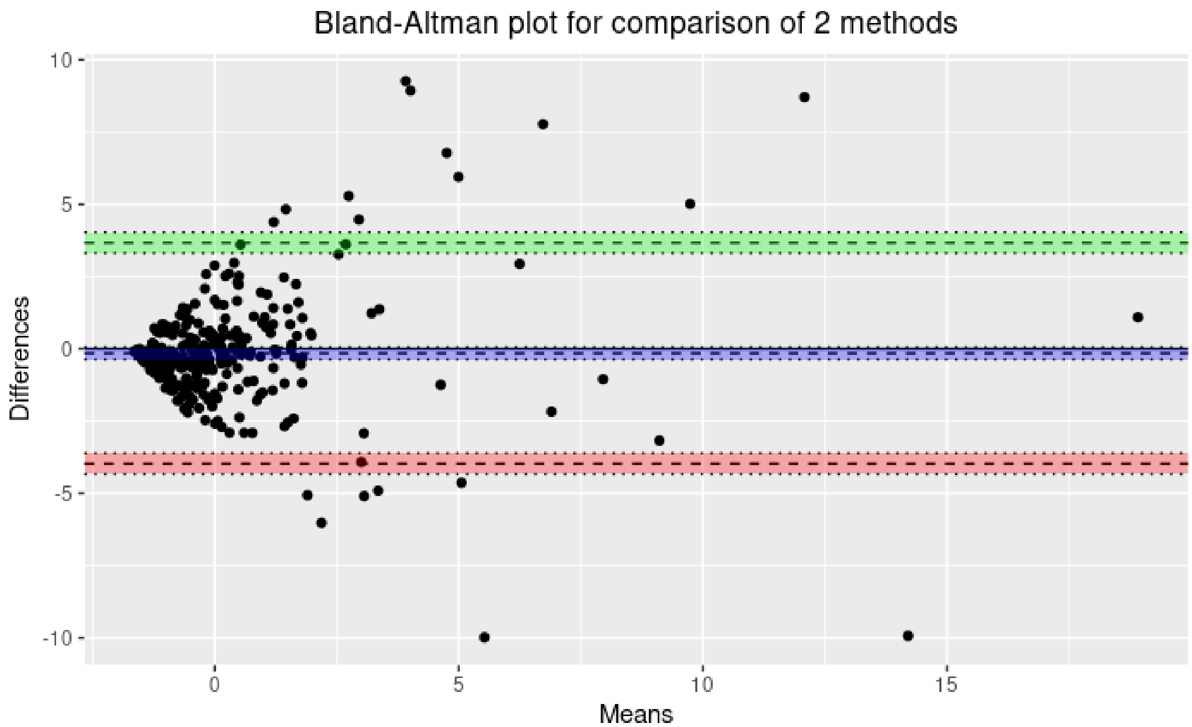
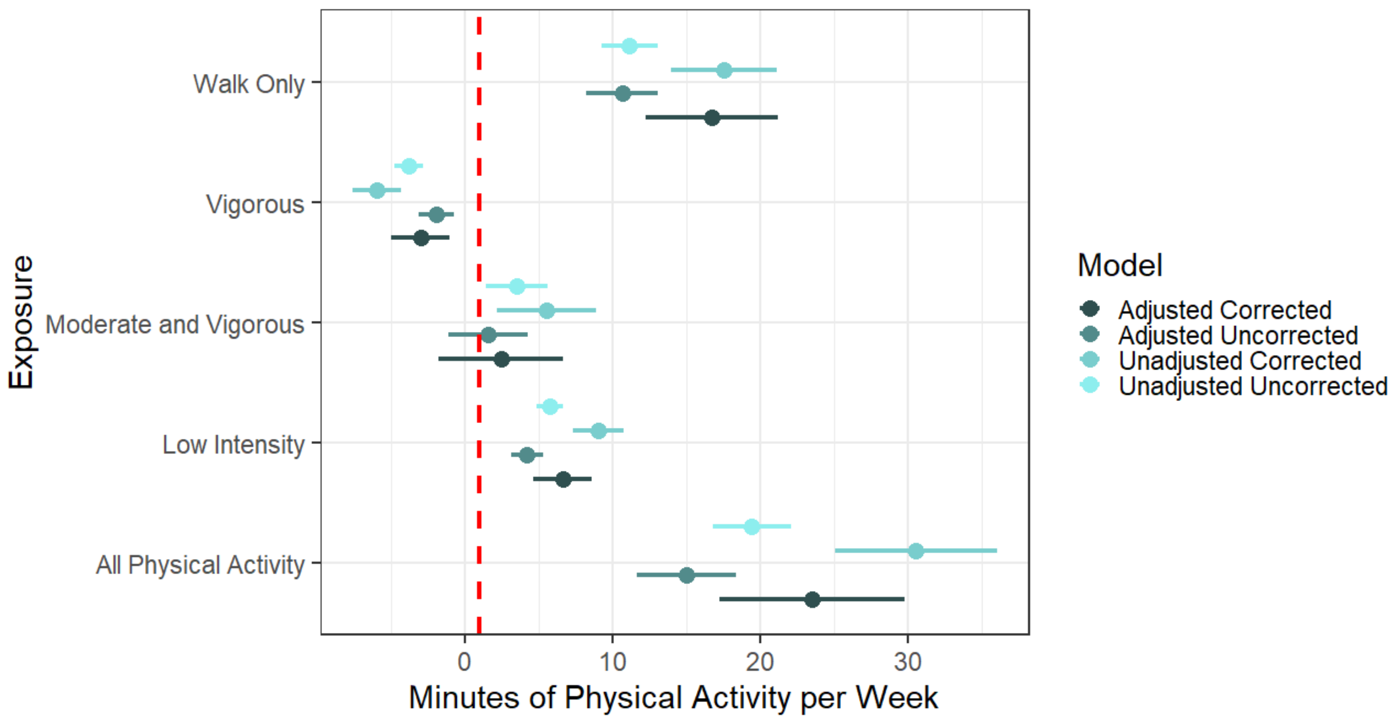


Figure 5.5 Uncorrected generalized linear models depicting association between residential walkability and weekly physical activity minutes and regression calibration GPS mobility corrected generalized linear models depicting association between residential walkability and weekly physical activity minutes for walk only, vigorous, moderate, and vigorous, low intensity, and all physical activity. GPS locations were used to calculate mobility-based walkability estimates for corrected exposure. Associations expressed in difference in minutes of weekly physical activity by 1 SD change in walkability z-score (SD 2.64). Adjusted models controlled for individual participant measures of age (years, continuous), socioeconomic status defined as: education level (advanced degree, binary), and marital status (binary), and area-level measures of neighborhood socioeconomic status (z-score, quartiles), greenness (normalized difference vegetation index (NDVI), continuous), mean yearly temperature (Celsius, quartiles), and total yearly precipitation (dichotomized to <0.01mm, ≥0.01mm).



Chapter 6.
Conclusion

Residential addresses as a proxy for exposure estimates and subjective and self-reported outcome data are ubiquitous in the literature surrounding greenness, walkability, and physical activity. To examine hypothesized mechanisms between greenness, walkability, and physical activity we require GPS mobility exposure and objective outcome data sources. This dissertation contributes to outstanding research gaps in the field, including enhanced understanding of: 1) the concordance between residential and mobility-based estimates of greenness; 2) the momentary associations of greenness and walkability with objective physical activity; and 3) the role of improved mobility estimates on measurement error correction in the association between walkability exposure and self-reported physical activity.

All studies (Chapter 2-5) bring the strength of highly resolved spatiotemporal data nested within a large nationwide prospective cohort (NHS3) (Fore et al., 2020). The first study (Chapter 2) utilized detailed residential and GPS information to compare greenness exposure measures based on residential address and GPS coordinates to quantify shortcomings of previous studies using solely residential address data to elucidate the exposure of interest. Given the availability of GPS data sampled over the course of a year we were able to compare regional, seasonal, and built environmental patterns between the two greenness measures. Differences were identified between mean annual residential-based NDVI and mean annual mobility-based NDVI values. Differences persisted when examining associations in datasets that 1) included GPS points only outside of the workplace location, 2) included GPS points with velocities that fell within human propelled movement (walk to run speed), and 3) included GPS points of individuals who provided at least 12 hours of daily data for 10 days in two sampling periods. Largest differences between mean annual residential-based NDVI and mean annual mobility-based NDVI values were observed in summer months and in low walkability census tracts.

Detecting differences in mobility-based and residential based NDVI confirms our hypotheses that one's home location exposure is not representative of one's throughout the day. Our findings suggest that regression calibration approaches similar to our fourth study (Chapter 5) should be used wherever possible to correct exposure outcome associations.

In the second study (Chapter 3), the dissertation examined associations between mobility-based NDVI and mean steps measured approximately every ten minutes. We identified a non-linear inverse association between GPS-based NDVI every 10 minutes and mean step count in the following 10 minutes after adjusting for a priori individual and area-level confounders. This association held consistent through sensitivity analyses with restricted datasets on 1) wear time, 2) selective daily mobility bias, and 3) workplace location. A sensitivity analysis on active transportation data showed a null result. We observed effect modification by region and season with non-linear positive associations found in the south and the strongest effect estimates of an inverse association found in spring and summer months.

The third study of the dissertation (Chapter 4) examined associations between mobility-based walkability and mean objectively measured steps measured approximately every ten minutes. We identified a non-linear positive association between GPS-based walkability every 10 minutes and mean step count in the subsequent 10 minutes after adjusting for a priori individual and area-level confounders. This association held consistent through sensitivity analyses with restricted datasets on 1) wear time, 2) selective daily mobility bias, and 3) workplace location. A sensitivity analysis on active transportation data showed a null result. We observed effect modification by neighborhood SES, region and season with strongest non-linear positive associations found in lowest quartiles of neighborhood SES, spring months and the Midwest.

In the second and third studies we focus mainly on assessing instantaneous pathways of association. We attempted to address if being located in a greener or more walkable environment was followed by a period of higher physical activity. Our studies take preliminary steps to address exposure outcome pathways between contextual built and natural exposures and physical activity, a frequent mediator of downstream health outcomes. Future studies can employ more rigorous designs and analysis to test causal hypotheses. These studies should investigate the mechanism at alternate temporal scales (e.g. daily exposure and daily outcome, weekly exposure and weekly outcome) and lags (previous day exposure, present day outcome) to understand the full implications of the complex spatial and temporal patterns of these exposure outcome associations.

Lastly, the fourth and final study of the dissertation (Chapter 5), applied measurement error correction to assess the association between residence-based walkability and physical activity corrected for mobility-based walkability. We identified a positive association between residential walkability exposure and self-reported weekly minutes of physical activity controlling for individual and area-level a priori confounders. After applying measurement correction measures on the exposure utilizing mobility-based measures of walkability we observed a stronger positive association between the exposure and outcome than in non-corrected models. A positive association was observed for all physical activity outcome measures except vigorous physical activity, where a null association was identified. These approaches should be replicated for additional exposures (temperature, precipitation, nSES) and outcomes including measurement correction using objectively measured physical activity.

In closing, it is crucial that environmental epidemiologists carefully consider the spatiotemporal scale of contextual exposures and align these estimates with outcomes of interest. Our first study highlighted that the reliance on residential exposure estimates may overestimate an individual's true greenness exposure. The second and third studies illustrated potential momentary contextual environmental drivers of physical activity. In the second study, we failed to reject our null hypothesis and found that higher greenness is not associated with higher steps per minute. In the third study we rejected our null hypothesis and found a positive association between higher walkability and higher steps per minute. This suggested a complex and nuanced relationship between built and natural environments and physical activity. Walkable environments were associated with higher physical activity, whereas green environments were associated with lower physical activity and it did not appear that there were interactions between the two exposures. Findings from the second and third studies highlighted the potential improved health effects of walkable areas. Although we did not find evidence for interaction between greenness and walkability on the outcome of physical activity, the evidence for benefits of green spaces on social cohesion and mental restoration suggest the potential implications of greening already walkable environments. While several cumulative exposures studies exist (James, Hart, et al., 2017; Marquet et al., 2022a; Roscoe et al., 2022), intervention studies are needed to continue to explore these hypotheses and examine the potential joint effects on multiple outcomes as well. Lastly, our fourth study shows the implications of utilizing GPS mobility-based exposure measures to correct estimates derived from large prospective cohorts. Integrating mobility-based estimates into regression calibration is underutilized for measurement correction regarding contextual exposures. Our findings highlight the importance of investing in digital health technology to address bias in environmental epidemiology. In the fifth chapter wee outline approaches future researchers can implement utilizing GPS data and regression calibration to address measurement error for environmental exposures.

This dissertation grapples with the emergence of mHealth data for environmental epidemiology and the historical use of prospective cohorts to identify interactions of place, the environment, and human health. The work illustrated here is only a fraction of what is possible with integrated GPS and wearable device data for health outcomes research. The future is already here. Integrated ecological momentary assessment (EMA) surveys in large prospective cohorts allow for researchers to study perceptions, attitudes, and behaviors associated with momentary environmental exposures. The Nurses' Health Studies are currently utilizing the Beiwe platform to roll out EMAs across the cohorts. The plethora of GPS and object data require undertaking validation studies to understand wear time and study duration requirements.

The creation of improved contextual environmental exposures such as those using Google Street View are necessary. These metrics will provide specific information on the ground elements of natural and built environments (e.g. trees or sidewalks). While NDVI can quantify one's surrounding greenness, it cannot provide information on how this green space is perceived by those who interact with it. Efforts to develop a measure of perceived green space at a national scale are underway. This would allow researchers to estimate how green space is actually viewed by individuals. This is an important distinction because while green spaces may be present within a community, issues of public vs. private lands, access, safety, traffic patterns and climate may drastically alter how green spaces are perceived. There is evidence to suggest that understanding differences in green space perception can lead to divergent health outcomes. Similar arguments can be made for the exposure of walkability. Sidewalk availability and quality, public transportation, lighting, and safety are all factors of the built environment that must be integrated into future iterations of walkability measures. In this dissertation we aimed to provide a comparison of the ubiquitous contextual exposures of greenness and walkability at the momentary and residential spatial scales. Recognizing the crucial differences we observed, the next iteration must include elevating the exposure measures themselves.

While this dissertation paves the way for future research, many conclusions can be drawn from the findings themselves. Evidence from this dissertation and the work of numerous scholars from the fields of geography, public health, urban planning, and design highlight tangible conclusions that must be drawn from the literature. Greenness and walkability exposures are inversely related in space and promote unique health behaviors. The identification of an inverse association between greenness and physical activity in chapter 3 does not seem to suggest that individuals choose to actively avoid physical activity in higher greenness. Roscoe et al. suggest this inverse relationship may be due to what drives the positive association in the walkability and physicality activity relationship: population, building, and intersection density (Roscoe et al., 2022). Thus, the critical inverse relationship between greenness and walkability suggests where interventions and policies will be most beneficial. Policies and ordinances that tackle promotion of housing density, mixed use development, and deemphasize motor-vehicle based travel may create the conditions for higher physical activity. Greening these spaces may not necessarily lead to increased physical activity but may lead to notable other health benefits for communities in terms of mental health or other outcomes. In 2020 Minneapolis, MN became the first major U.S. city to implement a ban on single-family zoning across the city. Prior to this, single family zoning accounted for ¾ of urban land in Minneapolis (*Influential Minneapolis Policy Shift Links Affordability*, *Equity*, 2020.). Implications of this policy change are far ranging from racial and financial equity creating more opportunities for first time home buyers to improved health outcomes as population density is a key component of walkability. After the 2021 Boston mayoral elections, the chief of streets outlined a vision for the city stretching beyond car-centric transportation. These policies include free fare transportation, creation of “low-stress” bicycle routes, and promoting and expanding green corridors like the Emerald Necklace and southwest corridor bicycle and walking path designed by Fredrick Law Olmsted that connects the city (*Boston’s New Chief of Streets to Steer City beyond Car-Centric Transportation*, 2021.). These policies have the opportunity to incentivize and drive behavior change related to physical activity and numerous other health behaviors including social interaction and cohesion.

LIST OF SUPPLEMENTAL INFORMATION

LIST OF SUPPLEMENTAL TABLES

LIST OF SUPPLEMENTAL FIGURES

SUPPLEMENTAL INFORMATION

Supplemental Information 2.1. NDVI derivation from Google Earth Engine.

NDVI raster imagery for the residential and mobility greenness exposure metrics were available through the Google Earth Engine (GEE) platform (Gorelick et al., 2017) and processed using Earth Engine Landsat-specific processing methods for Landsat Tier 1 Raw Scene collection for Landsat 8. We used the LandsatSimpleComposite algorithm to compute a Landsat Top of Atmosphere (TOA) composite from 30m raw scenes, assigned a cloud score to each pixel using the SimpleLandsatCloudScore algorithm, and selected only the least-cloudy scenes in regions where more than one input scene was available on a seasonal scale. By selecting the best imagery for each season, we minimize missing data due to cloud cover or other environmental interference. We then applied the "NormalizedDifference" algorithm to compute the normalized difference between the near-infrared and red bands to calculate the Normalized Difference Vegetative Index (NDVI=(NIR-RED/(NIR+RED)) at a 30m spatial resolution. To estimate 270m and 1230m buffers around residential addresses, we applied the Earth Engine "reduceNeighborhood" algorithm to create two textures at varying spatial resolutions, with each 30x30m pixel representing the mean of 270m or 1230m neighborhood radius around that pixel in the input image from the 30m NDVI composite.

Supplemental Information 2.2. Stratified Variables.

Age was obtained from the full NHS3 cohort study dataset from participants initial questionnaire return (Module 1), which predated enrollment in the Substudy.

We measured neighborhood Socioeconomic Status (nSES) using a composite score of 7 Census tract level variables representing domains that have been previously associated with health outcomes, including education, employment, housing, wealth, racial composition, and population density (*DeVille et al. 2022, in review*). Variables were carried forward from the prior U.S. Census (2010) and each variable was z-standardized. We summed the z-scores for each component variable to create a nSES score. Higher scores indicated higher nSES (i.e. less socioeconomic deprivation). We joined quartiles of nSES score using the location of each 10-minute GPS point to create a mobility-based nSES score at the same temporal resolution as the mobility-based greenness exposure assessment.

We defined neighborhood walkability as a composite 3-item score of intersection density, calculated from 2019 Tiger/Line shapefiles of all roads with interstates removed (Bureau, 2020), population density, from 2019 ACS population data (*Explore Census Data*, 2020), and business density, from 2018 Infogroup US Historical Business Data (Infogroup, 2020). Variables were z-standardized for each Census tract. We summed the z-scores for each component variable (3-items) to create a neighborhood walkability index. Higher scores indicated more walkable areas. We joined quartiles of walkability score using the location of each 10-minute GPS point to create a mobility-based walkability score at the same temporal resolution as the mobility-based greenness exposure assessment

SUPPLEMENTAL TABLES

Table S2.1. Bland Altman Agreement Bias and 95% Confidence Intervals for Residential and Mobility Based Greenness Exposure Per 0.1 NDVI in both the primary analytic dataset (N=337) and secondary analytic dataset (N=208)

	Primary Analytic Dataset (N = 337)					**Secondary Analytic Dataset (N = 208)**	
Exposure Measure	270m Residential Buffer	1230m Residential Buffer	30m Mobility Average	30m Non-Work Mobility Average	30m Walk Only Mobility Average	270m Residential Buffer	1230m Residential Buffer
270m Residential Buffer	---------------	0.005 (-0.001, 0.010)	0.078 (0.067, 0.089)	0.070 (0.055, 0.084)	0.115 (0.105, 0.125)	---------------	0.005 (-0.001, 0.010)
1230m Residential Buffer	0.005 (-0.001, 0.010)	---------------	0.073 (0.061, 0.084)	0.066 (0.051, 0.081)	0.110 (0.099, 0.120)	0.005 (-0.001, 0.010)	---------------
30m Mobility Average	0.078 (0.067, 0.089)	0.073 (0.061, 0.084)	--------------------	0.020 (-0.015, -0.024)	0.037 (0.031, 0.043)	0.074 (0.061, 0.086)	0.070 (0.057, 0.083)
30m non-work Mobility Average	0.070 (0.055, 0.084)	0.066 (0.051, 0.081)	0.020 (-0.015, -0.024)	---------------	0.047 (0.038, 0.056)	0.063 (0.048, 0.078)	0.062 (0.051, 0.079)
30m Walk Only Mobility Average	0.115 (0.105, 0.125)	0.110 (0.099, 0.120)	0.037 (0.031, 0.043)	0.047 (0.038, 0.056)	---------------	0.100 (0.082, 0.122)	0.107 (0.095, 0.119)

Table S3.1. Participant Greenness and Physical Activity Distributions Across Seasons for Secondary Analytic Cohort (N=208)

	N	Steps/Min Mean (SD)	Steps/Min Min, Max	Greenness Mean (SD)	Greenness Min Max
Total Participants	208	7.04 (14.98)	0.00, 263.78	0.31 (0.21)	0.00, 0.84
Fall	192	6.79 (14.67)	0.00, 181.00	0.27 (0.20)	0.00, 0.82
Winter	175	6.60 (14.26)	0.00, 183.13	0.21 (0.15)	0.00, 0.73
Spring	144	7.42 (15.25)	0.00, 219.00	0.37 (0.20)	0.00, 0.84
Summer	196	7.26 (15.25)	0.00, 263.78	0.38 (0.21)	0.00, 0.84

Table S4.1. Participant Walkability and Physical Activity Distributions Across Seasons for Secondary Analytic Cohort (n=208)

	N	Steps/Min Mean (SD)	Steps/Min Min, Max	Walkability Mean (SD)	Walkability Min Max
Total Participants	208	7.04 (14.98)	0.00, 263.78	0.18 (2.89)	-1.66, 52.84
Fall	192	6.79 (14.67)	0.00, 181.00	0.15 (2.85)	-1.66, 52.84
Winter	175	6.60 (14.26)	0.00, 183.13	0.24 (3.32)	-1.66, 52.84
Spring	144	7.42 (15.25)	0.00, 219.00	0.21 (2.65)	-1.66, 33.27
Summer	196	7.26 (15.25)	0.00, 263.78	0.15 (2.76)	-1.66, 44.60

SUPPLEMENTAL FIGURES

Figure S2.1 Yearly Normalized Difference Vegetation Index (NDVI) histogram distributions for 270m and 1230m residential measures and averaged 30m GPS mobility exposure measures in primary restriction criteria

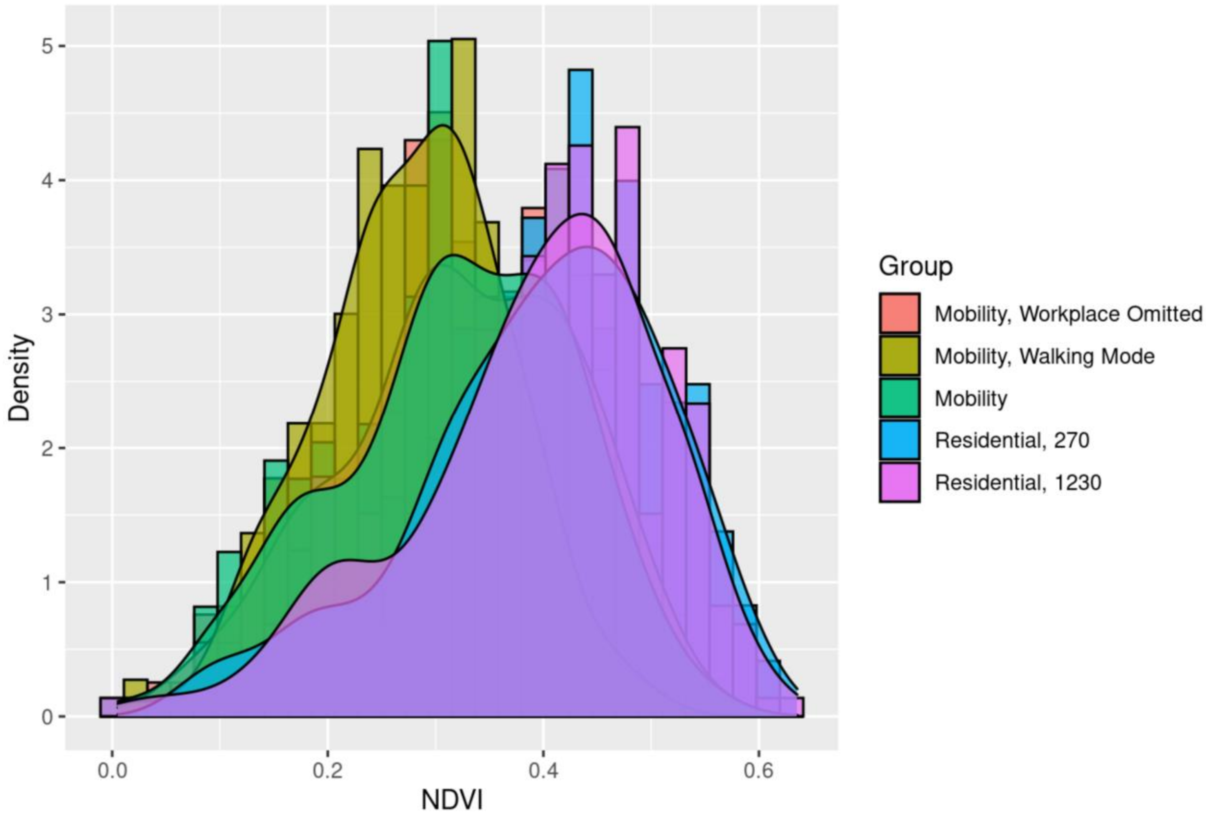


Figure S2.2 Yearly Normalized Difference Vegetation Index histogram distributions for 270m and 1230m residential measures and averaged 30m mobility exposure measures in secondary restriction criteria

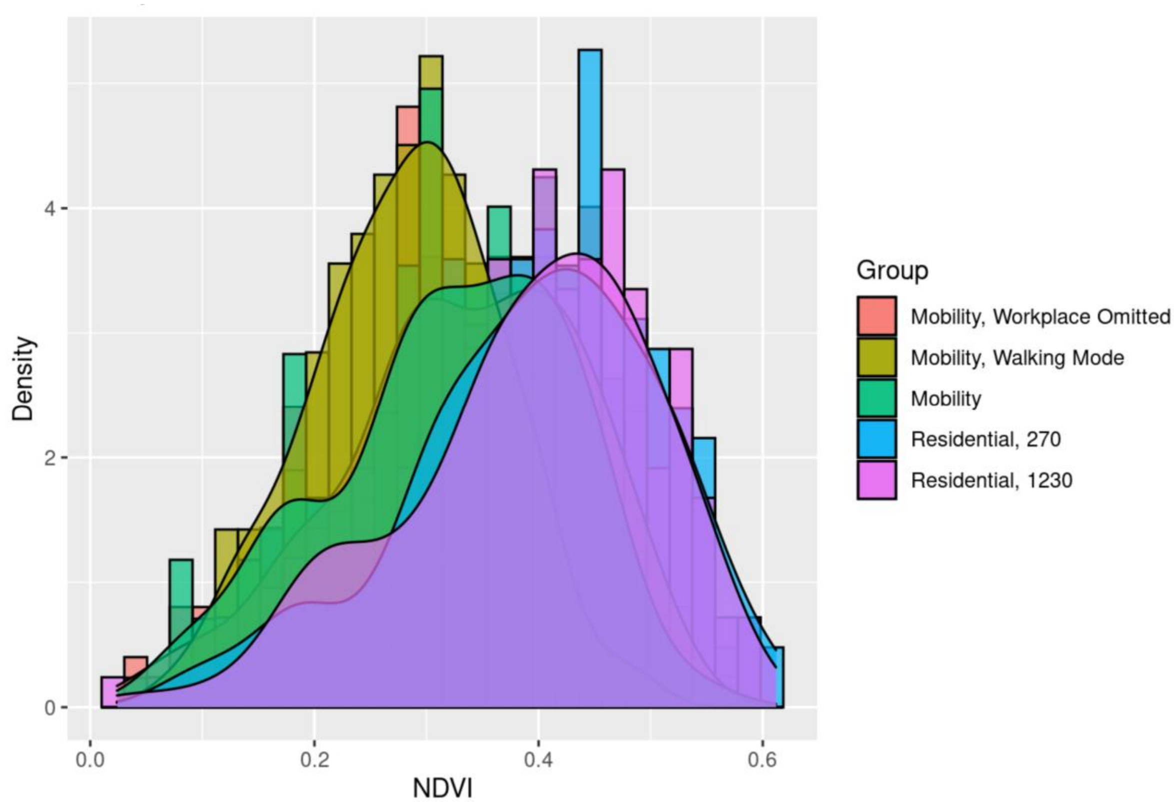


Figure S2.3 Bland Altman plots comparing mobility and residential Normalized Difference Vegetation Index (NDVI) measures. Purple band indicates 95% confidence level with red and green bands as lower and upper limit of agreement confidence bands in the secondary analytics dataset (N=208)

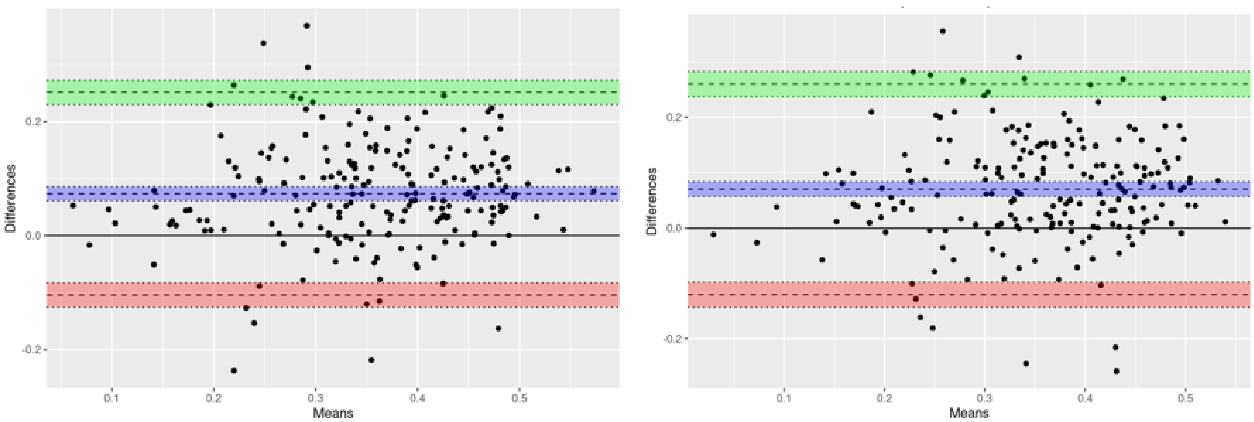


Figure S2.4 Generalized linear models depicting association between residential (270 and 1230 m buffers) and GPS mobility-based greenness exposures using NDVI in the secondary analytic dataset

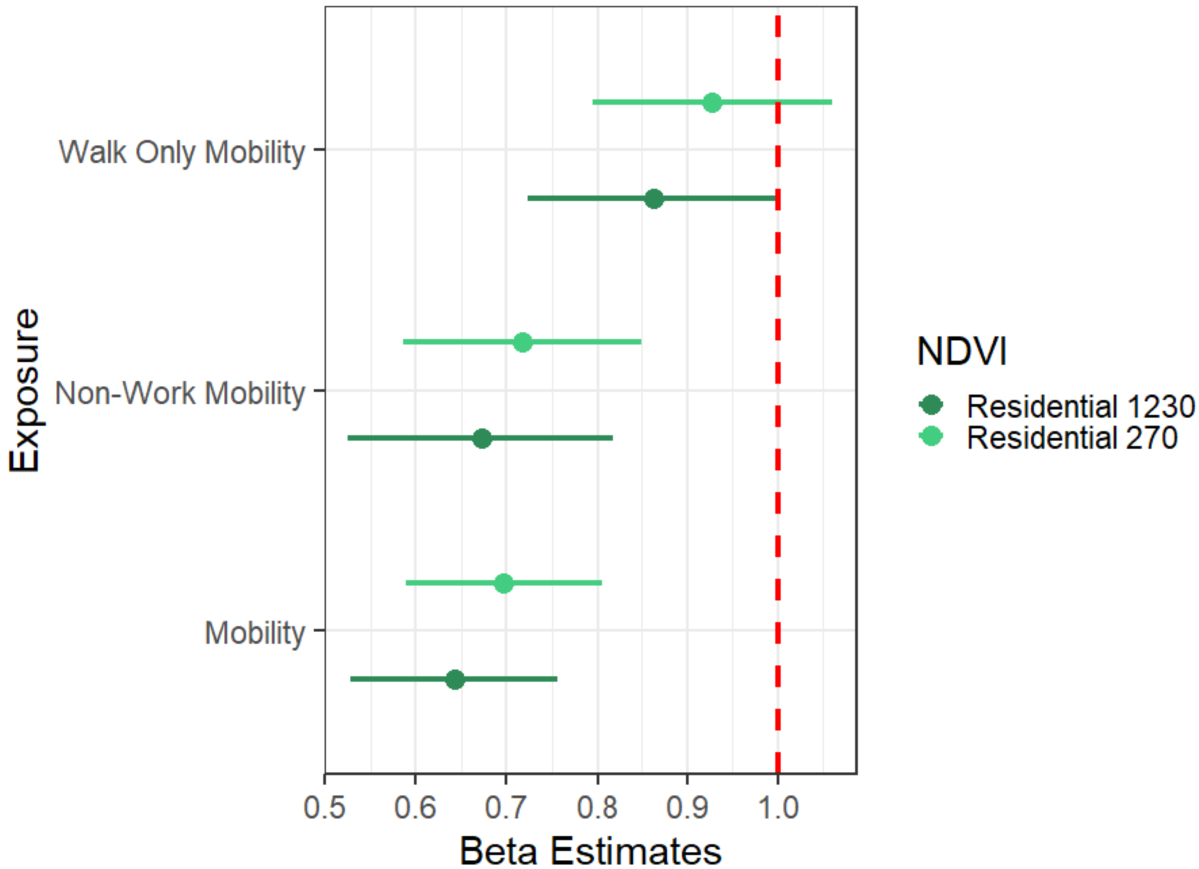


Figure S3.1. Nonlinear associations between Normalized Difference Vegitation Index (NDVI) and average steps per minute across a 10-minute period, restricting to selective daily mobility bias geographies, controlling for age (years; continuous), socioeconomic status defined as: education level (masters in nursing or higher; binary), and marital status (never [never married]]/ever [married, widowed, divorced]; binary), walkability (z-scores; quartiles), mean daily temperature (Celsius; quartiles), daily precipitation (millimeters; binary), NSES (z-scores, quartiles) season and Census region.

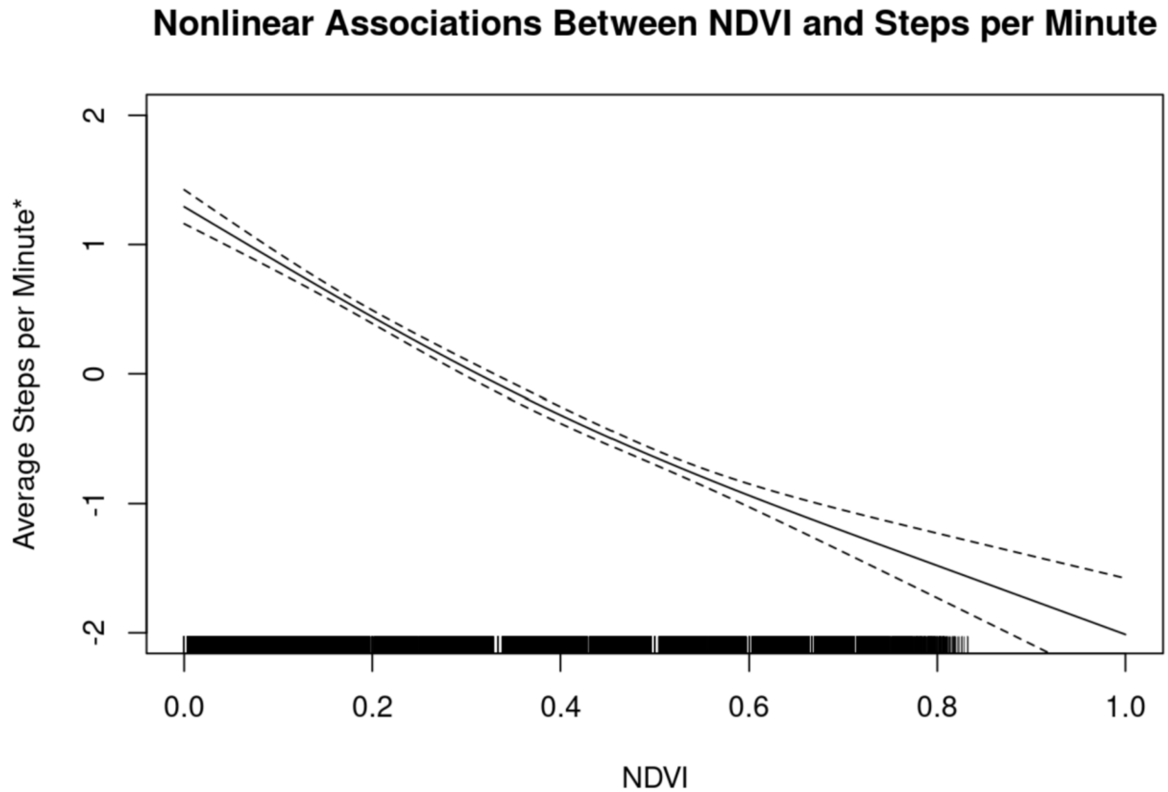


Figure S3.2. Nonlinear associations between Normalized Difference Vegitation Index (NDVI) and average steps per minute across a 10-minute period, restricting to workplace omitted geographies, controlling for age (years; continuous), socioeconomic status defined as: education level (masters in nursing or higher; binary), and marital status (never [never married]]/ever [married, widowed, divorced]; binary), walkability (z-scores; quartiles), mean daily temperature (Celsius; quartiles), daily precipitation (millimeters; binary), NSES (z-scores, quartiles) season and Census region.

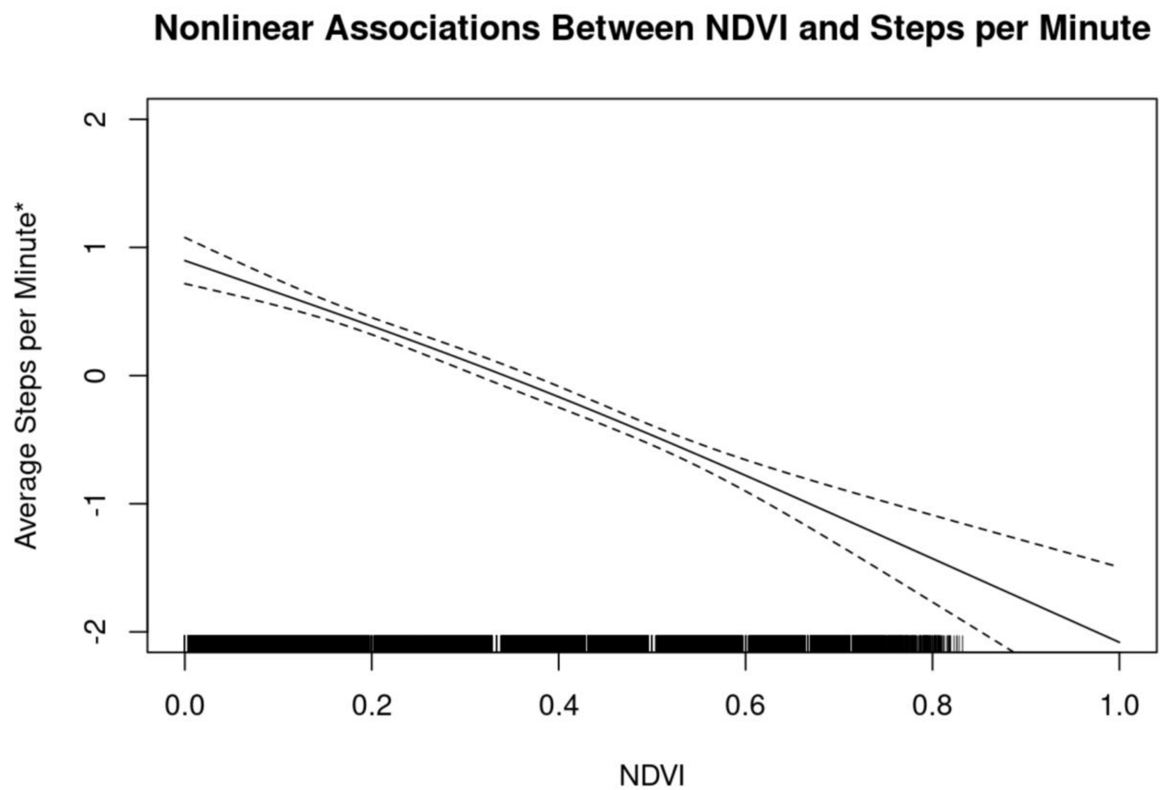


Figure S3.3. Nonlinear Nonlinear associations between Normalized Difference Vegitation Index (NDVI) and average steps per minute across a 10-minute period, restricting on the secondary analyic dataset, controlling for age (years; continuous), socioeconomic status defined as: education level (masters in nursing or higher; binary), and marital status (never [never married]]/ever [married, widowed, divorced]; binary), walkability (z-scores; quartiles), mean daily temperature (Celsius; quartiles), daily precipitation (millimeters; binary), NSES (z-scores, quartiles) season and Census region.

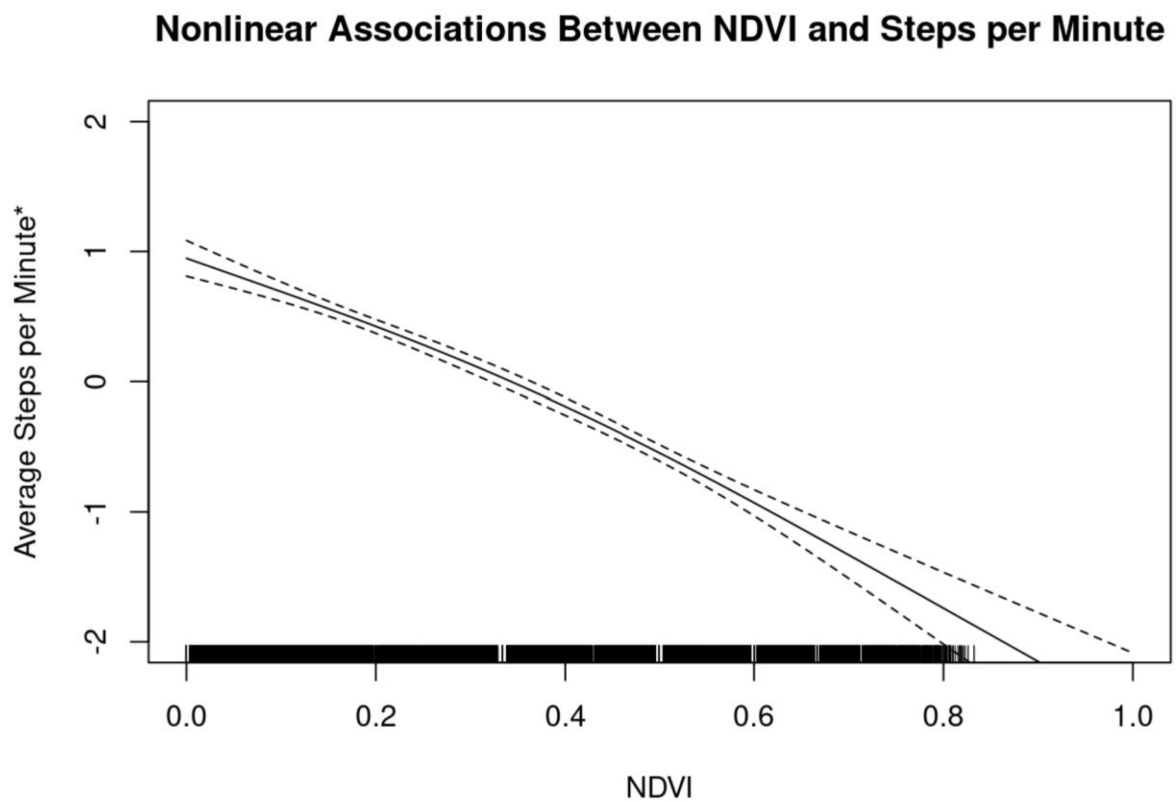


Figure S4.1. Nonlinear Associations Between Walkability Z-Score and Average Steps per Minute Across a 10-minute Period Controlling for age (years; continuous), socioeconomic status defined as: education level (masters in nursing or higher; binary), and marital status (never [never married]]/ever [married, widowed, divorced]; binary), and greenness (NDVI; quartiles), mean daily temperature (Celsius; quartiles), daily precipitation (millimeters; binary), season and Census region. Stratified on NSES Quartiles

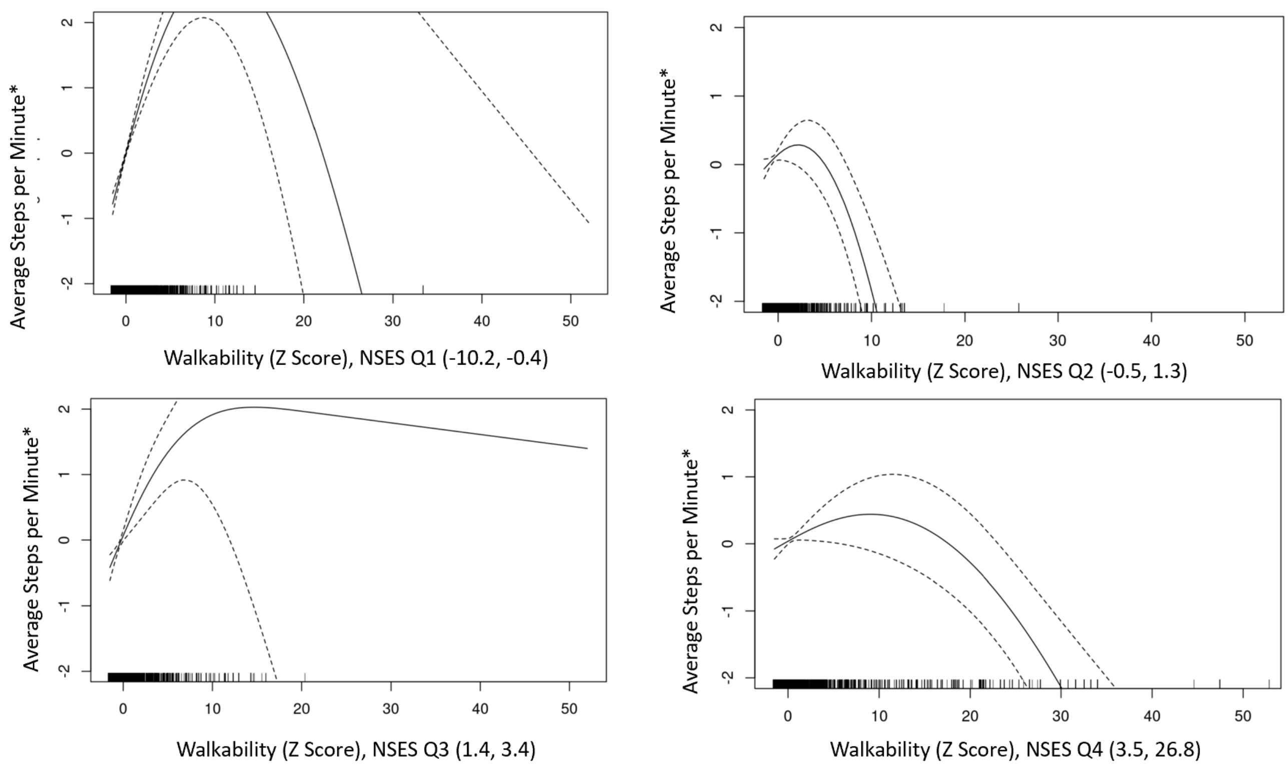


Figure S4.2. Nonlinear Associations Between Walkability Z-Score and Average Steps per Minute Across a 10-minute Period Controlling for age (years; continuous), socioeconomic status defined as: education level (masters in nursing or higher; binary), and marital status (never [never married]]/ever [married, widowed, divorced]; binary), and area-level measures of neighborhood socioeconomic status (z-score; quartiles), greenness (NDVI; quartiles), mean daily temperature (Celsius; quartiles), daily precipitation (millimeters; binary), and season. Stratified on census region.

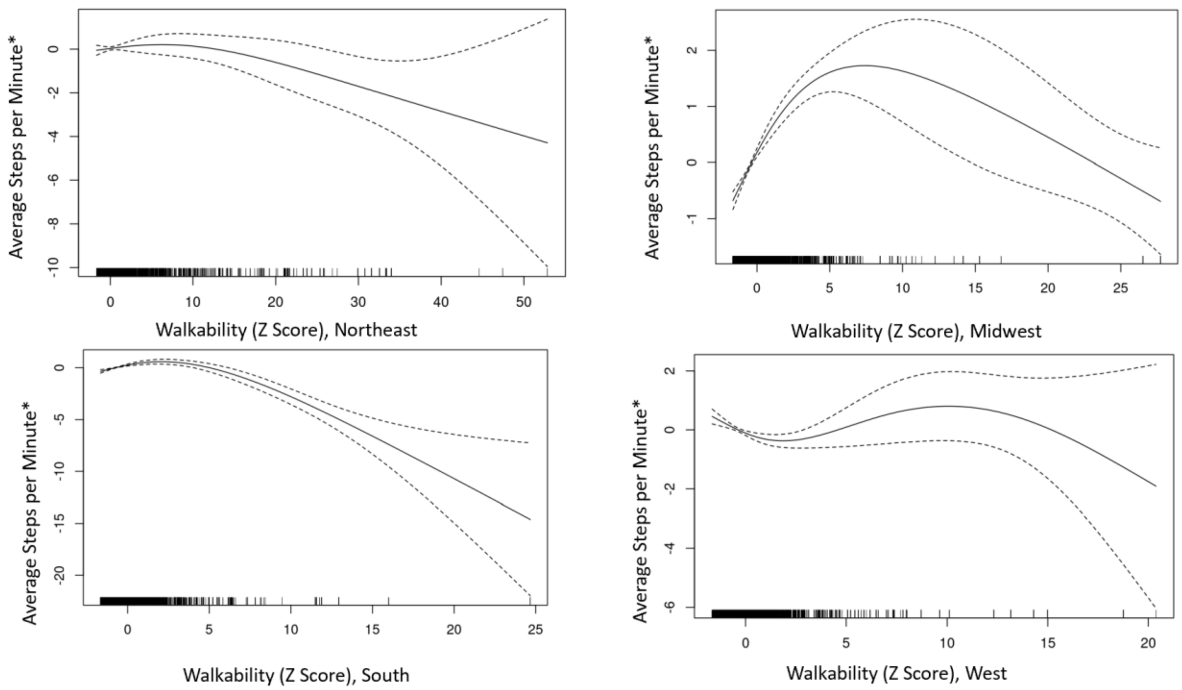


Figure S4.3. Nonlinear Associations Between Walkability Z-Score and Average Steps per Minute Across a 10-minute Period Controlling for age (years; continuous), socioeconomic status defined as: education level (masters in nursing or higher; binary), and marital status (never [never married]]/ever [married, widowed, divorced]; binary), and area-level measures of neighborhood socioeconomic status (z-score; quartiles), greenness (NDVI; quartiles), mean daily temperature (Celsius; quartiles), daily precipitation (millimeters; binary), and Census region. Minute Stratified on Season

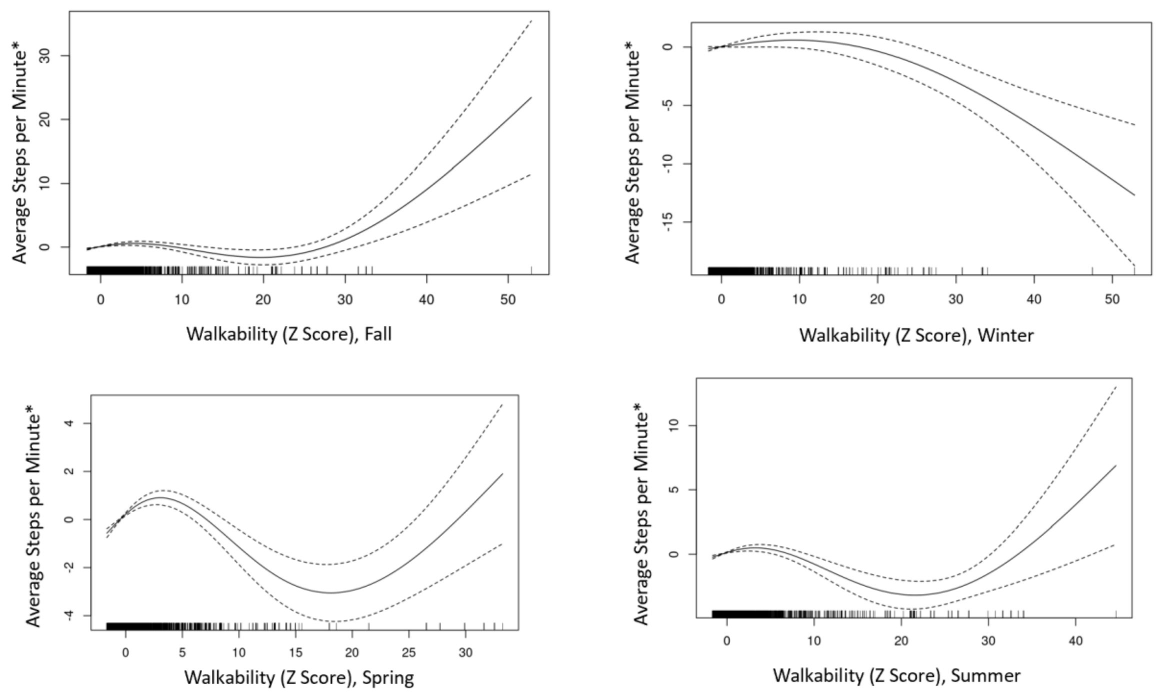


Figure S4.4. Nonlinear Associations Between Walkability and Steps per Minute Restricted to Selective Daily Mobility Bias Geographies, controlling for age (years; continuous), socioeconomic status defined as: education level (masters in nursing or higher; binary), and marital status (never [never married]]/ever [married, widowed, divorced]; binary), and area-level measures of neighborhood socioeconomic status (z-score; quartiles), greenness (NDVI; quartiles), mean daily temperature (Celsius; quartiles), daily precipitation (millimeters; binary), season and Census region. *Average steps per minute across each ten-minute collection period

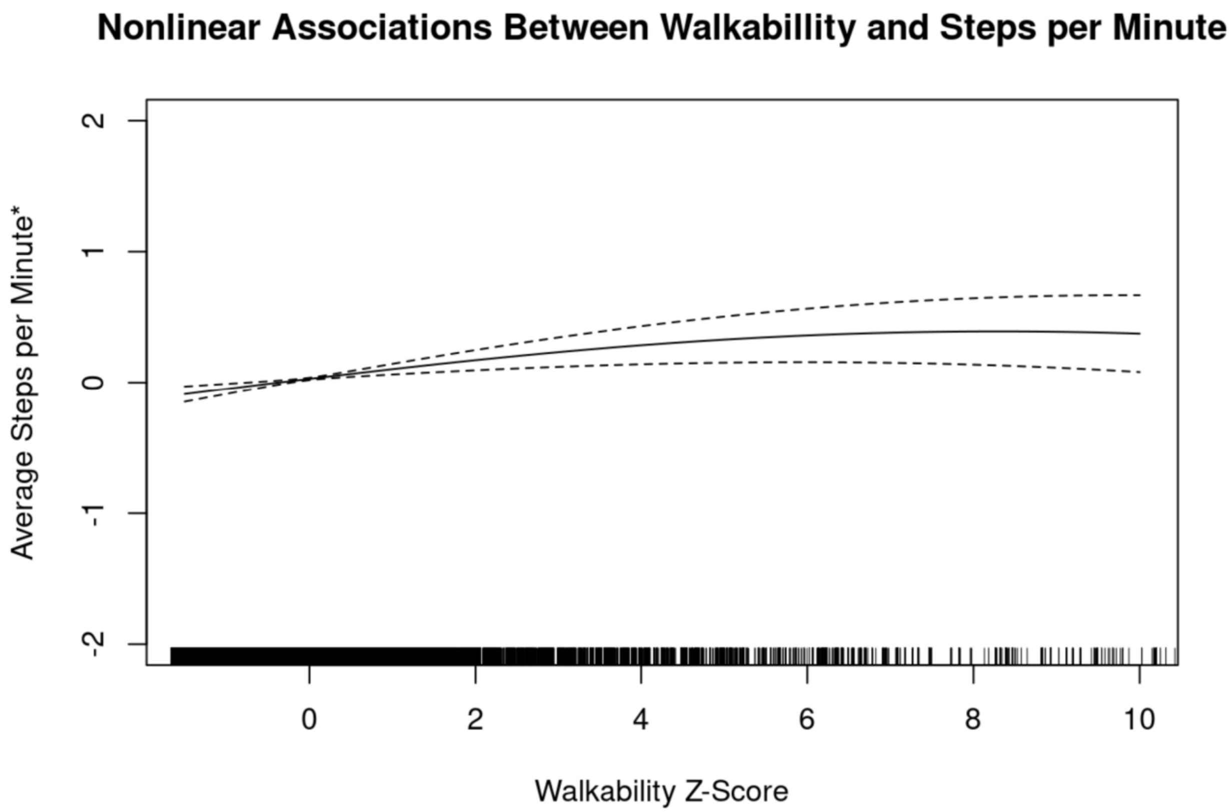


Figure S4.5. Nonlinear Associations Between Walkability and Steps per Minute Restricted to Workplace Omited Geographies, controlling for age (years; continuous), socioeconomic status defined as: education level (masters in nursing or higher; binary), and marital status (never [never married]]/ever [married, widowed, divorced]; binary), and area-level measures of neighborhood socioeconomic status (z-score; quartiles), greenness (NDVI; quartiles), mean daily temperature (Celsius; quartiles), daily precipitation (millimeters; binary), season and Census region. *Average steps per minute across each ten-minute collection period

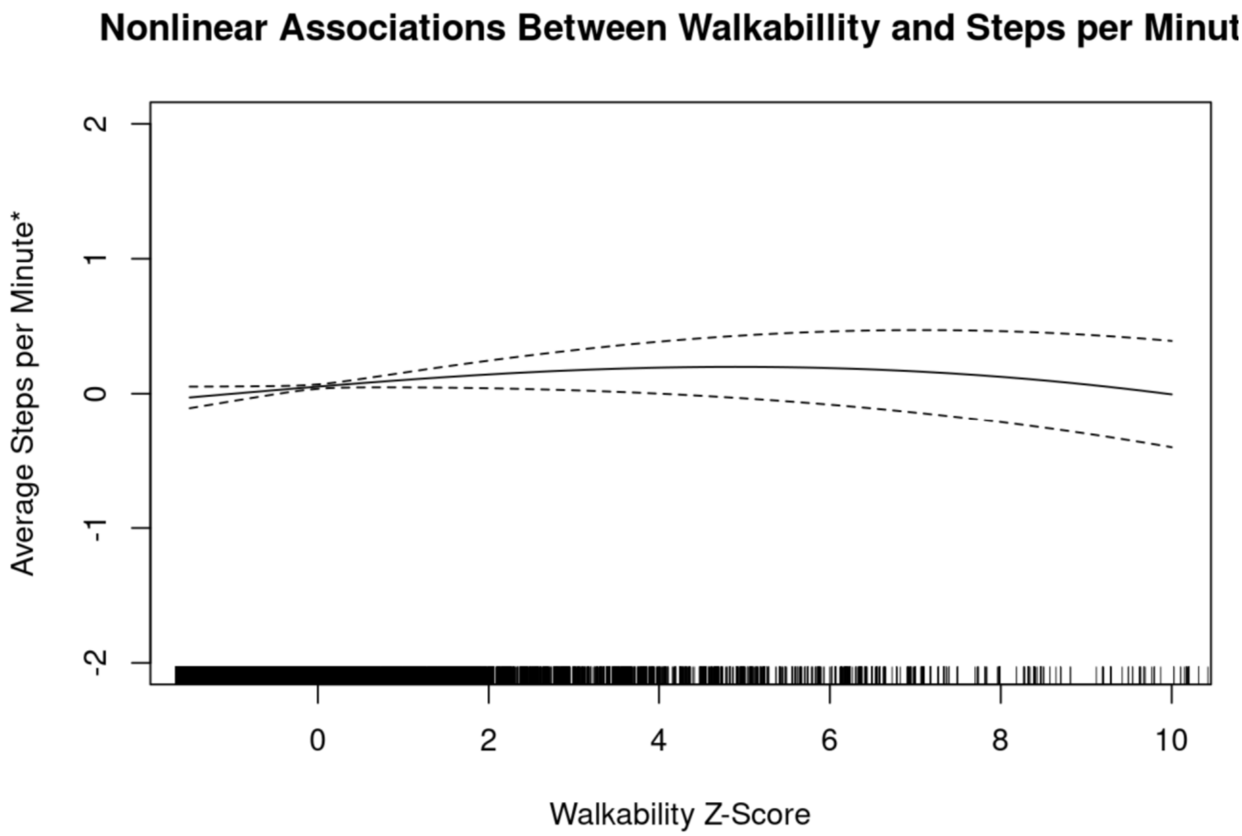


Figure S4.6. Nonlinear Associations Between Walkability and Steps per Minute Restricted to the Secondary Analyic Dataset, controlling for age (years; continuous), socioeconomic status defined as: education level (masters in nursing or higher; binary), and marital status (never [never married]]/ever [married, widowed, divorced]; binary), and area-level measures of neighborhood socioeconomic status (z-score; quartiles), greenness (NDVI; quartiles), mean daily temperature (Celsius; quartiles), daily precipitation (millimeters; binary), season and Census region. *Average steps per minute across each ten-minute collection period

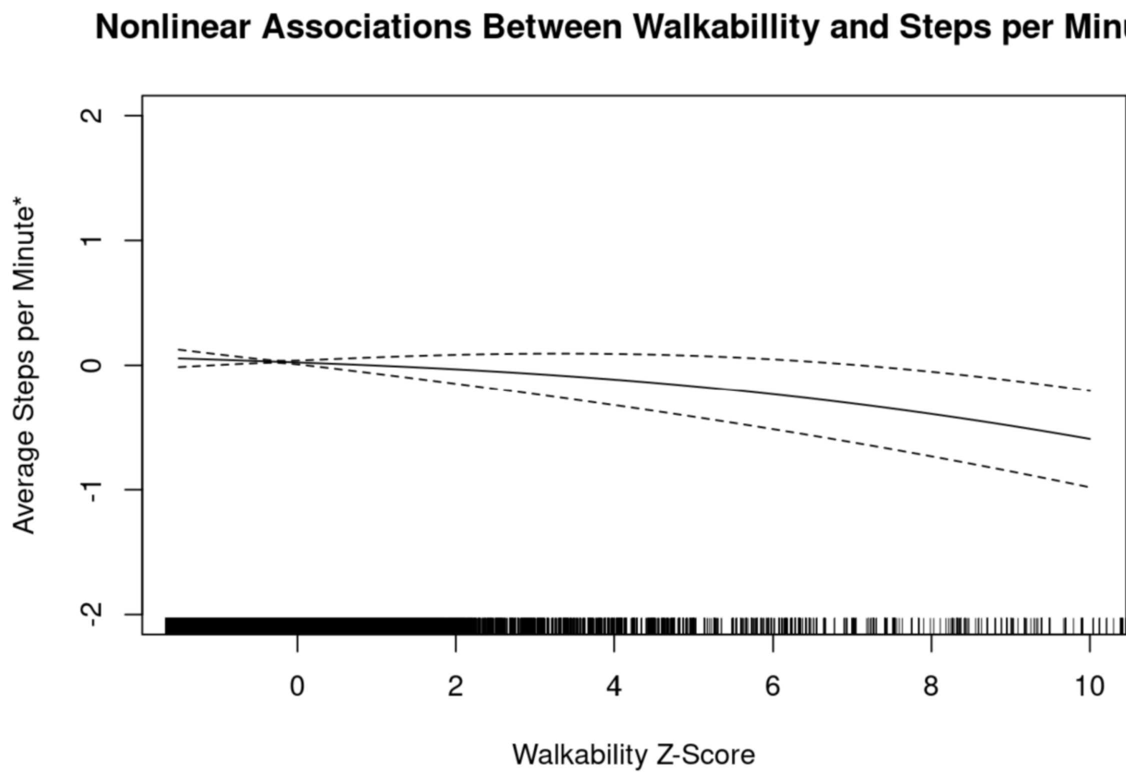


Figure S5.1. Histogram distributions for Residential Measures and Averaged Mobility Exposure Measures of Walkability Z-Score in Secondary Restriction Criteria Dataset (N=208)

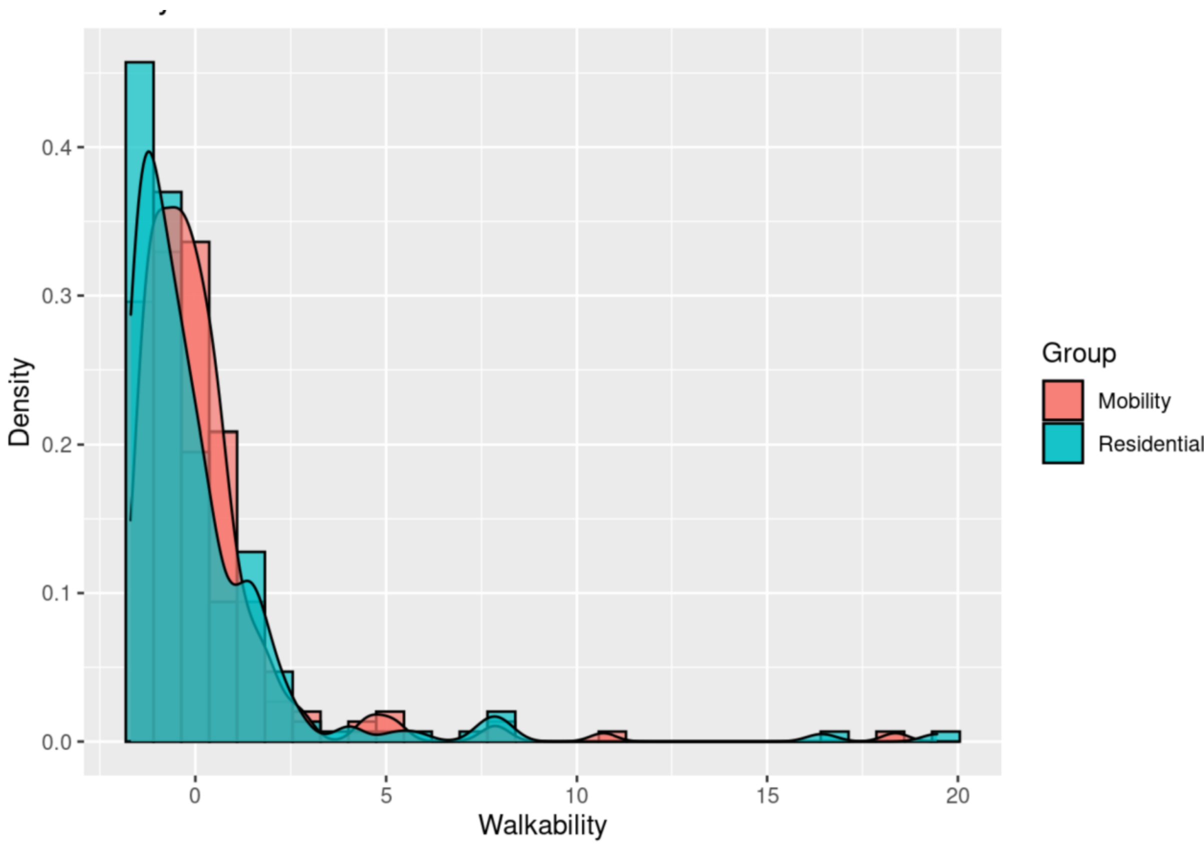
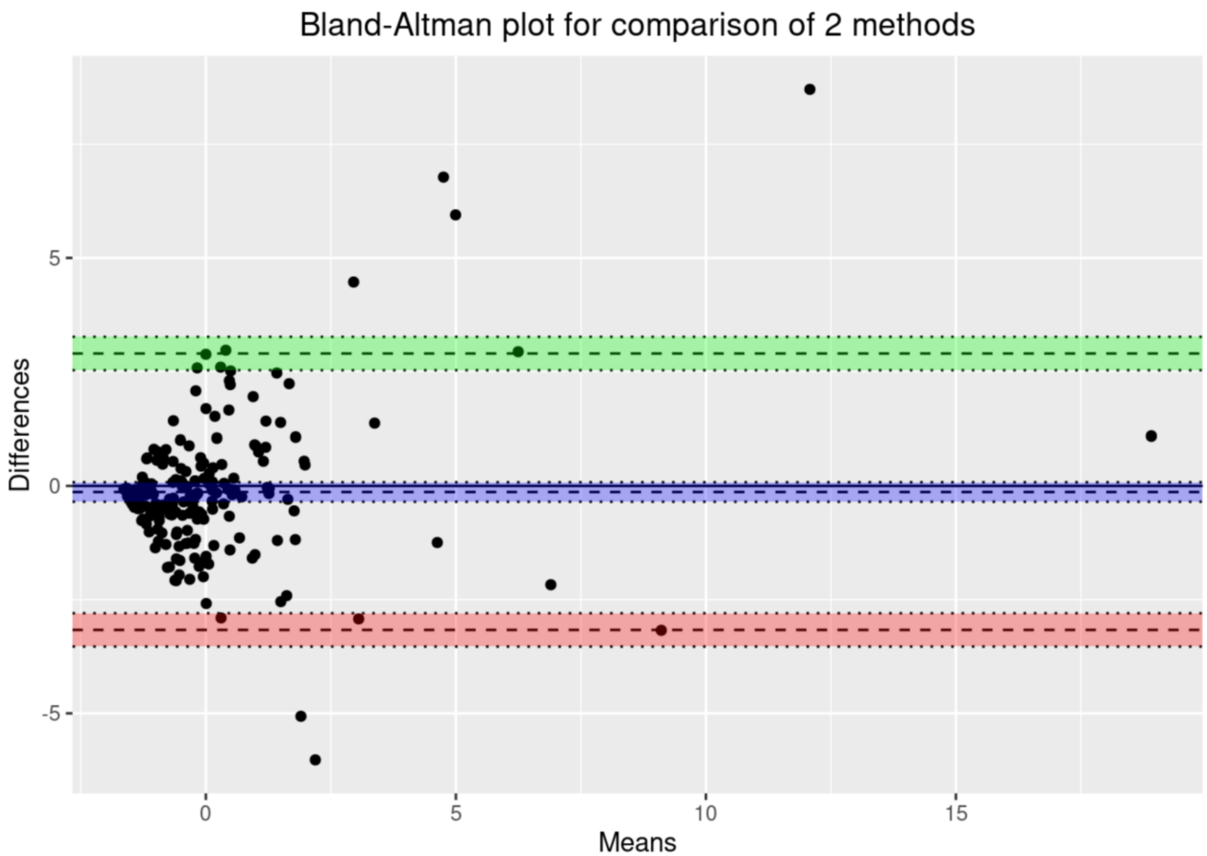


Figure S5.2. Bland Altman Plots Comparing Mobility and Residential Walkability Z-Score Measures. Purple band indicates 95% confidence level with red and green bands as lower and upper limit of agreement confidence bands



REFERENCES

Almanza, E., Jerrett, M., Dunton, G., Seto, E., & Ann Pentz, M. (2012). A study of community design, greenness, and physical activity in children using satellite, GPS and accelerometer data. *Health & Place*, *18*(1), 46–54. https://doi.org/10.1016/j.healthplace.2011.09.003

Ambrey, C. L. (2016). Urban greenspace, physical activity and wellbeing: The moderating role of perceptions of neighbourhood affability and incivility. *Land Use Policy*, *57*, 638–644. https://doi.org/10.1016/j.landusepol.2016.06.034

Armstrong, M., Winnard, A., Chynkiamis, N., Boyle, S., Burtin, C., & Vogiatzis, I. (2019). Use of pedometers as a tool to promote daily physical activity levels in patients with COPD: A systematic review and meta-analysis. *European Respiratory Review*, *28*(154). https://doi.org/10.1183/16000617.0039-2019

Bao, Y., Bertoia, M. L., Lenart, E. B., Stampfer, M. J., Willett, W. C., Speizer, F. E., & Chavarro, J. E. (2016). Origin, Methods, and Evolution of the Three Nurses' Health Studies. *American Journal of Public Health*, *106*(9), 1573–1581. https://doi.org/10.2105/AJPH.2016.303338

Baobeid, A., Koç, M., & Al-Ghamdi, S. G. (2021). Walkability and Its Relationships With Health, Sustainability, and Livability: Elements of Physical Environment and Evaluation Frameworks. *Frontiers in Built Environment*, *7*. https://www.frontiersin.org/articles/10.3389/fbuil.2021.721218

Bhopal, R. S. (2016). Error, bias, and confounding in epidemiology. In *Concepts of Epidemiology*. Oxford University Press.

https://oxfordmedicine.com/view/10.1093/med/9780198739685.001.0001/med-9780198739685-chapter-4

Blair, A., Stewart, P., Lubin, J. H., & Forastiere, F. (2007). Methodological issues regarding confounding and exposure misclassification in epidemiological studies of occupational exposures. *American Journal of Industrial Medicine*, *50*(3), 199–207. https://doi.org/10.1002/ajim.20281

Boston's new chief of streets to steer city beyond car-centric transportation. (n.d.). Smart Cities Dive. Retrieved January 31, 2023, from https://www.smartcitiesdive.com/news/bostons-new-chief-of-streets-to-steer-city-beyond-car-centric-transportati/611285/

Brokamp, C., Lemasters, G. K., & Ryan, P. H. (2016). Residential mobility impacts exposure assessment and community socioeconomic characteristics in longitudinal epidemiology studies. *Journal of Exposure Science and Environmental Epidemiology; Tuxedo*, *26*(4), 428–434. http://dx.doi.org.ezp-prod1.hul.harvard.edu/10.1038/jes.2016.10

Bureau, U. C. (n.d.). *TIGER/Line Shapefiles*. Census.Gov. Retrieved February 15, 2022, from https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.html

Carr, L. J., Dunsiger, S. I., & Marcus, B. H. (2010). Walk Score™ As a Global Estimate of Neighborhood Walkability. *American Journal of Preventive Medicine*, *39*(5), 460–463. https://doi.org/10.1016/j.amepre.2010.07.007

CDC. (2013, August 6). *CDC VitalSigns—More People Walk to Better Health*. Centers for Disease Control and Prevention. https://www.cdc.gov/vitalsigns/walking/index.html

CDC. (2022, June 2). *Move More; Sit Less*. Centers for Disease Control and Prevention. https://www.cdc.gov/physicalactivity/basics/adults/index.htm

CDCMMWR. (2019). QuickStats: Percentage of Adults Who Met Federal Guidelines for Aerobic Physical Activity Through Leisure-Time Activity, by Race/Ethnicity — National Health Interview Survey, 2008–2017. *MMWR. Morbidity and Mortality Weekly Report*, *68*. https://doi.org/10.15585/mmwr.mm6812a6

Chaix, B., Billaudeau, N., Thomas, F., Havard, S., Evans, D., Kestens, Y., & Bean, K. (2011). Neighborhood effects on health: Correcting bias from neighborhood effects on participation. *Epidemiology (Cambridge, Mass.)*, *22*(1), 18–26. https://doi.org/10.1097/EDE.0b013e3181fd2961

Chaix, B., Kestens, Y., Perchoux, C., Karusisi, N., Merlo, J., & Labadi, K. (2012). An Interactive Mapping Tool to Assess Individual Mobility Patterns in Neighborhood Studies. *American Journal of Preventive Medicine*, *43*(4), 440–450. https://doi.org/10.1016/j.amepre.2012.06.026

Chaix, B., Méline, J., Duncan, S., Merrien, C., Karusisi, N., Perchoux, C., Lewin, A., Labadi, K., & Kestens, Y. (2013). GPS tracking in neighborhood and health studies: A step forward for environmental exposure assessment, a step backward for causal inference? *Health & Place*, *21*, 46–51. https://doi.org/10.1016/j.healthplace.2013.01.003

Chavarro, J. E., Rich-Edwards, J. W., Gaskins, A. J., Farland, L. V., Terry, K. L., Zhang, C., & Missmer, S. A. (2016). Contributions of the Nurses' Health Studies to Reproductive Health Research. *American Journal of Public Health*, *106*(9), 1669–1676. https://doi.org/10.2105/AJPH.2016.303350

Chong, S., Mazumdar, S., Ding, D., Morgan, G., Comino, E. J., Bauman, A., & Jalaludin, B. (2019). Neighbourhood greenspace and physical activity and sedentary behaviour among older adults with a recent diagnosis of type 2 diabetes: A prospective analysis. *BMJ Open*, *9*(11), e028947. https://doi.org/10.1136/bmjopen-2019-028947

Coombes, E., Jones, A. P., & Hillsdon, M. (2010). The relationship of physical activity and overweight to objectively measured green space accessibility and use. *Social Science & Medicine (1982)*, *70*(6), 816–822. https://doi.org/10.1016/j.socscimed.2009.11.020

Coutts, C., Chapin, T., Horner, M., & Taylor, C. (2013). County-Level Effects of Green Space Access on Physical Activity. *Journal of Physical Activity and Health*, *10*(2), 232–240. https://doi.org/10.1123/jpah.10.2.232

Cruciani, F., Cleland, I., Nugent, C., McCullagh, P., Synnes, K., & Hallberg, J. (2018). Automatic Annotation for Human Activity Recognition in Free Living Using a Smartphone. *Sensors*, *18*(7), Article 7. https://doi.org/10.3390/s18072203

Cusack, L., Larkin, A., Carozza, S. E., & Hystad, P. (2017). Associations between multiple green space measures and birth weight across two US cities. *Health & Place*, *47*, 36–43. https://doi.org/10.1016/j.healthplace.2017.07.002

Dadvand, P., Villanueva, C. M., Font-Ribera, L., Martinez, D., Basagaña, X., Belmonte, J., Vrijheid, M., Gražulevičienė, R., Kogevinas, M., & Nieuwenhuijsen, M. J. (2014). Risks and Benefits of Green Spaces for Children: A Cross-Sectional Study of Associations with Sedentary Behavior, Obesity, Asthma, and Allergy. *Environmental Health Perspectives*, *122*(12), 1329–1335. https://doi.org/10.1289/ehp.1308038

Dai, D. (2011). Racial/ethnic and socioeconomic disparities in urban green space accessibility: Where to intervene? *Landscape and Urban Planning*, *102*(4), 234–244. https://doi.org/10.1016/j.landurbplan.2011.05.002

DeVille, N. V. (n.d.). Neighborhood Socioeconomic Status and Mortality in the Nurses' Health Study (NHS) and the Nurses' Health Study II (NHSII). *Environmental Epidemiology*, *In Press*.

Dewulf, B., Neutens, T., Van Dyck, D., De Bourdeaudhuij, I., Broekx, S., Beckx, C., & Van de Weghe, N. (2016). Associations between time spent in green areas and physical activity among late middle-aged adults. *Geospatial Health*, *11*(3), 411. https://doi.org/10.4081/gh.2016.411

Diaz, K. M., Krupka, D. J., Chang, M. J., Shaffer, J. A., Ma, Y., Goldsmith, J., Schwartz, J. E., & Davidson, K. W. (2016). Validation of the Fitbit One® for physical activity measurement at an upper torso attachment site. *BMC Research Notes*, *9*(1), 213. https://doi.org/10.1186/s13104-016-2020-8

Dockery, D. W., Pope, C. A., Xu, X., Spengler, J. D., Ware, J. H., Fay, M. E., Ferris, B. G., & Speizer, F. E. (1993). An Association between Air Pollution and Mortality in Six U.S. Cities. *New England Journal of Medicine*, *329*(24), 1753–1759. https://doi.org/10.1056/NEJM199312093292401

Explore Census Data. (n.d.). Retrieved February 15, 2022, from https://data.census.gov/cedsci/

Feng, Y., Wei, Y., Coull, B. A., & Schwartz, J. D. (2023). Measurement error correction for ambient PM2.5 exposure using stratified regression calibration: Effects on all-cause mortality. *Environmental Research*, *216*, 114792. https://doi.org/10.1016/j.envres.2022.114792

Fong, K. C., Hart, J. E., & James, P. (2018). A Review of Epidemiologic Studies on Greenness and Health: Updated Literature Through 2017. *Current Environmental Health Reports*, *5*(1), 77–87. https://doi.org/10.1007/s40572-018-0179-y

Fore, R., Hart, J. E., Choirat, C., Thompson, J. W., Lynch, K., Laden, F., Chavarro, J. E., & James, P. (2020). Embedding Mobile Health Technology into the Nurses' Health Study 3 to Study Behavioral Risk Factors for Cancer. *Cancer Epidemiology Biomarkers & Prevention*, *29*(4), 736–743. https://doi.org/10.1158/1055-9965.EPI-19-1386

Frank, L. D., Kerr, J., Sallis, J. F., Miles, R., & Chapman, J. (2008). A hierarchy of sociodemographic and environmental correlates of walking and obesity. *Preventive Medicine*, *47*(2), 172–178. https://doi.org/10.1016/j.ypmed.2008.04.004

Garrett, J. K., White, M. P., Elliott, L. R., Wheeler, B. W., & Fleming, L. E. (2020). Urban nature and physical activity: Investigating associations using self-reported and accelerometer data and the role of household income. *Environmental Research*, *190*, 109899. https://doi.org/10.1016/j.envres.2020.109899

Gaskins, A. J., Rich-Edwards, J. W., Lawson, C. C., Schernhammer, E. S., Missmer, S. A., & Chavarro, J. E. (2015). Work schedule and physical factors in relation to fecundity in nurses. *Occupational and Environmental Medicine*, *72*(11), 777–783. https://doi.org/10.1136/oemed-2015-103026

Gaskins, A. J., Rich-Edwards, J. W., Missmer, S. A., Rosner, B., & Chavarro, J. E. (2015). Association of Fecundity With Changes in Adult Female Weight. *Obstetrics and Gynecology*, *126*(4), 850–858. https://doi.org/10.1097/AOG.0000000000001030

Giles-Corti, B., Vernez-Moudon, A., Reis, R., Turrell, G., Dannenberg, A. L., Badland, H., Foster, S., Lowe, M., Sallis, J. F., Stevenson, M., & Owen, N. (2016). City planning and population health: A global challenge. *Lancet (London, England)*, *388*(10062), 2912–2924. https://doi.org/10.1016/S0140-6736(16)30066-6

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment, 202*, 18–27. https://doi.org/10.1016/j.rse.2017.06.031

Grigsby-Toussaint, D. S., Chi, S.-H., & Fiese, B. H. (2011). Where they live, how they play: Neighborhood greenness and outdoor physical activity among preschoolers. *International Journal of Health Geographics*, *10*(1), 66. https://doi.org/10.1186/1476-072X-10-66

Halonen, J. I., Pulakka, A., Pentti, J., Kallio, M., Koskela, S., Kivimäki, M., Kawachi, I., Vahtera, J., & Stenholm, S. (2020). Cross-sectional associations of neighbourhood socioeconomic disadvantage and greenness with accelerometer-measured leisure-time physical activity in a cohort of ageing workers. *BMJ Open*, *10*(8), e038673. https://doi.org/10.1136/bmjopen-2020-038673

Hart, J. E., Liao, X., Hong, B., Puett, R. C., Yanosky, J. D., Suh, H., Kioumourtzoglou, M.-A., Spiegelman, D., & Laden, F. (2015a). The association of long-term exposure to PM2.5 on all-cause mortality in the Nurses' Health Study and the impact of measurement-error correction. *Environmental Health: A Global Access Science Source*, *14*, 38. https://doi.org/10.1186/s12940-015-0027-6

Hart, J. E., Liao, X., Hong, B., Puett, R. C., Yanosky, J. D., Suh, H., Kioumourtzoglou, M.-A., Spiegelman, D., & Laden, F. (2015b). The association of long-term exposure to PM2.5 on all-cause mortality in the Nurses' Health Study and the impact of measurement-error correction. *Environmental Health*, *14*(1), 38. https://doi.org/10.1186/s12940-015-0027-6

Hartig, T. (2008). Green space, psychological restoration, and health inequality. *Lancet (London, England)*, *372*(9650), 1614–1615. https://doi.org/10.1016/S0140-6736(08)61669-4

Hillsdon, M., Panter, J., Foster, C., & Jones, A. (2006). The relationship between access and quality of urban green space with population physical activity. *Public Health*, *120*(12), 1127–1132. https://doi.org/10.1016/j.puhe.2006.10.007

Hooper, P. L., Middleton, N., Knuiman, M., & Giles-Corti, B. (2013). Measurement Error in Studies of the Built Environment: Validating Commercial Data as Objective Measures of Neighborhood Destinations. *Journal of Physical Activity and Health*, *10*(6), 792–804. https://doi.org/10.1123/jpah.10.6.792

Humpel, N., Owen, N., & Leslie, E. (2002). Environmental factors associated with adults' participation in physical activity: A review. *American Journal of Preventive Medicine*, *22*(3), 188–199. https://doi.org/10.1016/S0749-3797(01)00426-3

Hystad, P., Amram, O., Oje, F., Larkin, A., Boakye, K., Avery, A., Gebremedhin, A., & Duncan, G. (n.d.). Bring Your Own Location Data: Use of Google Smartphone Location History Data for Environmental Health Research. *Environmental Health Perspectives, 130*(11), 117005. https://doi.org/10.1289/EHP10829

Influential Minneapolis Policy Shift Links Affordability, Equity. (n.d.). LILP. Retrieved January 31, 2023, from https://www.lincolninst.edu/publications/articles/2020-01-rezoning-history-minneapolis-policy-shift-links-affordability-equity

Infogroup. (2020). *Infogroup US Historical Business Data* [Data set]. Harvard Dataverse. https://doi.org/10.7910/DVN/PNOFKI

Insights from a Healthcare Architect's Journal. (2019, October 28). BSA Lifestructures. https://www.bsalifestructures.com/insights-from-a-healthcare-architects-journal

James, P. a, Kioumourtzoglou, M.-A. d, Hart, J. E. b, Banay, R. F. b, Kloog, I. e, & Laden, F. a. (2017). Interrelationships Between Walkability, Air Pollution, Greenness, and Body

Mass Index. [Miscellaneous Article]. *Epidemiology*, *28*(6), 780–788. https://doi.org/10.1097/EDE.0000000000000724

James, P., Banay, R. F., Hart, J. E., & Laden, F. (2015). A Review of the Health Benefits of Greenness. *Current Epidemiology Reports*, *2*(2), 131–142. https://doi.org/10.1007/s40471-015-0043-7

James, P., Hart, J. E., Banay, R. F., & Laden, F. (2016). Exposure to Greenness and Mortality in a Nationwide Prospective Cohort Study of Women. *Environmental Health Perspectives*, *124*(9), 1344–1352. https://doi.org/10.1289/ehp.1510363

James, P., Hart, J. E., Hipp, J. A., Mitchell, J. A., Kerr, J., Hurvitz, P. M., Glanz, K., & Laden, F. (2017). GPS-Based Exposure to Greenness and Walkability and Accelerometry-Based Physical Activity. *Cancer Epidemiology Biomarkers & Prevention*, *26*(4), 525–532. https://doi.org/10.1158/1055-9965.EPI-16-0925

James, P., Wilt, G. E., & Jimenez, M. P. (2022). Invited Perspective: Can My Smartphone Assess My Exposure? The Potential to Retroactively Estimate Personalized Exposures using Smartphone Location Data. *Environmental Health Perspectives*, *130*(11), 111304–1. https://doi.org/10.1289/EHP12237

Jimenez, M. P., DeVille, N. V., Elliott, E. G., Schiff, J. E., Wilt, G. E., Hart, J. E., & James, P. (2021). Associations between Nature Exposure and Health: A Review of the Evidence. *International Journal of Environmental Research and Public Health*, *18*(9), Article 9. https://doi.org/10.3390/ijerph18094790

Jimenez, M. P., Wellenius, G. A., James, P., Subramanian, S. V., Buka, S., Eaton, C., Gilman, S. E., & Loucks, E. B. (2020). Associations of types of green space across the life-course with blood pressure and body mass index. *Environmental Research*, *185*, 109411. https://doi.org/10.1016/j.envres.2020.109411

Jones, A., Hillsdon, M., & Coombes, E. (2009). Greenspace access, use, and physical activity: Understanding the effects of area deprivation. *Preventive Medicine*, *49*(6), 500–505. https://doi.org/10.1016/j.ypmed.2009.10.012

Kajosaari, A., & Pasanen, T. P. (2021). Restorative benefits of everyday green exercise: A spatial approach. *Landscape and Urban Planning*, *206*, 103978. https://doi.org/10.1016/j.landurbplan.2020.103978

Kaplan, S. (1995). The restorative benefits of nature: Toward an integrative framework. *Journal of Environmental Psychology*, *15*(3), 169–182. https://doi.org/10.1016/0272-4944(95)90001-2

Kärmeniemi, M., Lankila, T., Ikäheimo, T., Koivumaa-Honkanen, H., & Korpelainen, R. (2018). The Built Environment as a Determinant of Physical Activity: A Systematic Review of Longitudinal Studies and Natural Experiments. *Annals of Behavioral Medicine: A Publication of the Society of Behavioral Medicine*, *52*(3), 239–251. https://doi.org/10.1093/abm/kax043

Kim, J., & Kwan, M.-P. (2018). Beyond Commuting: Ignoring Individuals' Activity-Travel Patterns May Lead to Inaccurate Assessments of Their Exposure to Traffic Congestion. *International Journal of Environmental Research and Public Health*, *16*(1), 89. https://doi.org/10.3390/ijerph16010089

Klepeis, N. E., Nelson, W. C., Ott, W. R., Robinson, J. P., Tsang, A. M., Switzer, P., Behar, J. V., Hern, S. C., & Engelmann, W. H. (2001). The National Human Activity Pattern Survey (NHAPS): A resource for assessing exposure to environmental pollutants.

Journal of Exposure Science & Environmental Epidemiology, *11*(3), 231–252. https://doi.org/10.1038/sj.jea.7500165

Klompmaker, J. O., Hoek, G., Bloemsma, L. D., Gehring, U., Strak, M., Wijga, A. H., van den Brink, C., Brunekreef, B., Lebret, E., & Janssen, N. A. H. (2018). Green space definition affects associations of green space with overweight and physical activity. *Environmental Research*, *160*, 531–540. https://doi.org/10.1016/j.envres.2017.10.027

Kwan, M.-P. (2012a). The Uncertain Geographic Context Problem. *Annals of the Association of American Geographers*, *102*(5), 958–968. https://doi.org/10.1080/00045608.2012.687349

Kwan, M.-P. (2012b). How GIS can help address the uncertain geographic context problem in social science research. *Annals of GIS*, *18*(4), 245–255. https://doi.org/10.1080/19475683.2012.727867

Kwan, M.-P. (2019). *Human Mobility, Spatiotemporal Context, and Environmental Health: Recent Advances in Approaches and Methods*. MDPI - Multidisciplinary Digital Publishing Institute.

Labib, S. M., Huck, J. J., & Lindley, S. (2020). Modelling and mapping eye-level greenness visibility exposure using multi-source data at high spatial resolutions. *Science of The Total Environment*, 143050. https://doi.org/10.1016/j.scitotenv.2020.143050

Labib, S. M., Lindley, S., & Huck, J. J. (2020). Scale effects in remotely sensed greenspace metrics and how to mitigate them for environmental health exposure assessment. *Computers, Environment and Urban Systems*, *82*, 101501. https://doi.org/10.1016/j.compenvurbsys.2020.101501

Lachowycz, K., & Jones, A. P. (2011). Greenspace and obesity: A systematic review of the evidence. *Obesity Reviews*, *12*(5), e183–e189. https://doi.org/10.1111/j.1467-789X.2010.00827.x

Laden, F., Schwartz, J., Speizer, F. E., & Dockery, D. W. (2006). Reduction in fine particulate air pollution and mortality: Extended follow-up of the Harvard Six Cities study. *American Journal of Respiratory and Critical Care Medicine*, *173*(6), 667–672. https://doi.org/10.1164/rccm.200503-443OC

Luzio, M. D., Johnson, G. L., Daly, C., Eischeid, J. K., & Arnold, J. G. (2008). Constructing Retrospective Gridded Daily Precipitation and Temperature Datasets for the Conterminous United States. *Journal of Applied Meteorology and Climatology*, *47*(2), 475–497. https://doi.org/10.1175/2007JAMC1356.1

Maas, J., Verheij, R. A., Spreeuwenberg, P., & Groenewegen, P. P. (2008). Physical activity as a possible mechanism behind the relationship between green space and health: A multilevel analysis. *BMC Public Health*, *8*(1), 206. https://doi.org/10.1186/1471-2458-8-206

Markevych, I., Schoierer, J., Hartig, T., Chudnovsky, A., Hystad, P., Dzhambov, A. M., de Vries, S., Triguero-Mas, M., Brauer, M., Nieuwenhuijsen, M. J., Lupp, G., Richardson, E. A., Astell-Burt, T., Dimitrova, D., Feng, X., Sadeh, M., Standl, M., Heinrich, J., & Fuertes, E. (2017). Exploring pathways linking greenspace to health: Theoretical and methodological guidance. *Environmental Research*, *158*, 301–317. https://doi.org/10.1016/j.envres.2017.06.028

Marquet, O., Floyd, M. F., James, P., Glanz, K., Jennings, V., Jankowska, M. M., Kerr, J., & Hipp, J. A. (2020). Associations Between Worksite Walkability, Greenness, and Physical

Activity Around Work. *Environment and Behavior*, *52*(2), 139–163. https://doi.org/10.1177/0013916518797165

Marquet, O., Hirsch, J. A., Kerr, J., Jankowska, M. M., Mitchell, J., Hart, J. E., Laden, F., Hipp, J. A., & James, P. (2022a). GPS-based activity space exposure to greenness and walkability is associated with increased accelerometer-based physical activity. *Environment International*, *165*, 107317. https://doi.org/10.1016/j.envint.2022.107317

Marquet, O., Hirsch, J. A., Kerr, J., Jankowska, M. M., Mitchell, J., Hart, J. E., Laden, F., Hipp, J. A., & James, P. (2022b). GPS-based activity space exposure to greenness and walkability is associated with increased accelerometer-based physical activity. *Environment International*, *165*, 107317. https://doi.org/10.1016/j.envint.2022.107317

McCormack, G., McLaren, L., Salvo, G., & Blackstaffe, A. (2017). Changes in Objectively-Determined Walkability and Physical Activity in Adults: A Quasi-Longitudinal Residential Relocation Study. *International Journal of Environmental Research and Public Health*, *14*(5), 551. https://doi.org/10.3390/ijerph14050551

Middelweerd, A., Van Der Ploeg, H. P., Van Halteren, A., Twisk, J. W. R., Brug, J., & Te Velde, S. J. (2017). A Validation Study of the Fitbit One in Daily Life Using Different Time Intervals. *Medicine & Science in Sports & Exercise*, *49*(6), 1270–1279. https://doi.org/10.1249/MSS.0000000000001225

Mnich, C., Weyland, S., Jekauc, D., & Schipperijn, J. (2019). Psychosocial and Physiological Health Outcomes of Green Exercise in Children and Adolescents—A Systematic Review. *International Journal of Environmental Research and Public Health*, *16*(21). https://doi.org/10.3390/ijerph16214266

Mooney, S. J., & Garber, M. D. (2019). Sampling and Sampling Frames in Big Data Epidemiology. *Current Epidemiology Reports*, *6*(1), 14–22. https://doi.org/10.1007/s40471-019-0179-y

Mooney, S. J., Hurvitz, P. M., Moudon, A. V., Zhou, C., Dalmat, R., & Saelens, B. E. (2020). Residential neighborhood features associated with objectively measured walking near home: Revisiting walkability using the Automatic Context Measurement Tool (ACMT). *Health & Place*, *63*, 102332. https://doi.org/10.1016/j.healthplace.2020.102332

Nardone, A., Rudolph, K. E., Morello-Frosch, R., & Casey, J. A. (2021). Redlines and Greenspace: The Relationship between Historical Redlining and 2010 Greenspace across the United States. *Environmental Health Perspectives*, *129*(1), 017006. https://doi.org/10.1289/EHP7495

National Academies of Sciences, E., Division, H. and M., Board, F. and N., & Surveillance, C. on S. for I. P. A. (2019). Community Supports for Physical Activity. In *Implementing Strategies to Enhance Public Health Surveillance of Physical Activity in the United States*. National Academies Press (US). https://www.ncbi.nlm.nih.gov/books/NBK545629/

Nieuwenhuijsen, M. J., Khreis, H., Triguero-Mas, M., Gascon, M., & Dadvand, P. (2017). Fifty Shades of Green: Pathway to Healthy Urban Living. *Epidemiology (Cambridge, Mass.)*, *28*(1), 63–71. https://doi.org/10.1097/EDE.0000000000000549

Office of the Surgeon General (US). (2010). *The Surgeon General's Vision for a Healthy and Fit Nation*. Office of the Surgeon General (US). http://www.ncbi.nlm.nih.gov/books/NBK44660/

Office of the Surgeon General (US). (2015). *Step It Up! The Surgeon General's Call to Action to Promote Walking and Walkable Communities*. US Department of Health and Human Services. http://www.ncbi.nlm.nih.gov/books/NBK538433/

Orstad, S. L., McDonough, M. H., James, P., Klenosky, D. B., Laden, F., Mattson, M., & Troped, P. J. (2018). Neighborhood walkability and physical activity among older women: Tests of mediation by environmental perceptions and moderation by depressive symptoms. *Preventive Medicine*, *116*, 60–67. https://doi.org/10.1016/j.ypmed.2018.08.008

Park, Y. M., & Kwan, M.-P. (2017). Individual exposure estimates may be erroneous when spatiotemporal variability of air pollution and human mobility are ignored. *Health & Place*, *43*, 85–94. https://doi.org/10.1016/j.healthplace.2016.10.002

Perchoux, C., Chaix, B., Brondeel, R., & Kestens, Y. (2016). Residential buffer, perceived neighborhood, and individual activity space: New refinements in the definition of exposure areas – The RECORD Cohort Study. *Health & Place*, *40*, 116–122. https://doi.org/10.1016/j.healthplace.2016.05.004

Perchoux, C., Chaix, B., Cummins, S., & Kestens, Y. (2013). Conceptualization and measurement of environmental exposure in epidemiology: Accounting for activity space related to daily mobility. *Health & Place*, *21*, 86–93. https://doi.org/10.1016/j.healthplace.2013.01.005

Persson, Å., Möller, J., Engström, K., Sundström, M. L., & Nooijen, C. F. J. (2019). Is moving to a greener or less green area followed by changes in physical activity? *Health & Place*, *57*, 165–170. https://doi.org/10.1016/j.healthplace.2019.04.006

Plue, R., Jewett, L., & Widener, M. J. (2020). Considerations When Using Individual GPS Data in Food Environment Research: A Scoping Review of ‘Selective (Daily) Mobility Bias’ in GPS Exposure Studies and Its Relevance to the Retail Food Environment. In Y. Lu & E. Delmelle (Eds.), *Geospatial Technologies for Urban Health* (pp. 95–112). Springer International Publishing. https://doi.org/10.1007/978-3-030-19573-1_6

Prince, S. A., Adamo, K. B., Hamel, M., Hardt, J., Connor Gorber, S., & Tremblay, M. (2008). A comparison of direct versus self-report measures for assessing physical activity in adults: A systematic review. *International Journal of Behavioral Nutrition and Physical Activity*, *5*(1), 56. https://doi.org/10.1186/1479-5868-5-56

Rojas-Rueda, D., Nieuwenhuijsen, M. J., Gascon, M., Perez-Leon, D., & Mudu, P. (2019). Green spaces and mortality: A systematic review and meta-analysis of cohort studies. *The Lancet Planetary Health*, *3*(11), e469–e477. https://doi.org/10.1016/S2542-5196(19)30215-3

Roscoe, C., Sheridan, C., Geneshka, M., Hodgson, S., Vineis, P., Gulliver, J., & Fecht, D. (2022). Green Walkability and Physical Activity in UK Biobank: A Cross-Sectional Analysis of Adults in Greater London. *International Journal of Environmental Research and Public Health*, *19*(7), 4247. https://doi.org/10.3390/ijerph19074247

Rundle, A. G., Sheehan, D. M., Quinn, J. W., Bartley, K., Eisenhower, D., Bader, M. M. D., Lovasi, G. S., & Neckerman, K. M. (2016). Using GPS Data to Study Neighborhood Walkability and Physical Activity. *American Journal of Preventive Medicine*, *50*(3), e65–e72. https://doi.org/10.1016/j.amepre.2015.07.033

Saelens, B. E., Sallis, J. F., Black, J. B., & Chen, D. (2003). Neighborhood-Based Differences in Physical Activity: An Environment Scale Evaluation. *American Journal of Public Health*, *93*(9), 1552–1558. https://doi.org/10.2105/AJPH.93.9.1552

Salvo, D., Durand, C. P., Dooley, E. E., Johnson, A. M., Oluyomi, A., Gabriel, K. P., van den Berg, A. E., Perez, A., & Kohl, H. W. (2019). Reducing the Uncertain Geographic Context Problem in Physical Activity Research: The Houston TRAIN Study: 1607 Board #4 May 30 1:30 PM - 3:30 PM. *Medicine & Science in Sports & Exercise*, *51*(Supplement), 437. https://doi.org/10.1249/01.mss.0000561808.49993.53

Slaght, J., Sénéchal, M., Hrubeniuk, T. J., Mayo, A., & Bouchard, D. R. (2017). Walking Cadence to Exercise at Moderate Intensity for Adults: A Systematic Review. *Journal of Sports Medicine*, *2017*, 1–12. https://doi.org/10.1155/2017/4641203

Smith, A. (n.d.). *Smartphone Ownership—2013 Update*. 13.

Smith, M., Hosking, J., Woodward, A., Witten, K., MacMillan, A., Field, A., Baas, P., & Mackie, H. (2017). Systematic literature review of built environment effects on physical activity and active transport—An update and new findings on health equity. *The International Journal of Behavioral Nutrition and Physical Activity*, *14*(1), 158. https://doi.org/10.1186/s12966-017-0613-9

Spiegelman, D. (2010). Approaches to Uncertainty in Exposure Assessment in Environmental Epidemiology. *Annual Review of Public Health*, *31*, 149–163. https://doi.org/10.1146/annurev.publhealth.012809.103720

Surgeon General's Report on Physical Activity and Health | JAMA | JAMA Network. (n.d.). Retrieved February 24, 2023, from https://jamanetwork.com/journals/jama/article-abstract/406551

Suzuki, E., Mitsuhashi, T., Tsuda, T., & Yamamoto, E. (2017). A typology of four notions of confounding in epidemiology. *Journal of Epidemiology*, *27*(2), 49–55. https://doi.org/10.1016/j.je.2016.09.003

Thurston, S. W., Williams, P. L., Hauser, R., Hu, H., Hernandez-Avila, M., & Spiegelman, D. (2005). A comparison of regression calibration approaches for designs with internal validation data. *Journal of Statistical Planning and Inference*, *131*(1), 175–190. https://doi.org/10.1016/j.jspi.2003.12.015

Ulrich, R. S., Simons, R. F., Losito, B. D., Fiorito, E., Miles, M. A., & Zelson, M. (1991). Stress recovery during exposure to natural and urban environments. *Journal of Environmental Psychology*, *11*(3), 201–230. https://doi.org/10.1016/S0272-4944(05)80184-7

Ussery, E. N. (2017). Walking for Transportation or Leisure Among U.S. Women and Men—National Health Interview Survey, 2005–2015. *MMWR. Morbidity and Mortality Weekly Report*, *66*. https://doi.org/10.15585/mmwr.mm6625a1

Villanueva, K., Knuiman, M., Nathan, A., Giles-Corti, B., Christian, H., Foster, S., & Bull, F. (2014). The impact of neighborhood walkability on walking: Does it differ across adult life stage and does neighborhood buffer size matter? *Health & Place*, *25*, 43–46. https://doi.org/10.1016/j.healthplace.2013.10.005

Wei, Y., Qiu, X., Yazdi, M. D., Shtein, A., Shi, L., Yang, J., Peralta, A. A., Coull, B. A., & Schwartz, J. D. (2022). The Impact of Exposure Measurement Error on the Estimated Concentration-Response Relationship between Long-Term Exposure to PM2.5 and Mortality. *Environmental Health Perspectives*, *130*(7), 77006. https://doi.org/10.1289/EHP10389

Wheeler, B. W., Cooper, A. R., Page, A. S., & Jago, R. (2010). Greenspace and children's physical activity: A GPS/GIS analysis of the PEACH project. *Preventive Medicine, 51*(2), 148–152. https://doi.org/10.1016/j.ypmed.2010.06.001

Yuenyongchaiwat, K. (2016). Effects of 10,000 steps a day on physical and mental health in overweight participants in a community setting: A preliminary study. *Brazilian Journal of Physical Therapy*, *20*(4), 367. https://doi.org/10.1590/bjpt-rbf.2014.0160

Zenk, S. N., Kraft, A. N., Jones, K. K., & Matthews, S. A. (2019). Convergent validity of an activity-space survey for use in health research. *Health & Place*, *56*, 19–23. https://doi.org/10.1016/j.healthplace.2019.01.009

Zenk, S. N., Matthews, S. A., Kraft, A. N., & Jones, K. K. (2018). How many days of global positioning system (GPS) monitoring do you need to measure activity space environments in health research? *Health & Place*, *51*, 52–60. https://doi.org/10.1016/j.healthplace.2018.02.004

Zhang, L., Zhou, S., Kwan, M.-P., Chen, F., & Lin, R. (2018). Impacts of Individual Daily Greenspace Exposure on Health Based on Individual Activity Space and Structural Equation Modeling. *International Journal of Environmental Research and Public Health, 15*(10). https://doi.org/10.3390/ijerph15102323

Zhang, Y., Murray, A. T., & Turner, B. L. (2017). Optimizing green space locations to reduce daytime and nighttime urban heat island effects in Phoenix, Arizona. *Landscape and Urban Planning*, *165*, 162–171. https://doi.org/10.1016/j.landurbplan.2017.04.009

Zhao, P., Kwan, M.-P., & Zhou, S. (2018). The Uncertain Geographic Context Problem in the Analysis of the Relationships between Obesity and the Built Environment in Guangzhou.

International Journal of Environmental Research and Public Health, *15*(2), 308.

https://doi.org/10.3390/ijerph15020308