Preventive Medicine 121 (2019) 55–61

Contents lists available at ScienceDirect

Preventive Medicine

journal homepage: www.elsevier.com/locate/ypmed

How race, ethnicity, and income moderate the relationship between urban vegetation and physical activity in the United States☆

Kevin Lanza[a,*], Brian Stone Jr.[a], Regine Haardörfer[b]

[a] Georgia Institute of Technology, 245 Fourth St. NW, Atlanta, GA 30332-0155, USA
[b] Emory Rollins School of Public Health, 1518 Clifton Road NE, Atlanta, GA 30322-4201, USA

ARTICLE INFO

Keywords:
Physical activity
Trees
Environment and public health
Social justice
City planning

ABSTRACT

To facilitate physical activity interventions, researchers identify which factors associate with physical activity, such as vegetation levels of the surrounding environment. While most studies examining vegetation and physical activity find a positive correlation, the literature does not investigate how vegetation may have a varied effect on physical activity based on demographic composition. This study examined how race, ethnicity, and income moderate the relationship between both non-tree vegetation and tree canopy on the percentage of individuals participating in leisure-time physical activity per census tract. Physical activity data from 2013 to 2014 for 7842 census tracts across 25 US cities originated from the CDC's 500 Cities project. Aerial images from the USDA's National Agriculture Imagery Program were used to classify vegetation levels per tract. Demographic variables originated from the American Community Survey 2011–2015 5-year estimates. Tracts were stratified into four types (Black + low income, Hispanic + low income, White + high income, and remaining) and assessed through multilevel modeling as to whether tract type moderated the relationship between vegetation and physical activity. Results showed that non-tree vegetation negatively associated with physical activity across all census tract types, while tree canopy exhibited a mixed association with physical activity, based on tract type. These findings can spur further research into how vegetation impacts physical activity of different demographic groups, and potentially inform greenspace and tree planting installments in those areas at greatest risk for physical inactivity-related diseases.

1. Introduction

The United States has been in a physical inactivity crisis for years, with just over half of adults meeting national guidelines for aerobic physical activity in 2016 (US Centers for Disease Control and Prevention, 2017). This pattern of physical inactivity extends past the US and has significant public health implications. I. M. Lee et al. (2012), for example, estimate that physical inactivity caused nine percent of premature mortality, or > 5.3 million of the 57 million deaths that occurred worldwide, in 2008. In particular, medical research has identified physical inactivity as a risk factor for a host of chronic diseases including obesity, type II diabetes, cardiovascular disease, and 13 types of cancer (Lauby-Secretan et al., 2016; Li and Siegrist, 2012; Sigal et al., 2013).

To combat the health burden attributed to physical inactivity, researchers employ theories of health behavior to identify which factors are associated with physical activity. The ecological model of health behavior recognizes that there are multiple influences across multiple levels (e.g., biological characteristics as an influence within the individual-level and built environment features as an influence within the environment-level) on specific health behaviors (e.g., physical activity), and that these influences interact across these different levels. Ecological models draw the central conclusion that significant changes in health behavior require a combination of both individual- and environment-level interventions (Glanz et al., 2008).

Research has investigated the connection between vegetation, one feature of the environment, and physical activity. Most of these studies find a positive association between vegetation and physical activity, and measure vegetation in several ways, including: 1) greenness, 2) observation, and 3) greenspace access. Greenness — the relative density and condition of vegetation — is commonly measured as normalized difference vegetation index (NDVI), a widely accepted indicator of vegetative cover that is calculated for each pixel of a satellite image (US Geological Survey, 2015). In a study examining physical activity among

☆ This work was supported by the National Science Foundation [grant number 1520803].

* Corresponding author.
E-mail addresses: lanza.kevin@gatech.edu (K. Lanza), stone@gatech.edu (B. Stone), regine.haardoerfer@emory.edu (R. Haardörfer).

https://doi.org/10.1016/j.ypmed.2019.01.022
Received 12 May 2018; Received in revised form 14 January 2019; Accepted 27 January 2019
Available online 30 January 2019
0091-7435/

children in California, children who participated in > 20 min of daily exposure to greener spaces engaged in almost five times more physical activity per day than those with close to no exposure to greener spaces (95% CI 3.09–7.20) (Almanza et al., 2012). Tilt et al. (2007) found subjective greenness, measured as the total number of natural features reported by each study participant, exhibited a significant positive association with walking trips per month ($B = 2.047$, $p < 0.001$).

In observational studies, researchers travel on-site to assess vegetation levels of various environments. In Stockholm, Sweden, children in preschool environments with trees and shrubs averaged 21.5 steps per minute, compared to 17.7 steps per minute in delimited environments with little vegetation ($p < 0.001$) (Boldemann et al., 2006). Furthermore, Coley et al. (1997) found public spaces with trees in Chicago, Illinois, attracted more people and a more mixed group of youth and adults than areas with fewer trees.

Lastly, researchers utilize access to greenspace — vegetated areas including forest, trees, parks, allotments, or cemeteries — to understand the relationship between vegetation and physical activity (Bastian et al., 2012). A study in Bristol, United Kingdom, found that when individuals lived > 2250 m from formal greenspace (i.e., well-maintained greenspace with an organized layout and structured path network), the odds of achieving physical activity guidelines decreased by 24% ($p < 0.01$) (Coombes et al., 2010). Similarly, the proximity to parks and the number of nearby parks have exhibited significant positive associations with physical activity (Lee and Moudon, 2006; Norman et al., 2006).

Yet while research demonstrates that access to vegetated areas correlates with increased physical activity, areas with a higher percentage of minority and/or low-income residents have significantly poorer access to greenspace (Astell-Burt et al., 2014; Dai, 2011; Heynen et al., 2006). These populations that lack greenspace access are the same populations that are most inactive: For US adults in 2014, 38.5% of Blacks and 40.1% of Hispanics/Latinos were inactive compared to 28.9% of Whites. Furthermore, 40.6% of those with annual family income less than $35,000 were inactive compared to 24.2% of those with income above $35,000 (Blackwell and Lucas, 2015). The inability to access greenspace for these at-risk populations is an environmental justice issue that requires better understanding if society is to lower mortality attributed to heart disease and cancer, which in 2014 accounted for nearly half of all US deaths (US National Center for Health Statistics, 2016).

Based on the current state of the literature, this investigation serves to fill two knowledge gaps: 1) No studies examine the associations between different types of vegetation and physical activity through classification of high-resolution satellite imagery, and 2) no studies examine how the effect of vegetation on physical activity differs by demographic group. Specifically, this study asks the question, "How do race, ethnicity, and income moderate the relationship between urban vegetation and leisure-time physical activity in the United States?" The analysis concentrates on three demographic groups (i.e., Black + low income, Hispanic + low income, and White + high income) positioned at the two ends of the environmental justice spectrum, and two categories of vegetation: non-tree vegetation and tree canopy.

2. Methods

2.1. Sample selection

This cross-sectional ecological study focused on 7842 census tracts for 25 US cities (Table 1). First, potential cities were selected from a list of the 50 most highly-populated cities, based on the 2010 US Census. Within a geographic information system (ESRI ArcGIS, version 10.5.1), vegetation percentage was calculated for each city based on the 2011 National Land Cover Database (i.e., class values 21, 41, 42, 43, 52, and 71) (Homer et al., 2015). The 50 cities were categorized into tertiles (i.e., low, medium, and high) of vegetation percentage. To reach 25 study cities, selection aimed for equal representation from each of the vegetation tertiles and the four US census regions. Because only three of the 50 most highly-populated cities were in the Northeast, cities bordering the Northeast (i.e., Baltimore, Cleveland, and District of Columbia) were selected.

Table 1
Census tracts (n = 7842) of the 25 study cities.

City	Vegetation tertile	Census region	# of census tracts
Atlanta	High	South	128
Austin	High	South	196
Baltimore	Medium	South	198
Boston	Low	Northeast	172
Chicago	Low	Midwest	793
Cleveland	Low	Midwest	173
Colorado Springs	High	West	103
Detroit	Low	Midwest	291
District of Columbia	Medium	South	178
Indianapolis	Medium	Midwest	214
Kansas City	Medium	Midwest	154
Las Vegas	Medium	West	153
Los Angeles	Medium	West	990
Miami	Low	South	100
Milwaukee	Low	Midwest	212
Nashville	High	South	156
New York	Low	Northeast	2109
Oakland	Medium	West	111
Oklahoma City	High	South	216
Philadelphia	Low	Northeast	375
Phoenix	High	West	355
Raleigh	High	South	92
Seattle	Low	West	133
Tucson	High	West	134
Wichita	Medium	Midwest	106

2.2. Variables

The dependent variable of interest was the percentage of adults aged ≥18 years who participated in leisure-time physical activity per census tract, which originated from the publicly available 500 Cities project, a joint project between the Robert Wood Johnson Foundation and the US Centers for Disease Control and Prevention that reports census tract-level health data through small area estimation methods for the 500 largest American cities (US Centers for Disease Control and Prevention, 2016). All estimated health data are based on the Behavioral Risk Factor Surveillance System (BRFSS), a state-based, random-digit-dialed telephone survey of the US civilian, noninstitutionalized population aged ≥18 years. The 500 Cities project packages these health data, including physical activity, at the census tract-level for all 500 cities as a shapefile, intended for analyses within a geographic information system (GIS). Specifically, the continuous dependent variable came from the BRFSS item (2013–2014), "During the past month, other than your regular job, did you participate in any physical activities or exercises such as running, calisthenics, golf, gardening, or walking for exercise?," with the dependent variable being the percentage of survey participants per census tract who responded "yes" to this question. This BRFSS question was selected because the 500 Cities project only estimates census tract-level data for this particular physical activity item.

There were two vegetation variables: 1) percentage of census tract covered by non-tree vegetation – all vegetation not considered trees (e.g., grasses and shrubs), and 2) percentage of census tract covered by tree canopy – the "layer of tree leaves, branches, and stems that provide coverage of the ground when viewed from above" (US Department of Agriculture, 2018). These continuous variables were created from aerial imagery acquired by the USDA's National Agriculture Imagery Program (NAIP) and taken during the US agricultural growing season from 2013 to 2016. NAIP products are available online as digital ortho quarter quad tiles at a relatively low cost from the USDA's Farm Service Agency website (US Department of Agriculture, n.d.). NAIP imagery was selected because of its display of vegetation at its peak greenness, complete spatial coverage of the 25 cities, high resolution (i.e., 60-cm to one-meter ground sample distance), and inclusion of four spectral bands (i.e., red, green, blue, and near infrared) for subsequent image classification, the process of categorizing image pixels to obtain a set of information classes.

To calculate vegetation variables, NAIP imagery and a census tract shapefile from the 500 Cities project were imported into GIS. A mosaicked, 4-band NAIP image for each city served as the input raster file for image classification within an image processing program (ERDAS IMAGINE, 2015). Utilizing unsupervised classification, ERDAS IMAGINE separated the NAIP image pixels based on their reflectance values into unique classes through K-means clustering – a method in which classes are statistically determined by assigning pixels to the nearest cluster mean based on all spectral bands – with no direction from the authors other than setting 100 classes for pixel grouping (Al-doski et al., 2013). Within GIS, the 100 classes were then reclassified based on visual inspection by the authors into three classes: non-tree vegetation, tree canopy, and non-vegetation. Lastly, zonal statistics were employed to calculate the percentage of pixels for each census tract occupied by non-tree vegetation and tree canopy.

An accuracy assessment of the image classification was performed for each of the 25 study cities. For randomly selected test points across each city, the accuracy assessment evaluated whether the assigned class in the classified image matched the landcover found in the high-resolution NAIP image. The assessment adopted an accuracy target of > 85% correct allocation, the universally accepted accuracy threshold developed by Anderson (1976). The formula for binomial probability theory (Eq. (1)) was used to determine the number of randomly selected test points per city (Jensen, 2011):

$$N = \frac{Z^2(p)(q)}{E^2} \tag{1}$$

N = sample size
Z = 2 from the standard normal deviate of 1.96 for the 95% two-sided confidence interval
p = expected percent accuracy
q = 100 − p
E = allowable error

Based on this formula, 203 randomly selected test points were required per study city for an expected map accuracy of 85% and an acceptable error of 5%. The accuracy assessment showed overall classification accuracy of each city's classified image was above the 85% target accuracy, with cities exhibiting average, minimum, and maximum classification accuracies of 92.9%, 85.6%, and 96.7%, respectively.

The study included the demographic characteristics of age, sex, race, ethnicity, and median household income, all of which have been found to exhibit significant associations with physical activity (Hallal et al., 2012; Kruger et al., 2005). These characteristics were collected at the census tract-level and originated from the American Community Survey (2011–2015) 5-year estimates by the US Census Bureau (Table ID #'s B03002, B19013, and S0101). The median age of individuals in a census tract and the percentage of females in a census tract were included as potential confounders. This work also controlled for population density, as the literature shows positive associations between compactness and physical activity (Ewing et al., 2008; Garden and Jalaludin, 2009). Density was defined as population per square mile per census tract, with population and area data originating from the 2010 US Census and the census tract shapefile provided by the 500 Cities project, respectively. Lastly, this work controlled for safety, as research shows vegetation of an open character positively associates with perceived safety, while extensive tree canopy is commonly perceived as potentially concealing threats and thus less safe (Haans and De Kort, 2012; Jansson et al., 2013; Loewen et al., 1993). As a proxy for safety, crime rate – the lone city-level variable in the analysis – was defined as violent crime rate per 100,000 population, in 2014, with crime data originating from Uniform Crime Reports (US Department of Justice, 2018).

For race, ethnicity, and income, census tracts were stratified into four types: 1) Black + low income, 2) Hispanic + low income, 3) White + high income, and 4) remaining. Tract types 1–3 were defined as having greater than two-thirds Black, Hispanic, and White populations, respectively, in addition to having a median household income in the bottom tertile (i.e., low income = $35,648) or top tertile (i.e., high income = $59,426) across all study tracts. These three combinations of race, ethnicity, and income elements were selected because 1) physical activity surveillance shows high minority and/or low-income areas have the least access to greenspace and are the most physically inactive, while low minority and/or high-income areas have the greatest access to greenspace and are the most physically active (Blackwell and Lucas, 2015; Astell-Burt et al., 2014; Dai, 2011; Heynen et al., 2006); and 2) each of these census tract types had a sufficient number of tracts (i.e., 12.1% Black + low income tracts, 6.9% Hispanic + low income tracts, and 16.5% White + high income tracts) based on available demographic data and chosen bounds for race, ethnicity, and income. Tracts without these combinations of race, ethnicity, and income were categorized as "remaining." The study included dummy variables for tract types 1–3.

To understand how race, ethnicity, and income moderate the relationship between vegetation and physical activity, the study included six interaction terms: 1) Black + low income ∗ non-tree vegetation percentage, 2) Black + low income ∗ tree canopy percentage, 3) Hispanic + low income ∗ non-tree vegetation percentage, 4) Hispanic + low income ∗ tree canopy percentage, 5) White + high income ∗ non-tree vegetation percentage, and 6) White + high income ∗ tree canopy percentage.

2.3. Statistical analysis

To assess how race, ethnicity, and income moderate the relationship between urban vegetation and physical activity, this work made use of a two-level hierarchical linear model (HLM) because of the nested structure of the data, i.e., census tracts (Level 1) grouped within cities (Level 2). Unlike ordinary least squares regression, HLM allows for simultaneous estimation of factors at multiple levels, accounting for nonindependence of observations within groups (Luke, 2004). To assess the appropriateness of HLM, the intraclass correlation (ICC) was computed from the unconditional model, a model with no variables introduced with only random error allowed to be free. The ICC was 31.5%, indicating that almost one third of the variance in the outcome is attributable to the city-level. This analysis selected the covariance structure of Variance Components (VC), in which a scaled identity structure is assigned to each specified effect. VC is the default structure in IBM SPSS Statistics (version 24), the statistical software used for this work.

The dependent variable was the percentage of adults participating in any leisure-time physical activity per census tract. The model included race, ethnicity, income, and vegetation as focal predictors, and age, sex, density, and crime rate as potential confounders (Table 2). To aid interpretation of model estimates, the variables of age and sex were mean-centered, and density was scaled by 10^3. To reduce the correlation between interaction terms and corresponding main effects, the continuous variables within interaction terms were mean-centered.

To better understand the combined effect of model estimates for interaction terms and corresponding main effects (e.g., Black + low income ∗ tree canopy, Black + low income, and tree canopy), the study predicted the impact of two hypothetical vegetation scenarios on mean

Table 2
Predictor variables included in multilevel model.

Predictor variable	Measurement description
Census tract variables (level 1)	
Demographic	
Sex[a]	% female per tract
Age[a]	Median age (in years) per tract
Black + low income	0 or 1 (1 ≥ two-thirds Black + lowest tertile of median household income tract)
Hispanic + low income	0 or 1 (1 ≥ two-thirds Hispanic + lowest tertile of median household income tract)
White + high income	0 or 1 (1 ≥ two-thirds White + highest tertile of median household income tract)
Environment	
Density	Population/mi^2/10^3 per tract
Non-tree vegetation	% non-tree vegetation per tract
Tree canopy	% tree canopy per tract
Interaction	
Black + low income ∗ non-tree vegetation[a]	0 or 1 (1 = Black + low income tract) ∗ % per tract
Black + low income ∗ tree canopy[a]	0 or 1 (1 = Black + low income tract) ∗ % per tract
Hispanic + low income ∗ non-tree vegetation[a]	0 or 1 (1 = Hispanic + low income tract) ∗ % per tract
Hispanic + low income ∗ tree canopy[a]	0 or 1 (1 = Hispanic + low income tract) ∗ % per tract
White + high income ∗ non-tree vegetation[a]	0 or 1 (1 = White + high income tract) ∗ % per tract
White + high income ∗ tree canopy[a]	0 or 1 (1 = White + high income tract) ∗ % per tract
City variables (level 2)	
Safety	
Crime rate	Violent crime rate per 100,000 population

[a] Mean-centered variables.

physical activity per census tract compared to a baseline scenario for each tract type. The two vegetation scenarios were a 10% increase in non-tree vegetation and tree canopy above baseline for each tract type, respectively. This study selected these percent increases based on adoption by other greenspace and tree canopy studies (Gill et al., 2007; Troy et al., 2012; University of Manchester, 2007). A baseline scenario, meant to represent a typical census tract, was developed separately for each tract type and defined as the mean values of predictor variables found to be significant ($p < 0.05$) in the final model. Based on the difference in physical activity from the vegetation scenarios to the baseline, this study utilized census tract-level population data from the 2010 US Census to predict the number of individuals per census tract that now participated in physical activity, stratified by tract type.

3. Results

Table 3 summarizes the descriptive statistics of the study census tracts. Across all tracts, the mean percentages of physical activity, non-tree vegetation, and tree canopy were 73.13, 13.96, and 15.87, respectively. Of the four tract types, Hispanic + low income tracts had the lowest mean percentages of physical activity, non-tree vegetation, and tree canopy. White + high income tracts had the highest percentages of physical activity and tree canopy, while Black + low income tracts had the highest percentage of non-tree vegetation. Hispanic + low income tracts were almost two and three times more densely populated than White + high income tracts and Black + low income tracts, respectively. Regarding safety, Black + low income tracts experienced violent crime rates almost double that of Hispanic + low income tracts.

A multilevel model predicted how race, ethnicity, and income moderated the relationship between leisure-time physical activity and both non-tree vegetation and tree canopy. Within the three-model set (Table 4), this analysis reported on Model 3 because this final, most elaborate model had the best fit, as determined by proportional reduction in variance, as well as lower values for Akaike information criterion (AIC) and Bayesian information criterion (BIC) from model to model signifying smaller deviance (i.e., better model fit). Model 3 accounted for −3.53% of the variance of the intercept (Level 2), and the statistically significant value for the unexplained variance at baseline indicates that there are other, omitted factors that can explain the

Table 3
Descriptive statistics: mean (M) and standard deviation (SD) for variables by census tract type.

	Black + low income (n = 948)	Hispanic + low income (n = 542)	White + high income (n = 1291)	Remaining[a] (n = 5061)
	M (SD)	M (SD)	M (SD)	M (SD)
Physical activity				
Participating in leisure-time physical activity (%)	63.11 (4.60)	62.84 (4.03)	84.43 (4.27)	73.23 (7.42)
Demographic				
Sex (% female)	54.61 (4.88)	50.31 (3.67)	51.33 (3.68)	51.03 (5.25)
Black (%)	88.96 (8.73)	9.75 (8.82)	3.71 (4.10)	23.54 (27.05)
Hispanic (%)	3.70 (5.33)	80.12 (9.01)	7.83 (5.02)	27.47 (22.86)
White (%)	4.69 (5.49)	6.09 (5.98)	79.38 (7.39)	35.28 (25.25)
Median age (years)	34.85 (6.87)	30.45 (4.85)	39.38 (6.65)	35.46 (6.24)
Median household income ($)	24,386 (6,260)	27,279 (5,254)	94,564 (29,248)	49,919 (20,497)
Environment				
Non-tree vegetation (%)	19.53 (9.39)	8.49 (6.28)	15.30 (12.27)	13.16 (9.66)
Tree canopy (%)	19.17 (11.31)	9.97 (5.53)	20.12 (14.38)	14.80 (10.19)
Density (population/mi^2/10^3)	6.89 (7.21)	20.05 (19.89)	10.83 (16.46)	13.87 (15.75)
Safety				
Crime rate (per 100,000 population)	1298.40 (437.16)	691.35 (306.22)	746.46 (283.07)	786.88 (334.91)

[a] Remaining = all tracts not categorized into the other tract types.

Table 4
HLM results.

	Model 1	Model 2	Model 3
Fixed effects			
Intercept	74.52***	78.02***	77.06***
Level 1-census tract variables			
Individual			
Sex (0 = 51.5% female)[a]		−0.03*	−0.04*
Age (0 = 35.7 years)[a]		0.11***	0.11***
Black + low income (0 = not Black + low income)		−9.51***	−9.34***
Hispanic + low income (0 = not Hispanic + low income)		−10.64***	−10.83***
White + high income (0 = not White + high income)		9.80***	10.09***
Environment			
Density (population/mi^2/10^3)		−0.01	−0.01
Non-tree vegetation (%)		−0.04***	−0.03*
Tree canopy (%)		0.03**	0.08***
Interaction			
Black + low income * non-tree vegetation			0.01
Black + low income * tree canopy			−0.13***
Hispanic + low income * non-tree vegetation			0.02
Hispanic + low income * tree canopy			−0.10
White + high income * non-tree vegetation			−0.06***
White + high income * tree canopy			−0.12***
Level 2-city variables			
Safety			
Crime rate (per 100,000 population)		−0.01*	−0.01*
Random effects			
τ_{00} (intercept)	65.62***	32.33***	31.80***
σ^2	30.15***	9.96***	10.31***
Model fit			
Reduction in τ_{00}		67.0%	−3.53%
Reduction in σ^2		50.7%	1.63%
AIC	55,180.7	49,549.6	49,449.9
BIC	55,194.6	49,563.6	49,463.8

Dependent variable: percentage of individuals participating in leisure-time physical activity.
* p < 0.05.
** p < 0.01.
*** p < 0.001.
[a] Mean-centered variables.

percentage of individuals participating in leisure-time physical activity. Model 3 accounted for 1.63% of the within group (Level 1) variance, with the remaining variance found to be statistically significant. The residuals for Model 3 were normally distributed.

The final model found that Black + low income tracts exhibited a 9.34 lower physical activity percentage than all other tracts ($p < 0.001$), Hispanic + low income tracts exhibited a −10.83 lower physical activity percentage than all other tracts ($p < 0.001$), and White + high income tracts exhibited a 10.09 higher physical activity percentage than all other tracts ($p < 0.001$), holding all other variables constant. For "remaining" tracts, which constituted approximately 65% of all study tracts, each one percent-increase in non-tree vegetation predicted a 0.03% decrease in physical activity ($p < 0.05$), and each one percent-increase in tree canopy predicted a 0.08% increase in physical activity ($p < 0.001$). Three of the six interaction terms were found to be statistically significant ($p < 0.001$), e.g., Model 3 found that every one percent-increase in tree canopy associated with a 0.13% decrease in physical activity over and above any effect predicted in Black + low income tracts.

Tables 5 and 6 show the predicted change in physically active individuals from the two vegetation scenarios over the baseline scenario per tract type across all sample tracts. A 10% addition of non-tree vegetation resulted in a decrease in physically active individuals over baseline for all tract types. If the 25 cities increase non-tree vegetation by 10% within all census tracts, then physical activity is predicted to decrease by 114,683 people overall. Conversely, a 10% addition of tree canopy had mixed effects based on tract type: Hispanic + low income tracts and "remaining" tracts exhibited increases in physically active individuals over baseline, while Black + low income tracts and White + high income tracts exhibited decreases in physically active individuals over baseline. If the 25 cities increase trees by 10% within all census tracts, then physical activity is predicted to increase by 99,718 people overall.

4. Discussion

This study found a negative association between non-tree vegetation and physical activity in all census tracts, and a mixed association between tree canopy and physical activity among the four tract types. The consistent negative association between non-tree vegetation and physical activity may relate to the link between vegetation and safety. The Prospect-Refuge theory posits that individuals prefer environments that balance outlook (i.e., prospect) and enclosure (i.e., refuge), which induce feelings of safety and spatially-derived pleasure (Appleton, 1996; Dosen and Ostwald, 2013). Vegetation of an open character has been shown to positively associate with perceived safety, as individuals are able to see potential threats from a distance (Jansson et al., 2013; Loewen et al., 1993). However, the negative association between non-tree vegetation and physical activity in this work may be a product of the non-tree vegetation quality within the open space. For instance, unkempt open spaces such as vacant lots – a common sight in marginalized communities – are oftentimes overgrown with unwanted vegetation, consequently attracting anti-social behavior (Garvin et al., 2013; Schukoske, 1999).

Regarding tree canopy and physical activity, the positive association in both Hispanic + low income tracts and "remaining" tracts may imply that trees serve as refuge from possible threats (Loewen et al., 1993). Trees may also positively correlate with physical activity in these tracts for their ability to 1) attract residents as a meeting place, 2) promote mental health, and 3) improve thermal comfort by cooling the ambient environment by 2–8 °C through shading and evapotranspiration (Coley et al., 1997; Doick and Hutchings, 2013; Huynh et al., 2013).

The negative association between tree canopy and physical activity in both Black + low income tracts and White + high income tracts may be linked to "refuge ambiguity," the phenomenon in which trees are perceived as places of concealment for potential offenders, rather than as safe havens (Haans and De Kort, 2012; Loewen et al., 1993). Another potential reason for the negative associations between both the

Table 5
Predicted percent change of physically active individuals from the two vegetation scenarios over the baseline scenario per tract type across all sample tracts (n = 7842).

	Black + low income	Hispanic + low income	White + high income	Remaining[a]
	% change	% change	% change	% change
+10% non-tree vegetation	−0.43	−0.41	−1.02	−0.35
+10% tree canopy	−0.69	1.29	−0.48	1.10

[a] Remaining = all tracts not categorized into the other three tract types.

Table 6
Predicted change in number of physically active individuals from the two vegetation scenarios over the baseline scenario per tract type across all sample tracts (n = 7842).

	Black + low income	Hispanic + low income	White + high income	Remaining[a]
	# change	# change	# change	# change
+10% non-tree vegetation	−11,250	−10,205	−57,278	−35,950
+10% tree canopy	−18,142	31,967	−26,719	112,612

[a] Remaining = all tracts not categorized into the other three tract types.

percentage of non-tree vegetation and tree canopy with physical activity is that areas with more vegetation lack infrastructure supporting leisure-time physical activity (e.g., sidewalks, bike lanes and paths, and public outdoor gyms). Large expanses of vegetation may also be characterized by recognized barriers to physical activity: poor road connectivity and limited access to scenic, interactive environments such as commercial centers (Boarnet et al., 2008; Chatman, 2009; Duncan et al., 2005).

If subsequent research discovers a causal link between demographics, vegetation, and physical activity, urban planners and landscape architects should design urban vegetation to maximize physical activity and safety, the latter of which has already been explored: Perceived safety can be improved by limbed-up trees and smooth, maintained ground surfaces (Herzog and Kutzli, 2002), and greening vacant lots can improve perceived safety and reduce gun crimes (Garvin et al., 2013). If vegetation has a varied effect on physical activity for different groups, then meaningful involvement of diverse populations throughout the decision-making process can maximize citywide physical activity levels. Furthermore, whether or not vegetation elicits physical activity, greenspace installation provides cities with stormwater management, heat management, air purification, carbon sequestration, and biodiversity enhancement (Food and Agriculture Administration, 2016; Kardan et al., 2015).

This work is not without limitations. A cause-and-effect relationship between predictors and physical activity cannot be ascertained from this work due to its cross-sectional nature and likelihood of omitted variables. In addition, the use of aggregated data at the census tract-level, in place of individual-level data, may incorrectly predict relationships between variables. The use of self-report for physical activity data has reliability and validity issues due to recall bias, social desirability bias, and misinterpretation of survey items, which results in measurement differences from direct measurement methods (e.g., accelerometer) (Prince et al., 2008). The research design assumes those living in a census tract are also performing activity within that census tract, and that the physical activity performed is outdoors where vegetation may impact activity. The analysis may exhibit temporal mismatch between vegetation and physical activity: BRFSS administers surveys throughout the year, and the physical activity item asks participants about activity "during the past month." Therefore, participants may not be responding during the US agricultural growing season, the period when the USDA captures NAIP imagery.

Subsequent exploration can build upon the design and methodology of this work. First, future research can control for vegetation density and quality (e.g., manicured versus unmaintained) that affect perceived safety and potentially physical activity (Bjerke et al., 2006; Sallis, 2009). Vacant parcels, if not maintained by the city, can serve as a proxy for vegetation quality. Second, research can distinguish between public and private greenspace, as park accessibility corresponds with physical activity (Fan et al., 2011). Additionally, future work can specify vegetation location, as the public cannot readily access backyard vegetation on private property. Third, studies can include built environment determinants of physical activity, such as physical activity infrastructure, road connectivity, and access to shops and services (Bauman et al., 2012). Lastly, future studies on the vegetation-physical activity relationship can categorize tract types based on different demographic characteristics, such as gender, as females express distinct relationships with physical activity and safety from males (Caspersen et al., 2000; Haans and De Kort, 2012).

5. Conclusion

The findings of this study assist researchers in understanding how different types of urban vegetation associate with areas of varying demographic composition, reveal new connections between factors correlated with physical activity within ecological models, and offer physical activity researchers a novel method for calculating vegetation levels. This work can potentially inform public health practitioners when implementing vegetation programming in cities, especially in Hispanic + low income tracts, the areas with the least vegetation and physical activity, although further confirmation of these associations is needed. For physical activity interventions, the importance of knowing which types of vegetation are suitable for which populations cannot be overstated, as physical inactivity is a modifiable risk factor for the world's deadliest chronic diseases. Vegetative enhancement, if applied with demographics in mind, can bring the ecosystem services of vegetation to marginalized communities, ultimately creating more equitable cities.

Conflicts of interest

The authors declare that there are no conflicts of interest.

References

Al-doski, J., Mansorl, S.B., Shafri, H.Z.M., 2013. Image Classification in Remote Sensing. Department of Civil Engineering, Faculty of Engineering, University Putra, Malaysia.

Almanza, E., Jerrett, M., Dunton, G., Seto, E., Pentz, M.A., 2012. A study of community design, greenness, and physical activity in children using satellite, GPS and accelerometer data. Health Place 18 (1), 46–54.

Anderson, J.R., 1976. A Land Use and Land Cover Classification System for Use With Remote Sensor Data. vol. 964 US Government Printing Office.

Appleton, J., 1996. The Experience of Landscape. Wiley, Chichester, pp. 66–67.

Astell-Burt, T., Feng, X., Mavoa, S., Badland, H.M., Giles-Corti, B., 2014. Do low-income neighbourhoods have the least green space? A cross-sectional study of Australia's most populous cities. BMC Public Health 14 (1), 292. https://doi.org/10.1186/1471-2458-14-292.

Bastian, O., Haase, D., Grunewald, K., 2012. Ecosystem properties, potentials and services–the EPPS conceptual framework and an urban application example. Ecol. Indic. 21, 7–16.

Bauman, A.E., Reis, R.S., Sallis, J.F., Wells, J.C., Loos, R.J., Martin, B.W., Lancet Physical Activity Series Working Group, 2012. Correlates of physical activity: why are some people physically active and others not? Lancet 380 (9838), 258–271.

Bjerke, T., Østdahl, T., Thrane, C., Strumse, E., 2006. Vegetation density of urban parks and perceived appropriateness for recreation. Urban For. Urban Green. 5 (1), 35–44.

Blackwell, D.L., Lucas, J.W., 2015. Tables of summary health statistics for U.S. adults: 2014 national health interview survey. Retrieved March 31, 2018 from. https://ftp.cdc.gov/pub/Health_Statistics/NCHS/NHIS/SHS/2014_SHS_Table_A-14.pdf.

Boarnet, M.G., Greenwald, M., McMillan, T.E., 2008. Walking, urban design, and health: toward a cost-benefit analysis framework. J. Plan. Educ. Res. 27 (3), 341–358.

Boldemann, C., Blennow, M., Dal, H., Mårtensson, F., Raustorp, A., Yuen, K., Wester, U., 2006. Impact of preschool environment upon children's physical activity and sun exposure. Prev. Med. 42 (4), 301–308.

Caspersen, C.J., Pereira, M.A., Curran, K.M., 2000. Changes in physical activity patterns in the United States, by sex and cross-sectional age. Med. Sci. Sports Exerc. 32 (9), 1601–1609.

Chatman, D.G., 2009. Residential choice, the built environment, and nonwork travel: evidence using new data and methods. Environ. Plan. A 41 (5), 1072–1089.

Coley, R.L., Sullivan, W.C., Kuo, F.E., 1997. Where does community grow? The social context created by nature in urban public housing. Environ. Behav. 29 (4), 468–494.
Coombes, E., Jones, A.P., Hillsdon, M., 2010. The relationship of physical activity and overweight to objectively measured green space accessibility and use. Soc. Sci. Med. 70 (6), 816–822.
Dai, D., 2011. Racial/ethnic and socioeconomic disparities in urban green space accessibility: where to intervene? Landsc. Urban Plan. 102 (4), 234–244. https://doi.org/10.1016/j.landurbplan.2011.05.002.
Doick, K.J., Hutchings, T., 2013. Air Temperature Regulation by Urban Trees and Green Infrastructure. Forestry Commission.
Dosen, A.S., Ostwald, M.J., 2013. Methodological characteristics of research testing prospect–refuge theory: a comparative analysis. Archit. Sci. Rev. 56 (3), 232–241.
Duncan, M.J., Spence, J.C., Mummery, W.K., 2005. Perceived environment and physical activity: a meta-analysis of selected environmental characteristics. Int. J. Behav. Nutr. Phys. Act. 2 (1), 11.
Ewing, R., Schmid, T., Killingsworth, R., Zlot, A., Raudenbush, S., 2008. Relationship between urban sprawl and physical activity, obesity, and morbidity. In: Urban Ecology. Springer, pp. 567–582.
Fan, Y., Das, K.V., Chen, Q., 2011. Neighborhood green, social support, physical activity, and stress: assessing the cumulative impact. Health Place 17 (6), 1202–1211.
Food and Agriculture Administration, 2016. Benefits of urban trees. Retrieved October 18, 2018 from. http://www.fao.org/resources/infographics/infographics-details/en/c/411348/.
Garden, F.L., Jalaludin, B.B., 2009. Impact of urban sprawl on overweight, obesity, and physical activity in Sydney, Australia. J. Urban Health 86 (1), 19–30.
Garvin, E.C., Cannuscio, C.C., Branas, C.C., 2013. Greening vacant lots to reduce violent crime: a randomised controlled trial. Inj. Prev. 19 (3), 198–203.
Gill, S.E., Handley, J.F., Ennos, A.R., Pauleit, S., 2007. Adapting cities for climate change: the role of the green infrastructure. Built Environ. 33 (1), 115–133.
Glanz, K., Rimer, B.K., Viswanath, K., 2008. Health Behavior and Health Education: Theory, Research, and Practice. John Wiley & Sons.
Haans, A., De Kort, Y.A., 2012. Light distribution in dynamic street lighting: two experimental studies on its effects on perceived safety, prospect, concealment, and escape. J. Environ. Psychol. 32 (4), 342–352.
Hallal, P.C., Andersen, L.B., Bull, F.C., Guthold, R., Haskell, W., Ekelund, U., 2012. Global physical activity levels: surveillance progress, pitfalls, and prospects. Lancet 380 (9838), 247–257. https://doi.org/10.1016/S0140-6736(12)60646-1.
Herzog, T.R., Kutzli, G.E., 2002. Preference and perceived danger in field/forest settings. Environ. Behav. 34 (6), 819–835.
Heynen, N., Perkins, H.A., Roy, P., 2006. The political ecology of uneven urban green space: the impact of political economy on race and ethnicity in producing environmental inequality in Milwaukee. Urban Aff. Rev. 42 (1), 3–25.
Homer, C., Dewitz, J., Yang, L., Jin, S., Danielson, P., Xian, G., ... Megown, K., 2015. Completion of the 2011 National Land Cover Database for the conterminous United States–representing a decade of land cover change information. Photogramm. Eng. Remote Sens. 81 (5), 345–354.
Huynh, Q., Craig, W., Janssen, I., Pickett, W., 2013. Exposure to public natural space as a protective factor for emotional well-being among young people in Canada. BMC Public Health 13 (1), 407.
Jansson, M., Fors, H., Lindgren, T., Wiström, B., 2013. Perceived personal safety in relation to urban woodland vegetation–a review. Urban For. Urban Green. 12 (2), 127–133.
Jensen, J.R., 2011. Introductory Digital Image Processing: A Remote Sensing Perspective.
Kardan, O., Gozdyra, P., Misic, B., Moola, F., Palmer, L.J., Paus, T., Berman, M.G., 2015. Neighborhood greenspace and health in a large urban center. Sci. Rep. 5, 11610.
Kruger, J., Ham, S.A., Kohl III, H.W., 2005. Trends in leisure-time physical inactivity by age, sex, and race/ethnicity-United States, 1994–2004. Morb. Mortal. Wkly Rep. 54 (39), 991–994.
Lauby-Secretan, B., Scoccianti, C., Loomis, D., Grosse, Y., Bianchini, F., Straif, K., 2016. Body fatness and cancer — viewpoint of the IARC Working Group. N. Engl. J. Med. 375 (8), 794–798. https://doi.org/10.1056/NEJMsr1606602.
Lee, C., Moudon, A.V., 2006. Correlates of walking for transportation or recreation purposes. J. Phys. Act. Health 3 (s1), S77–S98.
Lee, I.M., Shiroma, E.J., Lobelo, F., Puska, P., Blair, S.N., Katzmarzyk, P.T., 2012. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. Lancet 380 (9838), 219–229. https://doi.org/10.1016/S0140-6736(12)61031-9.
Li, J., Siegrist, J., 2012. Physical activity and risk of cardiovascular disease—a meta-analysis of prospective cohort studies. Int. J. Environ. Res. Public Health 9 (2), 391–407.
Loewen, L.J., Steel, G.D., Suedfeld, P., 1993. Perceived safety from crime in the urban environment. J. Environ. Psychol. 13 (4), 323–331.
Luke, D.A., 2004. Multilevel Modeling. vol. 143 Sage.
Norman, G.J., Nutter, S.K., Ryan, S., Sallis, J.F., Calfas, K.J., Patrick, K., 2006. Community design and access to recreational facilities as correlates of adolescent physical activity and body-mass index. J. Phys. Act. Health 3 (s1), S118–S128.
Prince, S.A., Adamo, K.B., Hamel, M.E., Hardt, J., Gorber, S.C., Tremblay, M., 2008. A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. Int. J. Behav. Nutr. Phys. Act. 5 (1), 56.
Sallis, J.F., 2009. Measuring physical activity environments: a brief history (Am. J. Prev. Med.). 36 (4), S86–S92.
Schukoske, J.E., 1999. Community development through gardening: state and local policies transforming urban open space. NYUJ Legis. Pub. Pol'y 3, 351.
Sigal, R.J., Armstrong, M.J., Colby, P., Kenny, G.P., Plotnikoff, R.C., Reichert, S.M., Riddell, M.C., 2013. Physical activity and diabetes. Can. J. Diabetes 37, S40–S44. https://doi.org/10.1016/j.jcjd.2013.01.018.
Tilt, J.H., Unfried, T.M., Roca, B., 2007. Using objective and subjective measures of neighborhood greenness and accessible destinations for understanding walking trips and BMI in Seattle, Washington. Am. J. Health Promot. 21 (4 suppl), 371–379.
Troy, A., Grove, J.M., O'Neil-Dunne, J., 2012. The relationship between tree canopy and crime rates across an urban–rural gradient in the greater Baltimore region. Landsc. Urban Plan. 106 (3), 262–270.
University of Manchester, 2007. Build parks to climate proof our cities. In: Science Daily, Retrieved April 4, 2018 from. www.sciencedaily.com/releases/2007/05/070514101534.htm.
US Centers for Disease Control and Prevention, 2016. 500 cities project data. Retrieved November 27, 2017 from. https://www.cdc.gov/500cities.
US Centers for Disease Control and Prevention, 2017. Exercise or physical activity. Retrieved March 30, 2018 from. https://www.cdc.gov/nchs/fastats/exercise.htm.
US Department of Agriculture, 2018. Urban tree canopy. Retrieved October 15, 2018 from. https://www.nrs.fs.fed.us/urban/utc/.
US Department of Agriculture NAIP imagery. Retrieved July 20, 2017 from. https://www.fsa.usda.gov/programs-and-services/aerial-photography/imagery-programs/naip-imagery/index (n.d.).
US Department of Justice, 2018. Uniform crime reporting statistics. Retrieved October 1, 2018 from. https://www.ucrdatatool.gov/index.cfm.
US Geological Survey, 2015. NDVI, the foundation for remote sensing phenology. Retrieved March 23, 2018 from. https://phenology.cr.usgs.gov/ndvi_foundation.php.
US National Center for Health Statistics, 2016. Health, United States, 2015: with special feature on racial and ethnic health disparities. Retrieved July 28, 2017 from. https://www.ncbi.nlm.nih.gov/books/NBK367640/.