Preventive Medicine Reports

Google Street View images as predictors of patient health outcomes, Utah 2017-2019

--Manuscript Draft--

Manuscript Number:	
Article Type:	Research paper
Keywords:	Google Street View; built environment; neighborhood characteristics; patient health; social determinants of health; computer vision
Corresponding Author:	Quynh Nguyen, PhD University of Maryland School of Public Health Riverdale, MD United States
First Author:	Quynh Nguyen, PhD
Order of Authors:	Quynh Nguyen, PhD
	Tom Belnap
	Pallavi Dwivedi
	Amir Hossein Nazem Deligani
	Abhinav Kumar
	Dapeng Li
	Ross Whitaker
	Tolga Tasdizen
	Kimberly D. Brunisholz
Abstract:	It is well-established that neighborhood characteristics can influence health outcomes. However, neighborhood data can both be time- and resource-intensive to collect, especially across broad geographies. The aim of this study was to leveraged publicly available Google Street View (GSV) images to construct indicators of neighborhood-built environment for the state of Utah. We then examine association between neighborhood features and individual-level health outcomes of about one third of people living in Utah by leveraging electronic medical records from one of the largest healthcare providers in Utah, Intermountain Healthcare. The use of electronic medical records allows for the assessment of associations between neighborhood characteristics and individual level health outcomes while controlling for predisposing factors, which distinguishes this study from previous GSV studies that were ecological in nature. We utilized close to 1.4 million GSV images to characterize neighborhood environments and implemented log Poisson models to assess their associations with 2017-2019 health outcomes among 938,085 adult patients from Intermountain Healthcare. We find that individuals living in communities in the highest tertiles of green streets and non-single family homes had 10-27% lower diabetes, uncontrolled diabetes, hypertension, and obesity but higher substance use disorders—controlling for age, white race, Hispanic ethnicity, religion, marital status, health insurance, and area deprivation index. Conversely, visible utility wires overhead was associated with 5-10% more diabetes, uncontrolled diabetes, hypertension, obesity, and substance use disorders. GSV images can aid in the identification of neighborhood characteristics that affect health.
Suggested Reviewers:	Roger Hyam r.hyam@rbge.org.uk Dr. Hyam did a study using GSV and Google's Computer Vision API to explore perceived naturalness of urban spaces.
	Xiaojiang Li xiaojiang.li@uconn.edu Dr. Li has used GSV to study urban greenery.
	Dragomir Anguelov dragomir@google.com

	Dr. Anguelov works at Google and has written a paper describes the technical challenges involved in capturing, processing, and serving street-level imagery on a global scale.
	Melvyn Hillsdon m.hillsdon@exeter.ac.uk Dr. Hillsdon has utilized GSV to assess environmental supportiveness for physical activity
	Zoe E. Petropoulos zep@bu.edu Dr. Petropoulos has utilized street imagery to assess neighborhood features.
	Michael Bader bader@american.edu Dr. Bader has worked extensively on using GSV to assess built environment features.
	Valter Silva v.silva@ymail.com Dr. Silva has implemented studies using GSV to study Obesogenic Built Environments
Opposed Reviewers:	

Cover Letter

UNIVERSITY OF MARYLAND

2234 School of Public Health Bldg
College Park, Maryland 20742-2611
301.405.3575 TEL 301.314.9366 FAX

SCHOOL OF PUBLIC HEALTH
Department of Epidemiology and Biostatistics

June 29, 2021
Eduardo L. Franco, MPH, DrPH, PhD (Hon)
McGill University, Montreal, Quebec, Canada

Dear Dr. Franco,

On behalf of my co-authors, I am pleased to submit our manuscript entitled, "Google Street View images as predictors of patient health outcomes, Utah 2017-2019" for consideration by the *Preventive Medicine Reports*.

A major barrier to investigating neighborhood influences on health is the lack of consistently constructed neighborhood indicators across large geographies that are also publicly available. Much neighborhood research has been conducted using census compositional characteristics. We know less about the potential influence of other place-based factors. In this study we construct neighborhood summaries from Google Street View images that provide eye-level views. We collected 1.4 million Google Street View images from Utah and leveraged computer vision to extract built environment features such presence of crosswalk, single lane road, green street, visible utility wires overhead, and buildings other than single family homes. Manually labeled- and algorithm-labeled images had good levels of agreement.

We leveraged clinical records of close to one million individuals aged 18 years and older in Utah in 2017-2019 to examine associations between neighborhood built environments and health outcomes. We found that individuals living in neighborhoods with more green streets and more buildings that were non-single family homes were less likely to have diabetes, uncontrolled diabetes, hypertension, and obesity but higher substance use disorders—controlling for individual age, white race, Hispanic ethnicity, any religion, marital status, and health insurance status. Conversely, visible utility wires overhead was associated with higher prevalence of all examined outcomes. Neighborhood characteristics can structure risks and resources that affect health. Utilizing existing big data sources and new technologies can enable large population studies on the potential impacts of built environments on health.

The authors have no conflict of interest to disclose. Thank you for considering our manuscript for publication. Please let me know if you have any questions.

Sincerely,

Quynh Nguyen, PhD
Associate Professor of Epidemiology and Biostatistics
University of Maryland School of Public Health

Research Highlights

- 1.4 million Google street view images were used to characterize neighborhoods in Utah
- Computer vision models labeled each image to create built environment indicators
- Electronic health records for 900,000+ patients were used to examine links with health
- Green streets and non-single family homes were connected with less chronic disease
- Visible utility wires were connected with more chronic disease and substance use

Original Research

Google Street View images as predictors of patient health outcomes, Utah 2017-2019

Manuscript word count: 3914 Abstract word count: 238 Tables: 4 Figures: 2
Appendix: eTable 1, eFigure1

ABSTRACT

It is well-established that neighborhood characteristics can influence health outcomes. However, neighborhood data can both be time- and resource-intensive to collect, especially across broad geographies. The aim of this study was to leveraged publicly available Google Street View (GSV) images to construct indicators of neighborhood-built environment for the state of Utah. We then examine association between neighborhood features and individual-level health outcomes of about one third of people living in Utah by leveraging electronic medical records from one of the largest healthcare providers in Utah, Intermountain Healthcare. The use of electronic medical records allows for the assessment of associations between neighborhood characteristics and individual level health outcomes while controlling for predisposing factors, which distinguishes this study from previous GSV studies that were ecological in nature. We utilized close to 1.4 million GSV images to characterize neighborhood environments and implemented log Poisson models to assess their associations with 2017-2019 health outcomes among 938,085 adult patients from Intermountain Healthcare. We find that individuals living in communities in the highest tertiles of green streets and non-single family homes had 10-27% lower diabetes, uncontrolled diabetes, hypertension, and obesity but higher substance use disorders—controlling for age, white race, Hispanic ethnicity, religion, marital status, health insurance, and area deprivation index. Conversely, visible utility wires overhead was associated with 5-10% more diabetes, uncontrolled diabetes, hypertension, obesity, and substance use disorders. GSV images can aid in the identification of neighborhood characteristics that affect health.

Keywords: Google Street View, built environment, neighborhood characteristics, patient health, social determinants of health, computer vision

INTRODUCTION

The importance of the built environment as a determinant of health is well-established in the literature (Macintyre and Ellaway, 2000). The quality of neighborhood conditions has been shown to influence the prevalence of obesity, diabetes, and risk of mortality (Nguyen et al., 2019a, Keralis et al., 2020). Features of the built environment may influence an individual's accessibility, and therefore likelihood, to engage in health promotion behaviors such as regular physical activity, obtaining adequate nutrition, and regularly visiting a healthcare provider, all of which may contribute to the improvement of physical and mental health (Chaiyachati et al., 2018, Morland et al., 2002, Laraia et al., 2004, Fein et al., 2004, Giles-Corti and Donovan, 2002, Penedo and Dahn, 2005). Previous research has reported the influence of neighborhood features such as presence of roadways, buildings, access to public transportation, green spaces, and walkability on both physical and mental health outcomes (Sallis et al., 2009, Burls, 2007, Nutsford et al., 2013, Browning et al., 2019a). Interconnected streets and mixed land use in urban neighborhoods have linked to increased physical activity (Frank et al., 2005). In our previous research, we found that ZIP code level built environment features such as green streets, crosswalks, and commercial buildings are associated with lower prevalence of individual level obesity and diabetes (Nguyen et al., 2018a).

Administrative data and neighborhood surveys serve as one of the sources of data on neighborhood conditions and provide insights regarding how residents perceive their neighborhood environment. While these data sources provide assessment of neighborhood features that are considered important for health by residents, these self-reported data are subject to same-source bias and social desirability bias (Chum et al., 2019, Krumpal, 2013). In-person audits are another source of data on the built-environment, but they can be expensive and time-consuming. As an alternative, Google Street View (GSV) images can serve as a reliable and cost-effective data source to capture neighborhood environments (Rundle et al., 2011c). Previous research utilizing GSV have found it to be consistent with field assessments (Rundle et al., 2011a, Kelly et al., 2013, Silva et al., 2015, Browning et al., 2019b, Rundle et al., 2011b), and can also be used to accurately identify built environment features of interest such as crosswalks, commercial buildings, highways, and grasslands (Nguyen et al., 2019b, Nguyen et al., 2018b).

Study aims

The aim of this study was to leveraged publicly available GSV images to construct indicators of neighborhood-built environment for the state of Utah. We then examine association between neighborhood features and individual-level health outcomes of about one third of people living in Utah by leveraging electronic medical records from one of the largest healthcare providers in Utah, Intermountain Healthcare. The use of electronic medical records allows for the assessment of associations between neighborhood characteristics and individual level health outcomes while controlling for predisposing factors, which distinguishes this study from previous GSV studies that were ecological in nature. Outcomes examined include obesity, diabetes, high blood pressure, and substance use disorders. Findings from this study can help inform clinical practice regarding neighborhood characteristics that are connected with patient health outcomes.

METHODS

Study Setting and Population

We conducted a cross-section study utilizing big data source of Google Street View images to create indicators of neighborhood quality and assess their associations with 2017-2019 health outcomes from patients at Intermountain Healthcare (IH). IH is a not-for-profit, integrated, community-based health system that provides services to about half of residents of Utah. Intermountain has 24 hospitals with 2,900 licensed beds and 215 owned or supported clinics. Annually IH provides 495,000 emergency department (ED) visits, 136,000 inpatient admissions and 160,000 inpatient and ambulatory surgeries. Patients included in the dataset were those who were 18 years and older, had a medical visit from 2017-2019, and were Utah residents (n= 1,433,316).

Study Measurement

Individual-level characteristics. From Intermountain Healthcare, we obtained individual-level health outcomes for eligible patients to study prevalence of type 2 diabetes, high blood pressure, and obesity (Body Mass Index $\geq$30 kg/m^2). Type 2 diabetes and hypertension were defined by the National Committee for Quality Assurance (NCQA)(National Committee for Quality Assurance (NCQA), 2013) through the Healthcare Effectiveness Data and Information Set (HEDIS) specifications. Type 2 diabetes specifications require only one of the following to be met along with a diagnosis code of diabetes (ICD-9 code: 250): (a) two outpatient encounters on different dates of service; (b) one acute inpatient encounter; (c) one emergency department visit; or (d) patients who were dispensed insulin or hypoglycemic/anti-hyperglycemics on an ambulatory basis. Individuals were identified with hypertension if they had one outpatient encounter with a hypertension diagnosis code during the study period and excluded those with evidence of end-stage renal disease, kidney transplant, pregnancy, or admission to a non-acute inpatient facility (e.g., skilled nursing facility). Other outcomes included type 2 diabetes control (HbA1c $\geq$ 7%) and substance use disorders (includes any of the following: alcohol, opioid, cannabis, sedative, hypnotics, anxiolytics, cocaine, other stimulates including caffeine, hallucinogens, inhalants, other psychoactive substances and multiple drug use). Sociodemographic characteristics included age (continuous), race (white yes/no), ethnicity (Hispanic yes/no), marital status (married vs. not married), religious affiliation (any versus no), insurance (yes/no) and area deprivation index (ADI). The ADI is a geographic area-based measure of the disadvantaged position of residents relative to the society (Phillips et al., 2016). The ADI was calculated for the state of Utah using a measure developed by Singh et al. (Singh, 2003) based upon 17 US Census measures associated with mortality, including living conditions, income, unemployment and education. Census measures were based upon the 2013 American Community Survey published by the US Census Bureau.

Google Street view image data

Google Street View image data collection. Using GSV Image API, we collected GSV image data for street intersections and for 50-meter apart sampled locations along the road segments for all the primary and secondary roads in Utah. For each location, we obtained four GSV images (west, east, north and south) to capture neighborhood environment. In total, 1,394,442 million images from Utah were obtained in November 2019.

Built environment indicators. Indicators for building type (the presence of any building that was not a single-family detached house), single-lane road (yes/no), presence of a crosswalk (yes/no), street greenness (street trees and street landscaping comprised at least 30% of the image - yes/no), and visible utility wires overhead (yes/no) were selected through an iterative process of considering what the literature has found to be important built environment characteristics and what is feasible for computer vision models. Neighborhood walkability (Rundle et al., 2008, Van Cauwenberg et al., 2016, Li et al., 2009), neighborhood disorder (Ross and Mirowsky, 2001, Molnar et al., 2004, Burdette and Hill, 2008), and mixed land use (Rundle et al., 2007, Renalds et al., 2010, Stevenson et al., 2016) have been identified in the literature as being important for health outcomes. The presence of crosswalks is a traditional indicator of walkability and was included to measure its influence on health behaviors and related health outcomes. We also considered sidewalks, but in urban areas, the prevalence of sidewalks is high and thus there is less variability with this indicator.

Moreover, we aimed to construct a measure of mixed land use because it's impact on travel behavior is well-studied. Areas that are single-use residential often lead individuals to use motorized transport to get to destinations. Conversely, areas that blend a mixture of residential, commercial and leisure destinations might allow individuals to walk or bike (Manaugh and Kreider, 2013) and be related to greater access to resources, physical activity and better health. We operationalized mixed land use such that labeling images was feasible for both humans (human coders manually labeled images to provide training data to the computer vision models) and machines. While looking through images, we noticed that an image could be classified as having only homes or a blend of homes and other building types. Thus, non-single-family home was created to distinguish between purely residential places with detached homes and places with different building types including businesses, schools, apartments, and cultural venues. The prevalence of this indicator with a median value of around 30% nationally suited the capacity of computer vision models, which can struggle if the feature is too common or rare. Single lane roads were selected to serve as an indicator of lower urban development to distinguish between areas with more capacity for cars and people versus areas with less capacity.

We operationalized street greenness as street trees and street landscaping comprising at least 30% of the image. A cut-point of approximately 30% was utilized to assist with inter-rater reliability in manual annotations of street greenness. Moreover, we found that most images had some street landscaping and aimed to create a neighborhood indicator to distinguish between ample versus sparse street landscaping.

From images, we also extracted the presence of visible wires. The literature on visible wires is nascent and more of this work has been done abroad, for instance in Rio de Janeiro, where the wires represent both an unattractive presence and a possible electrocution/electrical fire risk (Remigio et al., 2019). In the United States, visible wires have mainly a visual impact on the landscape. We chose this indicator to further the literature and to investigate whether visible wires as an indicator of physical disorder might have links to important health outcomes. Visible utility wires hanging overhead are visually striking and may impact residents' aesthetic sense of their environment, altering perceptions of safety or pleasurability and influencing both mental health (by increasing stress levels) and physical health (by discouraging walking). Other neighborhood indicators of physical disorder were considered, such as litter or trash. However, we found that computer vision models struggled with small objects. In addition, these objects were also difficult to label by humans as well (low inter-rater reliability due to litter appearing like leaves). Thus, while litter is a classic built environment feature for neighborhood disorder, we could not include this indicator.

Image data processing. Images were processed using trained Visual Geometry Group (VGG-19 model) deep convolutional networks (Krizhevsky et al., 2012, Simonyan and Zisserman, 2014) previously detailed by Nguyen et al. (Nguyen et al., 2018b)) to identify the built environment features of interest (one network per feature). Accuracy of the recognition tasks (comparing the images labeled using this machine learning approach compared with assessment by a human reviewer) ranged from 82-97%, and these figures were consistent with a separate, semi-supervised learning approach.

Neighborhood definitions. Census tracts were chosen as the neighborhood unit given that they represent relatively homogenous units with respect to population characteristics, economic status, and living conditions (Census Bureau, 2020). Census tracts generally range in population size between 1,200 and 8,000 people with an optimum size of 4,000. To arrive at the neighborhood indicators, we processed street imagery and then combined information on all street imagery within a census tract to arrive at census tract-level summaries (e.g., % of images in a census tract that contain a crosswalk). We derived aggregated measures for green streets, crosswalks, non-single family homes, single lane roads, and visible wires and created tertiles for all the built-environment indicators based on these measures. Tertiles were utilized to allow for nonlinearities in the relationship between built environment characteristics and health outcomes.

Statistical Analyses

The data on neighborhood features were merged with the individual-level health outcomes and sociodemographic data for patients. We implemented log Poisson regression models to examine the association between tertiles of built-environment indicators and individual chronic disease prevalence after adjusting for individual-level sociodemographic characteristics. Outcomes examined included diabetes prevalence, uncontrolled diabetes, high blood pressure, obesity, and substance use disorder. A variety of health outcomes were chosen to determine the range with which GSV images can predict patient health outcomes. Main predictors included tertiles for greenspace, crosswalk, non-single family homes, single lane road, and visible utility wires. Health outcomes were compared for patients living in neighborhoods in the third tertile (and second tertile) of built environment characteristics vs. the first tertile (lowest level). Models also adjusted for age, race, ethnicity, religious affiliation, health insurance status, and area-level deprivation index. Separate models were run for each health outcome, but all models included all the environment predictors and covariates simultaneously. Statistical significance was assessed with an alpha level of 0.05. SAS 9.4 software was utilized for analyses (SAS Institute Inc., Cary, NC, USA).

Given that potential impacts of neighborhood characteristics could differ in urban and rural areas, analyses were stratified by urban status. Census tracts that were over 50% rural were defined as rural. Nonetheless, the majority of patients served by Intermountain lived in urban areas (n=938,085 with nonmissing data on covariates and health outcomes), and hence those analyses are presented in the main tables.

RESULTS

Table 1 summarizes descriptive statistics of our study population and their census tract neighborhood environment derived from GSV images. The mean age was 47 years old with about 57% being female, 58% being married, 11% being Hispanic/Latinx, and 5% being non-white. About 28% were self-pay (uninsured) and 68% reported a religious affiliation. The prevalence of obesity was 47% and the prevalence of diabetes was 6%. Figure 1 displays the distribution of the GSV-derived built environment characteristics. Single lane roads and visible utility wires were unimodal and relatively common characteristics. Street greenness was right skewed with most census tracts have prevalences of 60% and above. Non-single family homes was left skewed, with the majority of census tracts having prevalences less than 40%. Crosswalks was also left skewed and also the rarest of the built environment characteristics with most census tracts having prevalences less than 10%.

Figure 2 presents the spatial distribution the GSV-derived built environment features across the Wasatch Front, which contains the major cities of Salt Lake City, West Valley City, Provo, West Jordan, Layton, and Ogden where the majority of Utah residents live. Single lane roads were concentrated in the areas such as the eastern part of Salt Lake City, Bountiful, West Valley City, Millcreek, Sandy, and Draper City (Utah County). Street greenness was concentrated to the east. Crosswalks were present only in a few locations (e.g., Salt Lake City, South Salt Lake, Murray, Ogden, and Provo) in the urban core. Visible utility and non-single family homes were present in the urban core (e.g., Salt Lake City and South Salt Lake) and also dispersed to the west.

Table 2 presents the estimated prevalence ratios (and 95% CI) for all the examined associations between tertiles of built environment indicators and individual health outcomes, controlling for individual age, white race, Hispanic ethnicity, any religion, marital status, lack of health insurance. In all models, GSV derived built environment variables statistically significantly predicted health outcomes, with green space and non-single family homes being protective. Comparing the third tertile with the first tertile, non-single family homes were associated with lower prevalence of diabetes (PR: 0.83; 95% CI: 0.81-0.85), uncontrolled diabetes (PR: 0.86; 95% CI:0.82-0.89), hypertension (PR: 0.73; 95% CI: 0.67-0.80), and obesity (PR: 0.89; 95% CI: 0.88-0.90). Green streets were associated with decreased diabetes (PR: 0.90), uncontrolled diabetes (PR: 0.89), hypertension (PR: 0.84), and obesity (PR: 0.90). However, both green streets and non-single family homes were tied to increased prevalence of substance use disorders, 17% and 12% respectively.

An increase in visible wires was associated with higher prevalence of all adverse outcomes, although not all comparisons for the 3rd and 2nd tertile reached statistical significance. More visible wires were associated with 9-10% higher prevalence of diabetes and uncontrolled diabetes and a 4-5% increase in obesity. Visible utility wires also were linked to increased hypertension and substance use. Surprisingly, more crosswalks were associated with 7-9% increased prevalence of hypertension and only weakly associated with other health outcomes. Single lane roads were generally not associated with health outcomes, except for a slight increase in diabetes (Table 2). Patterns are similar in rural areas but associations were more attenuated and the statistical power was less given the fewer number of Intermountain patients living in rural areas (n=53,414; eTable 1).

Individual characteristics were also associated with health outcomes and all tended to be statistically significant except for English as a primary language which had little effect and was removed from the final model. White race was associated with better health outcomes including lower prevalence of diabetes, uncontrolled diabetes, hypertension, and obesity (Table 2). Hispanic ethnicity was associated with increased diabetes, uncontrolled diabetes, and obesity. Religious affiliation was associated with more diabetes, more uncontrolled diabetes, and obesity, but was protective of hypertension. Marital status (married) was associated with hypertension.

To examine whether individual-level disadvantage was associated with certain built environments, we implemented log Poisson models to examine predictors of uninsured status among Intermountain patients. Uninsured patients were less likely to live in neighborhoods with green streets and to live in non-single-family-home neighborhoods. They were more likely to live in neighborhoods with visible utility wires overhead and were slightly more likely to live in neighborhoods with single lane roads and crosswalks (Table 3).

We examined associations between GSV derived built environment indicators and other census tract level characteristics. Percent of non-Hispanics blacks was related to less exposure to green space and single lane roads and more exposure to visible utility wires and non-single family homes. Median household income was related to more green space and less visible utility wires and non-single family homes (Table 4).

DISCUSSION

While a large body of literature has connected neighborhood characteristics with an array of health outcomes, neighborhood data beyond sociodemographic characteristics can be time consuming and expensive to gather, and thus largely unavailable for large areas of the country. In this study, we leverage high resolution images from GSV across the state of Utah to construct indicators of the built environment. We then examine whether these built environment characteristics were associated with patient health outcomes. Working with Intermountain Healthcare, a major provider of care in Utah, we examined health patterns for close to 1 million patients. Our study found that non-single-family homes (indicator of mixed land use and urban development) and green streets were related to lower chronic conditions. Conversely, visible utility wires and single lane roads were connected with higher burden of chronic conditions. This aligns with previous studies conducted at the census tract, county, and state level that have found similar associations for non-single family home, single lane roads, and visible utility wires (Phan et al., 2020, Keralis et al., 2020). For example, a previous state-level GSV study have linked non-single family homes and decreased diabetes and premature mortality and increased physical activity (Phan et al., 2020). Additionally, previous county-level analyses found urban development to generally rate to lower chronic conditions and premature mortality (Nguyen et al., 2019c). However, those studies were ecological and the current study is one of the few utilizing individual-level data.

In this study with individual-level patient data, we found that crosswalks (a potential indicator of walkability) were related to worse health outcomes, which is counter to our study hypotheses. Previous involving the 500 Cities Project, found complex results with crosswalks (Keralis et al., 2020). Areas with the most crosswalks (third tertile) experienced a reduction in obesity, diabetes and physical inactivity but the second tertile experienced higher rates of obesity, diabetes and physical inactivity compared to the first (lowest) tertile. While an increase in crosswalks is likely to facilitate walking and physical activity, an increase in area-level crime would deter walking. Thus, these complex relationships between crosswalks and health outcomes might be influenced by factors, such as neighborhood crime, which were not considered in this study. The distribution of crosswalks, was more left skewed and rarer than any other variable (Figure 1). Crosswalks might also be more likely placed in core urban centers where the most disadvantage individuals might live (Figure 2). Also, individuals without health insurance were slightly more likely to live in areas with more crosswalks (Table 2).

We additionally find that green streets and non-single family homes was related to higher prevalence of substance use disorders. Street landscaping and presence of other building types besides single detached family homes might indicate higher urbancity. The landscape of Utah with its sandy deserts, red rocks, and deep canyons has generally less natural greenness, which might mean that areas with more green landscaping denote higher urban development. In previous GSV analyses, we found that higher urban development was related to more excessive drinking (Nguyen et al., 2019c).

This study also examined predictors of built environment by health insurance status. Uninsured patients were more likely to live in areas with visible utility wires, single lane roads and crosswalks. Uninsured patients were less likely to live in areas with green street and non-single homes. In one of our previous studies, we found that greater county-level economic disadvantage was associated with lower prevalence of non-single-family homes and visible wires at the county level after adjusting for violent crime rate, age, race/ethnicity, percentage of population not proficient in English and ratio of population to primary care providers (Nguyen et al., 2020).

Study strengths and limitations. This is among the few studies examining GSV derived predictors of individual level outcomes, controlling for individual level predisposing characteristics. Previous studies with GSV images have utilized ecological frameworks; for instance, county level built environment predictors of county health outcomes. In partnership with one of the largest healthcare providers in Utah, in this study, we included close to a third of people in Utah. We find that GSV-derived built environment characteristics were linked with an array of important health outcomes. Study findings could suggest structuring neighborhoods to locate amenities where people live and adding street landscaping could improve chronic disease health. Conversely, physical disorder could increase health risks, through potential mechanisms such as decreased perception of safety and social cohesion, decreased physical activity, and poorer mental health status (Burdette and Hill, 2008, Casciano and Massey, 2012, Bjornstrom et al., 2013, Molnar et al., 2004).

Nonetheless, our study is subject to limitations. While utilized data from one of the main healthcare providers in Utah, there may differences between the composition of patients at Intermountain and Utah as a whole. For example, females are slightly over-represented, comprising 54.4% of the Intermountain sample versus 49.6% of the Utah population according to census estimates (U.S. Census Bureau, 2019). Additionally, a higher proportion of Intermountain patients are white versus those in Utah (95.4% vs. 90.6) (U.S. Census Bureau, 2019).

We leveraged Google Street View images and computer vision to characterize neighborhood environments. Nonetheless, other neighborhood characteristics not examined could also be important for health, including noise level, air pollution levels, and perceived safety of neighborhoods. For instance, previous studies have reported a positive association between perceived neighborhood safety and higher levels of walking activity (Li et al., 2005). Furthermore, using computer vision limits the types and number of features that can be examined. Computer vision is more accurate for features that are relatively large in size which makes detection easier. For instance, computer vision can more easily detect cars compared to litter. Moreover, the number of features that could be examined was limited given that training datasets were compiled from manual annotations of each neighborhood feature and separate computer vision models were built for each feature. Thus, unlike onsite neighborhood inventories that can consist of potentially hundreds of neighborhood features, we focus on a few core neighborhood features that have theoretically and empirically connected to health outcomes. These contextual characteristics can better help healthcare organizations understand the drivers' of their patients' health by further considering patients' residential environments which present both risks and resources. Patients living in communities with multiple factors of disadvantage may necessitate additional services and support to counter those disadvantages.

TABLES

Table 1. Descriptive statistics of study population, Utah, 2019

	N	Mean (Standard Deviation)
Individual level covariates		
Age (years)	1,433,316	46.53 (19.03)
% Female	1,433,316	54.36% (54.28-54.45)
% Married	1,069,207	58.06% (57.98-58.14)
% Nonwhite	1,346,584	4.61% (4.58-4.65)
% White	1,346,584	95.39% (95.35-95.42)
% Hispanic ethnicity	1,357,627	10.83% (10.78-10.88)
% Uninsured	1,433,316	28.39% (28.31-28.46)
% Religious affiliation	1,069,207	68.17% (68.08 – 68.25)
Area deprivation index	1,433,298	97.51 (18.61)
Health outcomes		
% Obese	1,374,731	47.28% (47.19-47.36)
% Diabetes	1,433,316	5.88% (5.84-5.92)
Hemoglobin A1c (%)	1,433,316	9.23% (9.18-9.28)
% Hypertension	1,433,316	0.69% (0.68-0.71)
Google Street View (Census tract)		
Green street	164,443,190	83.76 (12.68)
Crosswalk	164,443,190	4.95 (3.82)
Non-single family home[a]	164,443,190	27.53 (17.24)
Single lane road	164,443,190	65.56 (11.65)
Visible utility wires	164,443,190	46.19 (14.36)

[a] Non-single family home = presence of a building that is not a single family home (e.g., schools, grocery stores and other businesses denoting mixed land use)

Table 2. Associations between built environment characteristics and individual level health outcomes

	Diabetes	Uncontrolled Diabetes	Hypertension	Obesity	Substance use disorder
	Prevalence Ratio (95% CI)[b]	Prevalence Ratio (95% CI)[b]	Prevalence Ratio (95% CI)[b]	Prevalence Ratio (95% CI)[b]	Prevalence Ratio (95% CI)[b]
GSV indicators					
Green streets, 3rd tertile	0.90 (0.88, 0.92)*	0.89 (0.86, 0.92)*	0.84 (0.78, 0.90)*	0.90 (0.89, 0.91)*	1.17 (1.13, 1.21)*
Green streets, 2nd tertile	0.99 (0.97, 1.01)	0.98 (0.95, 1.01)	0.98 (0.93, 1.05)	0.98 (0.97, 0.98)*	1.06 (1.03, 1.09)*
Crosswalks, 3rd tertile	1.02 (1.00, 1.05)*	1.01 (0.98, 1.04)	1.07 (1.00, 1.14)*	1.01 (1.00, 1.02)*	1.00 (0.97, 1.03)
Crosswalks, 2nd tertile	1.01 (0.99, 1.03)	1.00 (0.98, 1.03)	1.09 (1.02, 1.16)*	1.02 (1.01, 1.02)*	0.99 (0.96, 1.02)
Non-single family home, 3rd tertile	0.83 (0.81, 0.85)*	0.86 (0.82, 0.89)*	0.73 (0.67, 0.80)*	0.89 (0.88, 0.90)*	1.12 (1.08, 1.17)*
Non-single family home, 2nd tertile	0.91 (0.89, 0.93)*	0.91 (0.88, 0.94)*	0.89 (0.83, 0.96)*	0.95 (0.95, 0.96)*	1.03 (0.99, 1.06)
Single lane roads, 3rd tertile	1.02 (0.99, 1.04)	1.00 (0.97, 1.04)	0.94 (0.87, 1.01)	1.00 (0.99, 1.01)	0.98 (0.95, 1.02)
Single lane roads, 2nd tertile	1.03 (1.01, 1.05)*	1.01 (0.99, 1.04)	0.98 (0.92, 1.04)	1.00 (1.00, 1.01)	0.97 (0.94, 1.00)
Visible wires, 3rd tertile	1.09 (1.06, 1.11)*	1.10 (1.06, 1.14)*	1.05 (0.97, 1.14)	1.04 (1.03, 1.06)*	1.05 (1.01, 1.09)*
Visible wires, 2nd tertile	1.09 (1.07, 1.12)*	1.10 (1.07, 1.13)*	1.08 (1.01, 1.16)*	1.05 (1.04, 1.05)*	0.99 (0.96, 1.02)
Covariates					
Age (years)	1.04 (1.04, 1.04)*	1.03 (1.03, 1.03)*	1.01 (1.01, 1.01)*	1.01 (1.01, 1.01)*	1.00 (1.00, 1.00)
White race	0.60 (0.58, 0.62)*	0.53 (0.51, 0.55)*	0.80 (0.72, 0.90)*	0.93 (0.91, 0.94)*	1.16 (1.10, 1.22)*
Hispanic ethnicity	1.15 (1.12, 1.18)*	1.34 (1.30, 1.39)*	0.96 (0.88, 1.05)	1.08 (1.07, 1.09)*	0.68 (0.65, 0.70)*
Any religion	1.21 (1.19, 1.23)*	1.18 (1.15, 1.21)*	0.86 (0.82, 0.91)*	1.07 (1.06, 1.07)*	0.65 (0.64, 0.67)*
Married	1.09 (1.07, 1.11)*	1.03 (1.01, 1.05)*	1.40 (1.33, 1.48)*	1.12 (1.11, 1.13)*	0.40 (0.39, 0.41)*
Uninsured	1.60 (1.57, 1.63)*	1.73 (1.69, 1.77)*	1.11 (1.05, 1.17)*	1.10 (1.09, 1.11)*	2.38 (2.33, 2.44)*
Area deprivation index	1.01 (1.01, 1.01)*	1.01 (1.01, 1.01)*	1.00 (1.00, 1.00)*	1.01 (1.01, 1.01)*	1.01 (1.01, 1.01)*

[a] For GSV indicators, reference category is 1[st] tertile. [b]Adjusted Log Poisson regression controlled for the following covariates: age, white race, Hispanic ethnicity, any religion, marital status, self-pay status for health insurance, area deprivation index. n = 938,085 *$p<0.05$

Table 3. Predicting uninsured status with neighborhood- and individual-level characteristics

	Prevalence Ratio (95% CI)
GSV indicators	
Green streets, 3rd tertile	0.89 (0.87, 0.92)*
Green streets, 2nd tertile	1.01 (0.99, 1.03)
Crosswalks, 3rd tertile	1.08 (1.05, 1.10)*
Crosswalks, 2nd tertile	1.06 (1.04, 1.08)*
Non-single family home, 3rd tertile	0.85 (0.83, 0.87)*
Non-single family home, 2nd tertile	0.88 (0.86, 0.90)*
Single lane roads, 3rd tertile	1.06 (1.03, 1.08)*
Single lane roads, 2nd tertile	1.04 (1.01, 1.06)*
Visible wires, 3rd tertile	1.32 (1.29, 1.35)*
Visible wires, 2nd tertile	1.23 (1.20, 1.25)*
Covariates	
Age (years)	1.04 (1.04, 1.04)*
White race	0.57 (0.55, 0.59)*
Hispanic ethnicity	1.33 (1.29, 1.36)*
Any religion	1.23 (1.21, 1.25)*
Married	1.03 (1.01, 1.05)*

Adjusted Poisson regression controlled for all variables listed simultaneously, n = 938,085 *$p<0.05$. For Google Street View indicators, the reference category is the 1st tertile.

Table 4. Associations between census tract sociodemographics and Google Street View-derived built environment characteristics, census tract level

	Built environment indicators				
Census tract characteristics[a]	Green space	Crosswalk	Not single family home	Single lane roads	Visible wire
	Prevalence (95% CI)	Prevalence (95% CI)	Prevalence (95% CI)	Prevalence (95% CI)	Prevalence (95% CI)
% non-Hispanic black	-43.68 (-60.61, -26.74)*	13.84 (9.08, 18.61)*	70.67 (48.88, 92.45)*	-67.12 (-84.09, -50.16)*	51.00 (32.75, 69.24)*
% Hispanic	0.16 (-2.00, 2.32)	-0.38 (-0.99, 0.23)	-3.50 (-6.28, -0.72)*	4.01 (1.85, 6.18)*	2.54 (0.21, 4.86)*
% Unemployed	1.72 (0.07, 3.36)*	0.34 (-0.13, 0.80)	0.83 (-1.29, 2.95)	-0.57 (-2.22, 1.08)	-0.26 (-2.04, 1.52)
Median household income	7.46 (5.75, 9.17)*	-0.70 (-1.18, -0.22)*	-11.59 (-13.79, -9.39)*	5.68 (3.97, 7.40)*	-10.55 (-12.39, -8.70)*
Household size	-2.96 (-3.89, -2.04)*	-0.76 (-1.02, -0.50)*	-2.56 (-3.75, -1.36)*	-0.33 (-1.26, 0.60)	-0.09 (-1.09, 0.91)
Population density	5.90 (5.00, 6.80)*	1.57 (1.32, 1.83)*	-5.65 (-6.81, -4.50)*	0.95 (0.05, 1.85)*	-2.69 (-3.66, -1.73)*

[a]All predictor variables are standardized to have a mean of 0 and standard deviation of 1

*p<0.05; N= 586 census tracts in Utah

Figure 1. Distribution of built environment characteristics in Utah

Histogram of Google Street View-derived built environment characteristics (presence of crosswalk, single lane road, green street, visible utility wires overhead, and buildings other than single family homes)

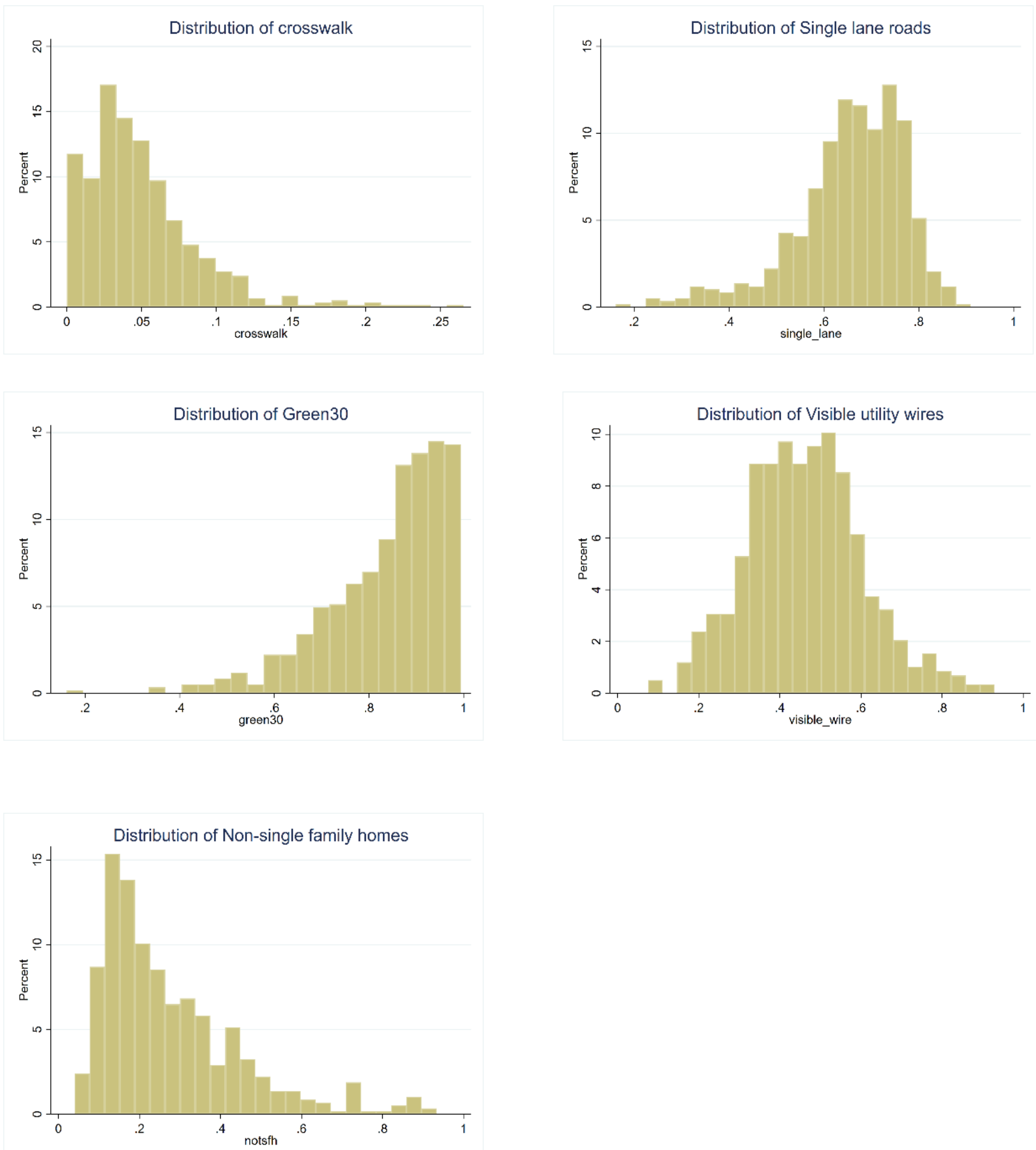
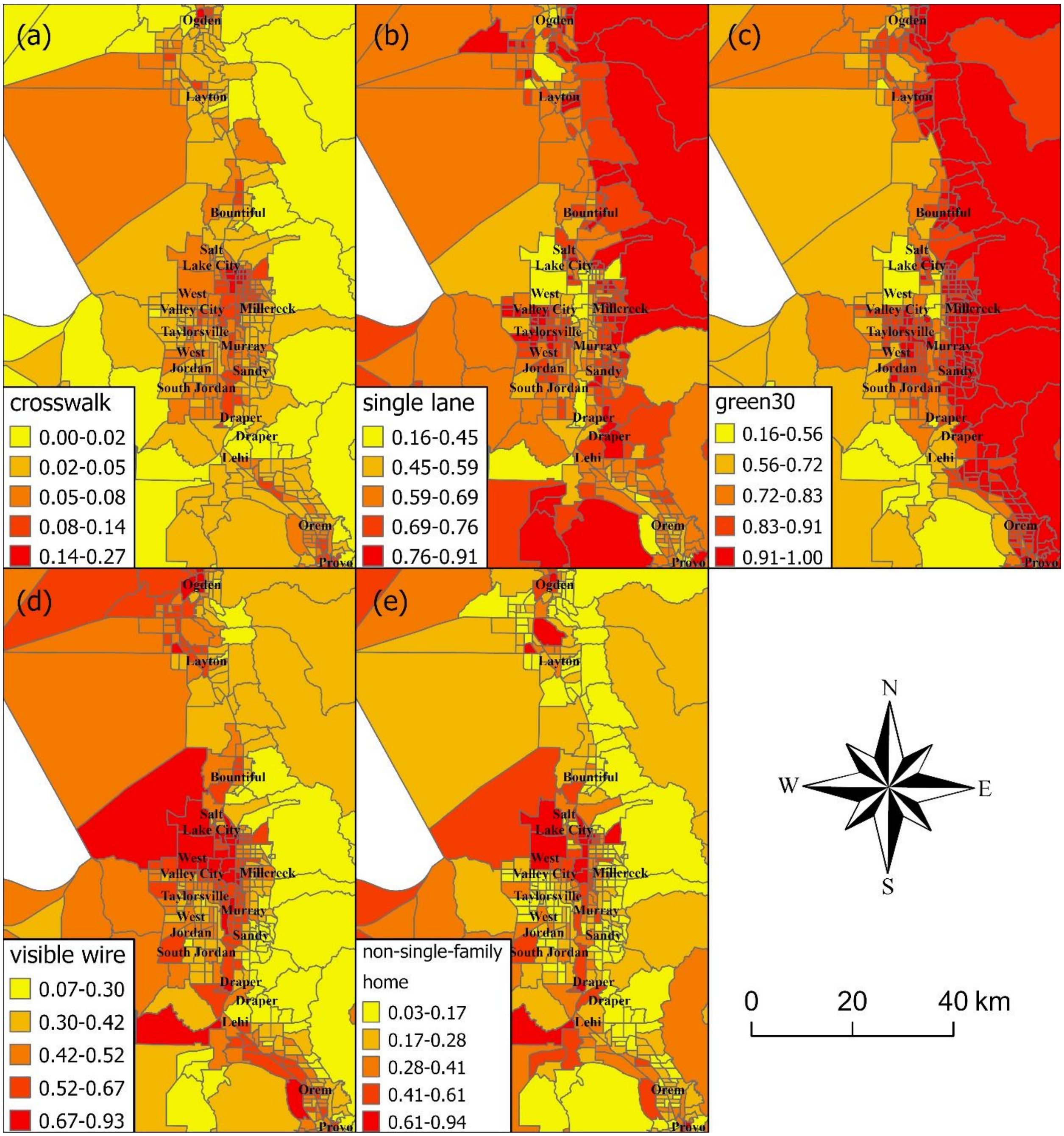


Figure 2. Geographical distribution of built environment characteristics in Utah
Spatial distribution of Google Street View derived built environment characteristics across zip codes in the Wasatch Front which contains the major cities of Salt Lake City, West Valley City, Provo, West Jordan, Layton, and Ogden where the majority of Utah residents live. Built environment features mapped include: presence of crosswalk, single lane road, green street, visible utility wires overhead, and buildings other than single family homes.



References

BJORNSTROM, E. E. S., RALSTON, M. L. & KUHL, D. C. 2013. Social Cohesion and Self-Rated Health: The Moderating Effect of Neighborhood Physical Disorder. *American Journal of Community Psychology,* 52**,** 302-312.

BROWNING, M. H., LEE, K. & WOLF, K. L. 2019a. Tree cover shows an inverse relationship with depressive symptoms in elderly residents living in US nursing homes. *Urban Forestry & Urban Greening,* 41**,** 23-32.

BROWNING, M. H. E. M., LEE, K. & WOLF, K. L. 2019b. Tree cover shows an inverse relationship with depressive symptoms in elderly residents living in U.S. nursing homes. *Urban Forestry & Urban Greening,* 41**,** 23-32.

BURDETTE, A. M. & HILL, T. D. 2008. An examination of processes linking perceived neighborhood disorder and obesity. *Social science & medicine,* 67**,** 38-46.

BURLS, A. 2007. People and green spaces: promoting public health and mental well-being through ecotherapy. *Journal of public mental health,* 6**,** 24.

CASCIANO, R. & MASSEY, D. S. 2012. Neighborhood disorder and anxiety symptoms: New evidence from a quasi-experimental study. *Health & Place,* 18**,** 180-190.

CENSUS BUREAU. 2020. *Census Tracts and Block Numbering Areas* [Online]. Available: https://www2.census.gov/geo/pdfs/reference/GARM/Ch10GARM.pdf [Accessed November 19, 2020].

CHAIYACHATI, K. H., HOM, J. K., HUBBARD, R. A., WONG, C. & GRANDE, D. 2018. Evaluating the association between the built environment and primary care access for new Medicaid enrollees in an urban environment using walk and transit scores. *Preventive medicine reports,* 9**,** 24-28.

CHUM, A., O'CAMPO, P., LACHAUD, J., FINK, N., KIRST, M. & NISENBAUM, R. 2019. Evaluating same-source bias in the association between neighbourhood characteristics and depression in a community sample from Toronto, Canada. *Social psychiatry and psychiatric epidemiology,* 54**,** 1177-1187.

FEIN, A. J., PLOTNIKOFF, R. C., WILD, T. C. & SPENCE, J. C. 2004. Perceived environment and physical activity in youth. *International journal of behavioral medicine,* 11**,** 135-142.

FRANK, L. D., SCHMID, T. L., SALLIS, J. F., CHAPMAN, J. & SAELENS, B. E. 2005. Linking objectively measured physical activity with objectively measured urban form: findings from SMARTRAQ. *Am J Prev Med,* 28**,** 117-25.

GILES-CORTI, B. & DONOVAN, R. J. 2002. The relative influence of individual, social and physical environment determinants of physical activity. *Social science & medicine,* 54**,** 1793-1812.

KELLY, C. M., WILSON, J. S., BAKER, E. A., MILLER, D. K. & SCHOOTMAN, M. 2013. Using Google Street View to audit the built environment: inter-rater reliability results. *Ann Behav Med,* 45 Suppl 1**,** S108-12.

KERALIS, J. M., JAVANMARDI, M., KHANNA, S., DWIVEDI, P., HUANG, D., TASDIZEN, T. & NGUYEN, Q. C. 2020. Health and the built environment in United States cities: measuring associations using Google Street View-derived indicators of the built environment. *BMC public health,* 20**,** 215.

KRIZHEVSKY, A., SUTSKEVER, I. & HINTON, G. E. 2012. ImageNet classification with deep convolutional neural networks. *Proceedings of the 25th International Conference on Neural Information Processing Systems* - *Volume 1.* Lake Tahoe, Nevada: Curran Associates Inc.

KRUMPAL, I. 2013. Determinants of social desirability bias in sensitive surveys: a literature review. *Quality & Quantity,* 47**,** 2025-2047.

LARAIA, B. A., SIEGA-RIZ, A. M., KAUFMAN, J. S. & JONES, S. J. 2004. Proximity of supermarkets is positively associated with diet quality index for pregnancy. *Preventive Medicine,* 39**,** 869-875.

LI, F., FISHER, K. J., BROWNSON, R. C. & BOSWORTH, M. 2005. Multilevel modelling of built environment characteristics related to neighbourhood walking activity in older adults. *Journal of epidemiology and community health,* 59**,** 558-564.

LI, F., HARMER, P., CARDINAL, B. J., BOSWORTH, M., JOHNSON-SHELTON, D., MOORE, J. M., ACOCK, A. & VONGJATURAPAT, N. 2009. Built environment and 1-year change in weight and waist circumference in middle-aged and older adults: Portland Neighborhood Environment and Health Study. *American journal of epidemiology,* 169**,** 401-408.

MACINTYRE, S. & ELLAWAY, A. 2000. Ecological approaches: rediscovering the role of the physical and social environment. *Social epidemiology,* 9**,** 332-348.

MANAUGH, K. & KREIDER, T. 2013. What is mixed use? Presenting an interaction method for measuring land use mix. *Journal of Transport and Land Use,* 6**,** 63-72.

MOLNAR, B. E., GORTMAKER, S. L., BULL, F. C. & BUKA, S. L. 2004. Unsafe to play? Neighborhood disorder and lack of safety predict reduced physical activity among urban children and adolescents. *American journal of health promotion,* 18**,** 378-386.

MORLAND, K., WING, S. & DIEZ ROUX, A. 2002. The Contextual Effect of the Local Food Environment on Residents' Diets: The Atherosclerosis Risk in Communities Study. *Am J Public Health,* 92**,** 1761-1768.

NATIONAL COMMITTEE FOR QUALITY ASSURANCE (NCQA) 2013. HEDIS 2014: Healthcare Effectiveness Data and Information Set.

NGUYEN, Q. C., KERALIS, J. M., DWIVEDI, P., NG, A. E., JAVANMARDI, M., KHANNA, S., HUANG, Y., BRUNISHOLZ, K. D., KUMAR, A. & TASDIZEN, T. 2020. Leveraging 31 Million Google Street View Images to Characterize Built Environments and Examine County Health Outcomes. *Public Health Reports***,** 0033354920968799.

NGUYEN, Q. C., KHANNA, S., DWIVEDI, P., HUANG, D., HUANG, Y., TASDIZEN, T., BRUNISHOLZ, K. D., LI, F., GORMAN, W. & NGUYEN, T. T. 2019a. Using Google Street View to examine associations between built environment characteristics and US health outcomes. *Preventive medicine reports,* 14**,** 100859.

NGUYEN, Q. C., KHANNA, S., DWIVEDI, P., HUANG, D., HUANG, Y., TASDIZEN, T., BRUNISHOLZ, K. D., LI, F., GORMAN, W., NGUYEN, T. T. & JIANG, C. 2019b. Using

Google Street View to examine associations between built environment characteristics and U.S. health outcomes. *Prev Med Rep,* 14**,** 100859.

NGUYEN, Q. C., KHANNA, S., DWIVEDI, P., HUANG, D., HUANG, Y., TASDIZEN, T., BRUNISHOLZ, K. D., LI, F., GORMAN, W., NGUYEN, T. T. & JIANG, C. 2019c. Using Google Street View to examine associations between built environment characteristics and U.S. health outcomes. *Preventive Medicine Reports,* 14**,** 100859.

NGUYEN, Q. C., SAJJADI, M., MCCULLOUGH, M., PHAM, M., NGUYEN, T. T., YU, W., MENG, H.-W., WEN, M., LI, F. & SMITH, K. R. 2018a. Neighbourhood looking glass: 360º automated characterisation of the built environment for neighbourhood effects research. *J Epidemiol Community Health,* 72**,** 260-266.

NGUYEN, Q. C., SAJJADI, M., MCCULLOUGH, M., PHAM, M., NGUYEN, T. T., YU, W., MENG, H. W., WEN, M., LI, F., SMITH, K. R., BRUNISHOLZ, K. & TASDIZEN, T. 2018b. Neighbourhood looking glass: 360-degree automated characterisation of the built environment for neighbourhood effects research. *J Epidemiol Community Health,* 72**,** 260-266.

NUTSFORD, D., PEARSON, A. & KINGHAM, S. 2013. An ecological study investigating the association between access to urban green space and mental health. *Public health,* 127**,** 1005-1011.

PENEDO, F. J. & DAHN, J. R. 2005. Exercise and well-being: a review of mental and physical health benefits associated with physical activity. *Current opinion in psychiatry,* 18**,** 189-193.

PHAN, L., YU, W., KERALIS, J. M., MUKHIJA, K., DWIVEDI, P., BRUNISHOLZ, K. D., JAVANMARDI, M., TASDIZEN, T. & NGUYEN, Q. C. 2020. Google Street View Derived Built Environment Indicators and Associations with State-Level Obesity, Physical Activity, and Chronic Disease Mortality in the United States. *International Journal of Environmental Research and Public Health,* 17**,** 3659.

PHILLIPS, R. L., LIAW, W., CRAMPTON, P., EXETER, D. J., BAZEMORE, A., VICKERY, K. D., PETTERSON, S. & CARROZZA, M. 2016. How other countries use deprivation indices—and why the United States desperately needs one. *Health Affairs,* 35**,** 1991-1998.

REMIGIO, R. V., ZULAIKA, G., RABELLO, R. S., BRYAN, J., SHEEHAN, D. M., GALEA, S., CARVALHO, M. S., RUNDLE, A. & LOVASI, G. S. 2019. A Local View of Informal Urban Environments: a Mobile Phone-Based Neighborhood Audit of Street-Level Factors in a Brazilian Informal Community. *Journal of Urban Health,* 96**,** 537-548.

RENALDS, A., SMITH, T. H. & HALE, P. J. 2010. A systematic review of built environment and health. *Family & community health,* 33**,** 68-78.

ROSS, C. E. & MIROWSKY, J. 2001. Neighborhood Disadvantage, Disorder, and Health. *Journal of Health and Social Behavior,* 42**,** 258-276.

RUNDLE, A., NECKERMAN, K. M., FREEMAN, L., LOVASI, G. S., PURCIEL, M., QUINN, J., RICHARDS, C., SIRCAR, N. & WEISS, C. 2008. Neighborhood food environment and walkability predict obesity in New York City. *Environmental health perspectives,* 117**,** 442-447.

RUNDLE, A., ROUX, A. V. D., FREEMAN, L. M., MILLER, D., NECKERMAN, K. M. & WEISS, C. C. 2007. The Urban Built Environment and Obesity in New York City: A Multilevel Analysis. *American Journal of Health Promotion,* 21**,** 326-334.

RUNDLE, A. G., BADER, M. D., RICHARDS, C. A., NECKERMAN, K. M. & TEITLER, J. O. 2011a. Using Google Street View to audit neighborhood environments. *American journal of preventive medicine,* 40**,** 94-100.

RUNDLE, A. G., BADER, M. D., RICHARDS, C. A., NECKERMAN, K. M. & TEITLER, J. O. 2011b. Using Google Street View to audit neighborhood environments. *Am J Prev Med,* 40**,** 94-100.

RUNDLE, A. G., BADER, M. D. M., RICHARDS, C. A., NECKERMAN, K. M. & TEITLER, J. O. 2011c. Using Google Street View to Audit Neighborhood Environments. *American Journal of Preventive Medicine,* 40**,** 94-100.

SALLIS, J. F., SAELENS, B. E., FRANK, L. D., CONWAY, T. L., SLYMEN, D. J., CAIN, K. L., CHAPMAN, J. E. & KERR, J. 2009. Neighborhood built environment and income: examining multiple health outcomes. *Social science & medicine,* 68**,** 1285-1293.

SILVA, V., GRANDE, A. J., RECH, C. R. & PECCIN, M. S. 2015. Geoprocessing via google maps for assessing obesogenic built environments related to physical activity and chronic noncommunicable diseases: validity and reliability. *J Healthc Eng,* 6**,** 41-54.

SIMONYAN, K. & ZISSERMAN, A. 2014. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv preprint arXiv:1409.1556*.

SINGH, G. K. 2003. Area deprivation and widening inequalities in US mortality, 1969–1998. *American journal of public health,* 93**,** 1137-1143.

STEVENSON, M., THOMPSON, J., DE SÁ, T. H., EWING, R., MOHAN, D., MCCLURE, R., ROBERTS, I., TIWARI, G., GILES-CORTI, B. & SUN, X. 2016. Land use, transport, and population health: estimating the health benefits of compact cities. *The lancet,* 388**,** 2925-2935.

U.S. CENSUS BUREAU. 2019. *QuickFacts: Utah* [Online]. Available: https://www.census.gov/quickfacts/UT [Accessed June 15, 2021].

VAN CAUWENBERG, J., VAN HOLLE, V., DE BOURDEAUDHUIJ, I., VAN DYCK, D. & DEFORCHE, B. 2016. Neighborhood walkability and health outcomes among older adults: The mediating role of physical activity. *Health & place,* 37**,** 16-25.

Corresponding Author's Declaration Form
Preventive Medicine Reports

(This form is fillable with adobe software or equivalent.)

Manuscript title:	Google Street View images as predictors of patient health outcomes, Utah 2017-2019
Corresponding author:	Quynh C. Nguyen, PhD
Additional authors in the order provided in the manuscript:	Quynh C. Nguyen, PhD, Tom Belnap, PhD, Pallavi Dwivedi, MPH, Amir Hossein Nazem Deligani, Abhinav Kumar, MS, Dapeng Li, PhD, Ross Whitaker, Tolga Tasdizen, PhD , Kimberly D. Brunisholz, PhD

The corresponding author must provide statements of authorship, originality, ethical compliance, conflicts of interest, and research funding on behalf of all authors of the manuscript.

Authorship and Originality:
The corresponding author warrants that all aforementioned authors fulfill the criteria of authorship as defined by the International Committee of Medical Journal Editors (ICMJE). The corresponding author further warrants that the work described in this manuscript has not been published before and is not (nor will be) under consideration elsewhere while under review in *Preventive Medicine Reports*; that all authors approved the present submitted version and their institutions have no objections to the manuscript's contents.

Ethical compliance:
The corresponding author warrants that if the manuscript describes research on human subjects the necessary ethical approval (or exemption) has been obtained and is on file with the authors' institutions. For empirical research papers, add a statement of ethical compliance or exemption to the Methods section.

Conflict of interest:
The corresponding author declares the following financial or other relationships with companies or organizations that are stakeholders on the topic of the manuscript. For details read this fact sheet. Insert the same information in the manuscript as a final disclosure section.

- [x] The authors have no conflicts of interest to disclose or
- [ ] The following authors report specific relationships that could be interpreted as implying a conflict (name author, nature of the relationship, and company or organization):

Funding source:
The corresponding author should declare all sources of funding and any involvement of study sponsors in the study design; collection, analysis and interpretation of data; the writing of the manuscript; or the decision to submit the manuscript for publication. If the study sponsors had no such involvement, this should be stated. Insert the same information in the manuscript as a final disclosure section.

- [ ] There were no sources of funding for this research
- [x] The following authors report financial support (name author, granting organization, grant number, or company sponsoring the study):

Quynh Nguyen, National Library of Medicine, R01LM012849

The corresponding author signed this statement on behalf of all coauthors to indicate that the above information is true, correct and complete.

Signature (both graphic and electronic validation signatures via software are acceptable): Quynh Nguyen Digitally signed by Quynh Nguyen DN: cn=Quynh Nguyen, o, ou, email=qtnguyen@umd.edu, c=US Date: 2021.06.29 14:45:24 -04'00'	Print name: Quynh Nguyen
	Date: 6/29/2021