Title Page

**INTER-RATER RELIABILITY OF RETROSPECTIVE VIRTUAL AUDITS OF THE BUILT ENVIRONMENT**

by

**Alyson B. Harding**

BA Anthropology, BA Chemistry, North Carolina State University, 2013, 2013

Submitted to the Graduate Faculty of

the Department of Epidemiology

Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Master of Public Health

University of Pittsburgh

2019

Committee Membership Page

UNIVERSITY OF PITTSBURGH

Graduate School of Public Health

This essay was submitted

by

**Alyson B. Harding**

on

April 19, 2019

and approved by

Essay Advisor:

Andrea L. Rosso, PhD, MPH, Assistant Professor

Department of Epidemiology, Graduate School of Public Health

University of Pittsburgh

Essay Reader:

Nancy W. Glynn, PhD, Assistant Professor

Department of Epidemiology, Graduate School of Public Health

University of Pittsburgh

Essay Reader:

Jennifer S. Brach, PhD, PT, Professor

Department of Physical Therapy, School of Health and Rehabilitation Sciences

University of Pittsburgh

Copyright © by Alyson B. Harding

2019

Abstract

Andrea L. Rosso, PhD, MPH

Title Page

**INTER-RATER RELIABILITY OF RETROSPECTIVE VIRTUAL AUDITS OF THE BUILT ENVRIONMENT**

Alyson B. Harding, MPH

University of Pittsburgh, 2019

**ABSTRACT:**

Background: Mobility decline is significantly associated with poor health outcomes and high health care costs, and impacts one third to one half of community-dwelling older adults. The built environment and neighborhood walkability play a major role in mobility and other health outcomes, and are of major interest in public health research. While many methods for assessing the built environment exist, the introduction of free, online, street-level images has made conducting virtual neighborhood audits a possibility. Previous studies have assessed the inter-rater reliability between virtual audits and field audits, and between multiple auditors conducting virtual audits with current imagery, but studies have not yet assessed the inter-rater reliability of audits conducted using archived Google Street View imagery.

Methods: Two raters independently conducted 50 neighborhood audits for a subset of participants in the Health, Aging, and Body Composition Study using a modified version of the Active Neighborhood Checklist and the earliest available archived Google Street View imagery. The 77 items of the Checklist were dichotomized, and the prevalence-adjusted bias-adjusted kappa (PABAK) and percent agreement were calculated.

Results: The PABAK values for the 77 items of the modified Active Neighborhood Checklist ranged from 0.20 to 1.00 with an average of 0.75. Eighty percent of items demonstrated substantial (PABAK ≥ 0.61) or nearly perfect agreement (PABAK ≥ 0.81).

Conclusion: Google Street View images provide an efficient, convenient, and safe way to conduct neighborhood audits to measure the built environment. Archived imagery can be used to assess past neighborhood characteristics, allowing for historic data about the built environment to be added to existing cohorts. With the important role of the built environment in health outcomes such as mobility, the ability to assess past neighborhood characteristics with Google Street View imagery has the potential to improve public health research and examination of novel outcomes.

Table of Contents

Preface ix

1.0 Introduction 1

1.1 Mobility in Older Adults 1

1.2 Mobility and the Built Environment 2

1.2.1 Mobility and Residential Characteristics 3

1.2.2 Assessing the Built Environment 4

1.2.2.1 Active Neighborhood Checklist 5

1.2.2.2 Virtual Assessment Tools 6

1.3 Gaps in Knowledge 7

1.4 Public Health Significance 8

2.0 Objective 9

3.0 Methods 10

3.1 Study Participants 10

3.2 Neighborhood Characteristics 11

3.3 Audit Protocol 11

3.4 Statistical Analysis 13

4.0 Results 15

5.0 Discussion 17

Appendix Tables 20

Bibliography 25

List of Tables

Table 1 Participant and Neighborhood Characteristics of subset included in inter-rater reliability analysis and eligible audited sample: Health, Aging, and Body Composition Cohort 20

Table 2 Prevalence-adjusted and bias-adjusted kappa (PABAK) summary statistics for inter-rater reliability analysis of the modified Active Neighborhood Checklist for subsample of 50 participants selected from the Health, Aging, and Body Composition cohort 21

Table 3 Observed agreement, prevalence-adjusted and bias-adjusted kappa (PABAK), and corresponding description of each item of the modified Active Neighborhood Checklist 21

List of Figures

[Figure 1: Flow chart for final analytical sample of Health, Aging, and Body Composition Study participants with valid addresses included in the inter-rater reliablity analysis 24](file:///G%3A%5CHarding%2520Alyson_MPH-1.docx#_Toc21105166)

Preface

I would like to express my sincerest gratitude to the members of my essay committee: Dr. Jen Brach for the valuable feedback provided; Dr. Nancy Glynn for the mentorship, guidance, and endless encouragement through this project and throughout this entire degree program; and Dr. Andi Rosso for being an incredible research mentor. I am grateful for the opportunity to have worked with you and learned from you.

This work would not have been possible without the effort and dedication of Dr. Phillipa Clarke, Ayushi Divecha, and Shala Ward. This study builds on the Health, Aging, and Body Composition Study, supported by the National Institute on Aging (NIA) Contracts N01-AG-6-2101; N01-AG-6-2103; N01-AG-6-2106; NIA grant R01-AG028050, NINR grant R01-NR012459, and NIA grant 5R21 AG054666. This research was funded in part by the Intramural Research Program of the NIH, National Institute on Aging.

To the entire Pitt Public Health community and to the amazing friends I have met through this program, thank you. I could not have gotten through it alone. I would also like to thank my parents and my sisters for the love and support. Finally, thank you to my partner, Hans, for the never-ending encouragement, love, and laughs.

# Introduction

## Mobility in Older Adults

Mobility, the ability of an individual to move about purposefully in their environment, is a key component of health in older adults. Mobility decline affects between one third and one half of adults over the age of 65 [1] and is associated with over $42 billion in additional health care costs annually [2]. Mobility is linked to overall health and quality of life, and is a predictor of disability [1].

Life space mobility measures an individual’s ability to move about in various environments, ranging from the home to beyond one’s town or city. Lower levels of life space mobility are associated with greater disability, or difficulty or inability conducting activities of daily living (ADL) (e.g. eating and walking) and instrumental activity of daily living (IADL) (e.g. preparing meals and managing money) [3]. Lower levels of life space mobility are also related to lower levels of social engagement [3]. Additionally, compared to those with unrestricted life space, those with restrictive life space scores were found to have 4.4 times greater risk of being admitted to a nursing home, independent of ADL scores [4].

Walking, an important form of exercise in older adults, is related to a number of health outcomes [5]. Among community dwelling older adults enrolled in the Health, Aging, and Body Composition Study, those who were excluded from or unable to complete a timed walk of 400 meters (approximately ¼ mile) had higher risk of mortality, cardiovascular disease, and mobility limitation and disability than those who completed the task [6]. Additionally, among those who completely the 400-m walk, each additional minute of time required to complete the task was associated with 29% higher rate of mortality, 20% higher rate of cardiovascular disease, and 52% higher rates of mobility limitation and disability [6]. A study of functionally limited women over the age of 65 revealed that those who walked at least eight blocks per week maintained better functional capacity and walking ability over a one year period than women who did not walk eight blocks per week [5]. In another longitudinal study, older individuals who walked at least the recommended 150 minutes per week at the beginning of the study and after six years of follow up had lower incidence of metabolic syndrome and lower likelihood of developing risk factors of metabolic syndrome compared to those who did not meet the recommendation at one or both time points [7]. Difficulty or inability to walk ¼ mile is associated with increased mortality, greater total annual health care costs, and more hospitalizations [2]. Compared to those with no difficulty walking, annual health care costs were $2,773 higher for those in difficulty walking ¼ mile and $3,919 higher for those who were unable to walk ¼ mile in a cohort of adults over the age of 65 [2]. Ability to walk is a crucial component of mobility and of health in older adults.

## Mobility and the Built Environment

Mobility limitations and restriction are not due to a single cause, but result from complex interactions between demographic, neurological, biomechanical, and environmental factors. The built environment, defined as human made or human altered space, has been shown to play a role in the disablement process [8]. Furthermore, older adults may be more vulnerable to factors in their immediate environment than younger adults, as older adults have endured a longer duration of exposure to their environment and are less likely than younger adults to travel outside their residential areas frequently [8]. Disabled community living older adults reported avoiding obstacles in the environment, including uneven surfaces, stairs, and curbs, and reported avoiding traveling alone and traveling to unfamiliar places [9]. Additionally, the presence of a street in very poor or fair condition (i.e. cracks, potholes, broken curbs) increases the odds of difficultly walking two to three blocks among individuals with neuromuscular impairment [10]. These studies illustrate the impact of the built environment and neighborhood walkability on mobility and the disablement process.

Beyond mobility, the built environment also plays a major role in physical activity, and in turn, body mass index (BMI) and obesity. Individuals living in more walkable areas spend more time walking or biking for transportation, and have lower BMIs than those in less walkable areas [11]. Access to cycle ways, footpaths, health clubs, and swimming pools is associated with physical activity, while neighborhoods that are less aesthetically pleasing have less reported walking [12]. The relationship between the built environment and obesity is an important and highly researched public health question. A recent systematic review discovered over 5,000 original studies examining the association between characteristics of the physical environment and adult weight status between 1995 and 2013 [13].

### Mobility and Residential Characteristics

 The walkability of a neighborhood is only relevant if an individual can access the outdoor environment. Barriers at the entries of homes, including steps and uneven surfaces, may pose as obstacles for older adults. Clarke and colleagues conducted in-person interviews with community dwelling adults over the age of 65, and had interviewers assess the area surrounding the residences for uneven walking surfaces or broken steps leading into the home, as well as presence of stairs or ramps [14]. The study demonstrated that stairs at the entryway of a residence were associated with greater likelihood of reporting difficulty going outside independently [14]. Presence of a ramp at the entryway of a residence also corresponded to higher odds of difficulty going outside independently among those who did not use assistive technology for mobility, including canes and walkers [14]. Another study evaluated barriers at the entry of residences using case managers’ in-person assessments, and determined that those living in a home with barriers at the entrance had nearly 50% higher odds of being homebound [15].

### Assessing the Built Environment

With the increasing understanding of the built environment’s significant contribution to mobility and to public health, there has been interest in developing strategies and tools to systematically assess the environmental characteristics. The built environment can be evaluated using either perceived measures or objective measures [16]. Perceived measures are self-reported assessments of residents, collected through surveys and interviews. While perceived measures can provide valuable information about the relationship between an individual and their environment that cannot be captured by objective measures, these measures are often subject to bias and low response rates [16].

Objective measures include both geographic information systems (GIS)-based measures and direct observational measures [8, 16]. GIS-based measures, or measures of the environment created from existing data sources, are useful for determining a wide variety of neighborhood characteristics, such as income, employment, population composition, housing, and education [16]. While some information, such as census data, is readily available, other GIS-based measures can be both labor and cost intensive to analyze. Direct observational measures, or community audits, involve researchers systematically assessing the features of a neighborhood using a standard checklist or survey. Initially developed for use in the field, surveyors observe the neighborhood on foot or by driving through an area. Audits are generally conducted for a street segment, defined as the section of road between two adjacent intersections [17]. More than 20 different auditing tools exist, which collect information about topics such as land use, sidewalks, streets and traffic, public spaces, parking, and bicycling facilities [16]. Audits may assess macro-level factors such as walkability and street connectivity, or micro-level factors, including presence and quality of sidewalks, intersections, and streets [18].

#### Active Neighborhood Checklist

One commonly used auditing tool, the Active Neighborhood Checklist (the Checklist) was designed in St. Louis, Missouri to assess street-level features that may be related to physical activity. An updated version of the Checklist, released in 2011, assesses 89 items in five areas of focus: land use, public transit, street characteristics, quality of environment for pedestrians, and places to walk and bicycle. The two-page Checklist was designed to be a straightforward and user-friendly way to conduct in-person neighborhood audits with minimal required training.

A 2006 study by Hoehler and colleagues analyzed the inter-rater reliability of the first version of the Active Neighborhood Checklist. Following a two-hour training, researchers conducted field audits for 64 street segments in St. Louis, and demonstrated a mean observed agreement of 0.87 and a mean kappa of 0.68 [19]. Land use items had the highest inter-rater reliability, while the sidewalks and the shoulders and bike lane sections had the lowest. Notably, presence of litter or broken glass, sidewalk width under three feet, and traffic calming devices all had only fair agreement, with kappa values between 0.20 and 0.39 [19].

#### Virtual Assessment Tools

 Although observational audits were designed to be conducted in the field, with the increasing availability of free digital satellite and omnidirectional imagery, many studies are now conducting audits virtually. Google Street View, a component of Google Maps, provides 360-degree panoramic street level images captured every ten to 25 meters [20]. Since its launch in five major United States cities in May of 2007, Google Street View coverage has expanded to 83 countries [21]. In 2014, the timeline feature was released, allowing access to archived Google Street View images, in addition to current images [21].

 Many studies have assessed inter-rater reliability of virtual audits and field audits, showing substantial to near perfect agreement for most audited items, suggesting that virtual audits provide a reliable and convenient way to conduct neighborhood audits [17, 18, 20, 22, 23]. In general, virtual and field audits show high agreement for objective measures, such as presence of infrastructure. Conversely, subjectively assessed items, such as aesthetics and quality of street or sidewalk, have lower agreement [17].

 Wilson and colleagues assessed the reliability of field and virtual audits using the Active Neighborhood Checklist in two urban environments in the United States. The study showed mean prevalence-adjusted bias-adjusted kappa (PABAK) of 0.81, with 86% of items having substantial (PABAK ≥ 0.61) or near perfect agreement (PABAK ≥ 0.81) [22]. Notably, sidewalk width showed poor agreement, and quality of environment items, including presence of liter and broken glass, had moderate agreement [22].

 Another study assessed the inter-rater reliability of two auditors who both used Google Street View to complete the Active Neighborhood Checklist. The mean PABAK for all Checklist items was 0.84, with 95% of the items having substantial or near perfect agreement [24]. Items with the lowest reliability were on-street parking, tree shade, sidewalk width, and curb cuts, while sidewalk continuity and presence of sidewalk obstructions demonstrated perfect agreement [24]. This study provides further evidence that utilizing Google Street View technology results in reliable neighborhood audits. The following study builds on this evidence by examining inter-rater reliability of virtual audits conducted using archived Google Street View images.

## Gaps in Knowledge

 Literature demonstrates that current Google Street View imagery provides a convenient, safe, and efficient way to conduct virtual neighborhood audits. However, studies have not yet analyzed the inter-rater reliability of virtual audits conducted using archived Google Street View images.

 Additionally, studies have demonstrated the importance of residential characteristics and barriers at the entryway of homes to mobility of older adults. However, previous studies have examined these factors in person, using interviewers or caseworkers to determine presence of stairs or uneven surfaces near the entrance of a residence [14, 15]. The possibility of using Google Street View or other free online imagery to assess residential characteristics has not been explored.

## Public Health Significance

Mobility decline is significantly associated with poor health outcomes and high health care costs. Numerous studies demonstrate the impact of environmental characteristics on mobility and other aspects of health, however most existing cohorts fail to consider the built environment of participants. Using Google Street View images provides a cost effective, efficient, safe, and valid way to conduct neighborhood audits to gain additional information about factors that may influence participant health outcomes.

Adding environmental characteristics to existing cohorts allows public health researchers to gain a better understanding of the impact of the built environment on health using existing data sources. Additionally, knowledge of specific elements of the built environment that positively or negatively impact mobility and physical activity could directly influence public policy decision-making regarding the build environment. Considering the influence of the built environment on health outcomes is essential, and the ability to retrospectively audit neighborhoods and add environmental characteristics to existing cohorts has major public health implications.

# Objective

This study aims to establish inter-rater reliability of a modified version of the Active Neighborhood Checklist for neighborhood audits in a cohort of community dwelling older individuals using archived web-based images.

# Methods

## Study Participants

A total of 3075 community dwelling older adults between the ages of 70-79 at baseline were enrolled in the Health, Aging, and Body Composition (Health ABC) Study in 1996-97 in Pittsburgh, PA and Memphis, TN [7]. All participants were otherwise healthy, able to walk ¼ mile, climb 10 steps, and were independent in performing all activities of daily living at baseline. Residents of Allegheny County and the contiguous surrounding counties who attended the tenth year of follow up (2007) were included in this study. Of the 795 participants who attended the tenth year of follow up at the Pittsburgh site, 45 of them lived outside the state of Pennsylvania and four lived outside the greater Pittsburgh area, resulting in 746 eligible participants in this study. One address could not be located, leaving a final total of 745 participants included in this study (Figure 1). Of these, the addresses of 13 participants lacked Google Street View imaging. In these cases, auditors utilized Google Satellite View imagery, with those street-level characteristics that could not be assessed from satellite images coded as missing.

The addresses of the first 50 eligible Health ABC participants were selected for inter-rater reliability analysis. Google Street View availability was checked for these addresses, and those without Google Street View imagery were replaced by the address corresponding to the next available identification number, until a total of 50 addresses were selected. Although this subsample was not randomly selected, the Health ABC ID numbers were assigned in order of recruitment, and were not representative of any outside factors.

## Neighborhood Characteristics

Census tract and census block corresponding to each participant address were obtained using the online Census Geocoder tool [25]. Publicly available 2010 census data from the United States Census Bureau provided neighborhood characteristics at the census tract level [26]. These characteristics included median individual income, median household income, median value of housing units, percentage of households receiving interest, dividend, or net rental income, percentage of adults with high school education, percent of adults with a bachelor’s degree, percentage of adults in managerial or professional specialty occupations, and median age of residents. Data collected from the United States Department of Agriculture (USDA) included population density and urban or rural setting [27]. The USDA Rural-Urban Commuting Area Codes were categorized into three tiers to describe urban or rural setting: metropolitan area, micropolitan area, and small town.

## Audit Protocol

The Active Neighborhood Checklist (the Checklist) has been used to conduct virtual audits with Google Street View technology and showed substantial to near perfect agreement with field audits [22]. The Checklist was modified to eliminate items with known difficulty of assessment using web-based images (e.g. litter, graffiti) [20, 22, 24], to remove bicycle related items that were not relevant to the study population, and to address residential entrance characteristics. The modified Active Neighborhood Checklist evaluated crucial environmental characteristics including land use (e.g. predominately residential, commercial), presence of public transport (e.g. bus stops, benches, and shelters), street characteristics (e.g. number of lanes, presence of stop signs), quality of physical environment (e.g. slope, lighting, trees), quality of sidewalk (e.g. presence of sidewalks, width), as well as residential characteristics. Residential characteristics included the following items from the National Health and Aging Trends Study: type of physical structure, presence of uneven surfaces or broken steps leading up to the home, presence and number of stairs at the entrance, and presence of a ramp [14] (for complete list of audited items, see Table 3).

Two trained auditors independently conducted retrospective virtual audits using the modified Active Neighborhood Checklist and archived Google Street View imagery. Raters performed trial audits before commencing the subsample of 50 audits to identify real time concerns that could limit the quality and completeness of web based audits.

Both auditors audited 1/8th mile or 660 feet distance on each side of a given residence using the earliest Google Street View images available. Although the Active Neighborhood Checklist was designed to audit a single street segment, defined as the distance between two consecutive intersections, the distance of ¼ mile was selected to be consistent with the mobility disability literature, which uses ability to walk ¼ mile as an indicator. Additionally, many streets in the greater Pittsburgh area are not laid out on a grid system, leading to great variation in the distance between adjacent intersections. The establishment of ¼ mile distance of the audits allowed for consistency between audits. A standard operating manual was created to ensure standardization of alternate audit decisions where warranted. For instance, the auditor turned right when encountering a T-intersection, and the auditor audited the remainder of the ¼ mile distance on the other segment when a cul-de-sac restricted the distance of the audited segment.

## Statistical Analysis

Descriptive statistics for demographic characteristics such as age, gender, race, education, body mass index, cognitive status, and difficulty walking ¼ mile were presented in order to understand how well the subsample represented the larger cohort.

Each auditor conducted left and right sides of an audit independently, with 660 feet on both the left side and the right side of a residence audited, for a total of 1320 feet (¼ mile) per audit. These audit segments were combined to a single audit, where if an item was present for either segment, it was considered present for the audit. For three specific sidewalk characteristics: continuous sidewalk, wide sidewalk, and ramps or curb cuts present, the items were coded as absent if they were absent for either segment. That is, if the sidewalk was not continuous for either segment, it was considered not continuous for the audit. Audits were analyzed as one audit per address rather than utilizing the separate segments to account for correlation between the two segments.

Consistent with previous studies, all items were dichotomized [22, 24]. Items that were recorded as present on one side of the street, present on both sides of the street, or absent were coded as present (one or both sides of the street) or absent. Similarly, slope was coded as present (moderate or steep slope) versus absent (flat) and street size was dichotomously coded as two lanes versus larger street. Finally, type of physical structure was coded as single family or row home versus multi-unit building or senior care home. A total of 77 items, including the residential characteristics, were analyzed.

The nine predominant land use variables were recoded to create three key predominant land use variables. If residential buildings were a predominant land use, and commercial buildings, schools, and parking lots and garages were absent, the predominant land use variable was coded as residential. If residential buildings were a predominant land use, and commercial buildings, schools, or parking lots and garages, or a combination of these were also predominant, the land use was labeled a mixed use. If residential buildings were not predominant, the predominant land use was labeled as non-residential. Inter-rater reliability was also calculated for these variables.

Inter-rater reliability was evaluated by computing Cohen’s kappa coefficient and observed agreement between two raters using SAS version 9.4 software (SAS Institute, Cary, NC). When the prevalence of one response is very high or very low, the kappa may suggest poor reliability, even if the observed agreement is high. To combat this issue, the prevalence-adjusted and bias-adjusted kappa (PABAK) was also calculated [28].

The following adjectives were used to describe PABAK values: <0.00 (poor agreement), 0.00 to 0.20 (slight agreement), 0.21 to 0.40 (fair agreement), 0.41 to 0.60 (moderate agreement), 0.61 to 0.80 (substantial agreement), and 0.81 to 1.00 (almost perfect agreement) [29]. These descriptive terms, introduced by Landis and Koch, are commonly used in the inter-rater reliability literature [18, 23].

# Results

The participants selected for the 50 audits used for the inter-rater reliability analysis were representative of the entire audited cohort for many participant characteristics, including age, Modified Mini-Mental State (3MS) score, and percentage of participants with difficulty walking ¼ mile (Table 1). However, differences in neighborhood characteristics exist between the groups. The subset of 50 participants included in the inter-rater reliability analysis lived in census tracts with a higher percentage of adults with bachelor’s degrees (64.3 vs. 30.0) and working in managerial or professional positions (57.5 vs. 42.0). The neighborhoods of the subset of 50 participants had a median population density of nearly 9,500 people per square mile, while the median population density of the neighborhoods represented in the entire sample was approximately 6,000 people per square mile. Additionally, the subset neighborhoods had higher median individual income and median value of housing units.

The results of the inter-rater reliability analysis of the subsample of 50 audits are summarized in Table 2 and presented in their entirety in Table 3. The PABAK values ranged from slight agreement (0.20) to perfect agreement (1.00). The mean PABAK of the 77 items of the modified Checklist was 0.75. Eighty percent of items had substantial (PABAK ≥ 0.61) or almost perfect agreement (PABAK ≥ 0.81).

Public recreational facilities and residential characteristics had the strongest agreement, with average PABAK values of 0.89 and 0.90 respectively. All four of the residential characteristics, and five of the six items in the public recreational facility category had almost perfect agreement, with the final variable demonstrating substantial agreement.

The non-residential use variables had PABAK values ranging from 0.20 to 1.00, with an average of 0.80. Sidewalk characteristics, quality of environment, and land use items had average PABAK values between 0.74 and 0.77. The three variables derived from the nine predominant land use variables of the modified Checklist, predominately residential, predominately mixed land use, and predominately non-residential, had PABAK values of 0.72, 0.76, and 0.96, respectively.

Residential use items ranged in PABAK values from 0.24 (multi-unit homes) to 0.96 (any residential use present), with an average of 0.67. Street characteristics had the poorest agreement, with an average PABAK of 0.59. Of the 12 items in the street characteristic category, no items had almost perfect agreement, with seven items demonstrating substantial agreement, three showing moderate agreement, and two having slight to fair agreement.

# Discussion

Of the 77 items included in the modified Active Neighborhood Checklist, 31 showed almost perfect agreement and 31 had substantial agreement. The PABAK values ranged from 0.20 to 1.00, with an average of 0.75. Overall, residential characteristics and non-residential uses had the strongest agreement, while street characteristics had the lowest agreement. These results suggest that retrospective neighborhood audits using archived Google Street View imagery yield inter-rater reliability scores consistent with field audits and audits conducted using current imagery.

Consistent with previous studies, the PABAK values of sidewalk width (0.40), presence of trees (0.56), and presence of traffic calming devices (0.20) demonstrated low inter-rater reliability [19, 22, 24]. These items should be interpreted with caution in existing studies, or potentially better defined in future study protocols to improve inter-rater reliability.

The non-residential use category contained 20 different items and had the largest range in PABAK values, from 0.20 (other non-residential use) to 1.00 (mall, indoor fitness facility, post office) with an average of 0.80. The variables with the lowest agreement were both the “other” categories (other non-residential use and other large building present). When these two variables were removed, the average PABAK for non-residential use increased to 0.85, and all remaining items in the non-residential use category show substantial or almost perfect agreement. The “other residential use” variable also only had moderate agreement, although the “other predominate land use” variable had substantial agreement. These “other” items should be examined with caution, since they are not as clearly defined and identifiable as the more specific items of the Checklist.

Although the majority of the residential uses demonstrated substantial or near perfect agreement between the raters, the city and infrastructure of Pittsburgh likely contributed to the lower agreement of multi-unit homes. In Pittsburgh, many historically single family homes have been converted into multi-unit apartments, even though these houses have one exterior door. Differences in knowledge about this style of housing units likely caused the fair agreement of multi-unit homes, which was the poorest agreement among types of residential uses. In other studies, the influence of background knowledge has been avoided by having virtual audits conducted by researchers who are not residents of the city being audited [18, 22].

One major limitation encountered with retrospective Google Street View analysis was the poor resolution in 2007 images, as the resolution of images taken in 2007 was lower than those taken in 2008 or later. The blurry images made it challenging to examine details such as signage (e.g. limited hour parking signs) and sidewalk quality. Additionally, Google Street View uses an omnidirectional camera to take pictures every ten to 25 meters, so the images may “hop” as an auditor moves down the street. Depending on how close together these images are taken, it is possible to miss features, such as signs. Trees, bushes, or cars parked along a street may also obscure important details.

Despite these limitations, 80% of the items from the modified Checklist had substantial or almost perfect agreement. The residential characteristics of type of residence, presence of stairs at the entrance, presence of ramp, and presence of uneven or broken surfaces leading to the entrance of residence all had almost perfect agreement. Although this information was missing for some residences due to limitations of Google Street View images, stairs, ramp, and uneven surfaces were assessed for 86% of the 50 addresses audited. While these items are not included in the original Checklist, it has been suggested that these factors contribute to difficulty with outdoor mobility in older adults [14, 15], so the ability to evaluate them using Google Street View imagery may play an important role in the mobility literature. Additionally, types of non-residential use (e.g. library, church) showed substantial or almost perfect agreement, as did public recreational facilities and many environmental, street, and sidewalk characteristics, suggesting that archived Google Street View images can be used with the validated Active Neighborhood Checklist to conduct neighborhood audits retrospectively.

The subset of 50 addresses used in the inter-rater reliability analysis was limited to urban environments. Each address was located in a census tract that the United States Department of Agriculture labeled as a metropolitan area, so these results may not be generalizable to suburban or rural environments. Additionally, variability exists in the availability of Google Street View images and the date of the earliest images available. In the subsample of 50 audits, 44 of the audits were conducted with images taken in 2007, while six audits utilized images from 2008. The differences in the availability of archived Google Street View images could pose a limitation to the method of using Google Street View for retrospective neighborhood audits, particularly in areas that have undergone drastic changes, or in cohorts where the earliest available images for participants vary by many years.

Images from Google Street View offer an efficient and effective way to assess street level characteristics and the built environment. Many cohorts exist that contain medical and laboratory data, but so far have failed to account for the importance of the built environment of participants. This study and the availability of archived Google Street View images demonstrate how street level characteristics and data about the built environment can be retroactively applied to and analyzed for existing cohorts, allowing health outcomes to be considered in an ecological framework and improving public health research methods.

Appendix Tables

Table 1 Participant and Neighborhood Characteristics of subset included in inter-rater reliability analysis and eligible audited sample: Health, Aging, and Body Composition Cohort

|  |  |  |
| --- | --- | --- |
|  | **Subset (50)** | **All participants (745)** |
| **Participant Characteristics** | Mean (SD) or n (%) | Mean (SD) or n (%) |
| Age of participants, years | 83.0 (2.7)  | 82.3 (2.7) |
| Females, % | 21 (42.0) | 427 (57.3) |
| Percent African-American, % | 19 (38.0) | 288 (38.7) |
| High School education or less, % | 16 (32.0) | 382 (53.0) |
| Body Mass Index, kg/m2 | 26.6 (4.2) | 27.8 (4.7) |
| Modified Mini-Mental State Score, 0-100 pts | 92.2 (6.9) | 92.1 (7.9) |
| Difficulty walking ¼ mile distance, % | 14 (28.0) | 186 (28.8) |
|  |  |  |
| **Neighborhood Characteristicsa** | Median (IQR) | Median (IQR) |
| Median household income, $ | 35,290 (26,139) | 41,531 (24,198) |
| Median individual income, $ | 33,567 (23,409) | 25,235 (13,290) |
| Median value of housing unit, $ | 154,300 (257,500) | 94,200 (88,600) |
| Households receiving interest, dividends, or net rental income, % | 26.8 (23.1) | 26.4 (24.2) |
| Adults with high school degree or higher, %  | 93.9 (7.3) | 93.3 (7.5) |
| Adults with bachelor’s degree or higher, % | 64.3 (48.2) | 30.0 (39.5) |
| Adults in managerial or professional positions, % | 57.5 (37.6) | 42.0 (28.6) |
| Median age of census tract residents, years | 37.2 (12.4) | 41.2 (9.6) |
| Population density, persons per square mile | 9460.8 (9146.8) | 5996.6 (6354.1) |
| Census tracts in metropolitan area, %  | 100 | 98.9 |

aNeighborhood characteristics of census track

Table 2 Prevalence-adjusted and bias-adjusted kappa (PABAK) summary statistics for inter-rater reliability analysis of the modified Active Neighborhood Checklist for subsample of 50 participants selected from the Health, Aging, and Body Composition cohort

|  |  |  |
| --- | --- | --- |
|  | Agreement | Number of items in Landis and Koch value range |
| Audit Tool Section | PABAK mean (range) | 0-0.40, slight to fair | 0.41-0.60 moderate | 0.61-0.80 substantial | 0.81-1.00 almost perfect |
| Residential Characteristics (4) | 0.90 (0.81-1.00) | 0 | 0 | 0 | 4 |
| Land Use (8) | 0.74 (0.40-0.96) | 1 | 0 | 5 | 2 |
| Residential Use (7) | 0.67 (0.24-0.96) | 1 | 1 | 3 | 2 |
| Public Recreational Facilities (6) | 0.89 (0.76-0.96) | 0 | 0 | 1 | 5 |
| Non-residential use (20) | 0.80 (0.20-1.00) | 2 | 0 | 5 | 13 |
| Quality of Environment (12) | 0.74 (0.56-0.88) | 0 | 3 | 7 | 2 |
| Street characteristics (12) | 0.59 (0.20-0.80) | 2 | 3 | 7 | 0 |
| Sidewalk characteristic (8) | 0.77 (0.40-1.00) | 1 | 1 | 3 | 3 |
| Total (77) | 0.75 (0.20-1.00) | 7 | 8 | 31 | 31 |

Table 3 Observed agreement, prevalence-adjusted and bias-adjusted kappa (PABAK), and corresponding description of each item of the modified Active Neighborhood Checklist

|  |  |  |  |
| --- | --- | --- | --- |
| Variable | Observed Agreement | PABAK | Descriptiona |
| **Residential Characteristics** |  |  |  |
| Type of Residence | 0.94 | 0.88 | Almost perfect |
| Stairs at entrance | 0.96 | 0.91 | Almost perfect |
| Ramp at entrance | 0.91 | 0.81 | Almost perfect |
| Uneven or broken | 1.00 | 1.00 | Perfect |
|  |  |  |  |
| **Land Use** |  |  |  |
| Residential buildings | 0.98 | 0.96 | Almost perfect |
| Commercial, institutional, office, industrial | 0.88 | 0.76 | Substantial |
| Schools | 0.84 | 0.68 | Substantial |
| Parking lots, garages | 0.70 | 0.40 | Fair |
| Parks, exercise facilities | 0.92 | 0.84 | Almost perfect |
| Abandoned buildings, vacant lots | 0.90 | 0.80 | Substantial |
| Designated green space | 0.88 | 0.76 | Substantial |
| Other non residential | 0.84 | 0.68 | Substantial |
| **Table 3 Continued** |  |  |  |
| Predominantly residentialb | 0.86 | 0.72 | Substantial |
| Predominantly mixed land useb | 0.88 | 0.76 | Substantial |
| Predominantly non-residentialb | 0.98 | 0.96 | Almost perfect |
|  |  |  |  |
| **Types of residential use** |  |  |  |
| Are any residential uses present? | 0.98 | 0.96 | Almost perfect |
| Abandoned homes | 0.84 | 0.68 | Substantial |
| Single families | 0.86 | 0.72 | Substantial |
| Multi unit | 0.62 | 0.24 | Fair |
| Small apartment | 0.84 | 0.68 | Substantial |
| Large apartment | 0.92 | 0.84 | Almost perfect |
| Other apartment | 0.78 | 0.56 | Moderate |
|  |  |  |  |
| **Public Recreational Facilities** |  |  |  |
| Are public recreational facilities present? | 0.88 | 0.76 | Substantial |
| Park | 0.92 | 0.84 | Almost perfect |
| Off road trail | 0.98 | 0.96 | Almost perfect |
| Sport or playing field | 0.98 | 0.96 | Almost perfect |
| Playground | 0.94 | 0.88 | Almost perfect |
| Pool | 0.96 | 0.92 | Almost perfect |
|  |  |  |  |
| **Non-residential uses** |  |  |  |
| Are any non-residential uses present? | 0.86 | 0.72 | Substantial |
| Convenience store | 0.90 | 0.80 | Substantial |
| Supermarket | 0.94 | 0.88 | Almost perfect |
| Church | 0.94 | 0.88 | Almost perfect |
| Food establishment | 0.88 | 0.76 | Substantial |
| Entertainment | 0.94 | 0.88 | Almost perfect |
| Library | 0.96 | 0.92 | Almost perfect |
| Post Office | 1.00 | 1.00 | Perfect |
| Bank | 0.98 | 0.96 | Almost perfect |
| Laundry | 0.98 | 0.96 | Almost perfect |
| Fitness | 1.00 | 1.00 | Perfect |
| Educational facilities | 0.82 | 0.64 | Substantial |
| Other non-residential uses | 0.60 | 0.20 | Slight |
| Big box store | 0.92 | 0.84 | Almost perfect |
| Mall | 1.00 | 1.00 | Perfect |
| Strip mall | 0.92 | 0.84 | Almost perfect |
| Offices, warehouses, factory, industry | 0.88 | 0.76 | Substantial |
| Hospital | 0.96 | 0.92 | Almost perfect |
| Medical Facility | 0.98 | 0.96 | Almost perfect |
| Other large building or business | 0.70 | 0.40 | Fair |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
| **Street Characteristics****Table 3 Continued** |  |  |  |
| Speed limit sign | 0.88 | 0.76 | Substantial |
| Special speed zone | 0.86 | 0.72 | Substantial |
| Street size | 0.90 | 0.80 | Substantial |
| Marked lanes | 0.86 | 0.72 | Substantial |
| Median | 0.88 | 0.76 | Substantial |
| Turn lane | 0.82 | 0.64 | Substantial |
| Limited hour parking sign | 0.72 | 0.44 | Moderate |
| Stop Sign | 0.68 | 0.36 | Fair |
| Stop light without walk sign | 0.72 | 0.44 | Moderate |
| Cross walk | 0.74 | 0.48 | Moderate |
| Traffic calming device | 0.60 | 0.20 | Slight |
| Cul-de-sac | 0.88 | 0.76 | Substantial |
|  |  |  |  |
| **Environmental Quality** |  |  |  |
| Ongoing construction | 0.84 | 0.68 | Substantial |
| Commercial buildings adjacent to sidewalk | 0.86 | 0.72 | Substantial |
| Pedestrian bench | 0.86 | 0.72 | Substantial |
| Lighting | 0.94 | 0.88 | Almost perfect |
| Trees of sidewalk or road | 0.78 | 0.56 | Moderate |
| Slope | 0.80 | 0.60 | Moderate |
| Steps to get to shops or houses | 0.94 | 0.88 | Almost perfect |
| Transit stop | 0.90 | 0.80 | Substantial |
| Transit bench | 0.88 | 0.76 | Substantial |
| Transit shelter | 0.88 | 0.76 | Substantial |
| Off Street Parking | 0.80 | 0.60 | Moderate |
| On Street Parking | 0.82 | 0.64 | Substantial |
|  |  |  |  |
| **Sidewalk characteristics** |  |  |  |
| Sidewalk present | 0.98 | 0.96 | Almost perfect |
| Buffer present | 0.80 | 0.60 | Moderate |
| Sidewalk continuous | 0.94 | 0.88 | Almost perfect |
| Sidewalk wide | 0.70 | 0.40 | Fair |
| Sidewalk narrow | 0.86 | 0.72 | Substantial |
| Ramps or curb cuts present | 0.90 | 0.80 | Substantial |
| Bumps, cracks, holes, weeds | 0.88 | 0.76 | Substantial |
| Permanent obstructions | 1.00 | 1.00 | Perfect |
| aLandis and Koch descriptions: <0.00 (poor agreement), 0.00-0.20 (slight agreement), 0.21-0.40 (fair agreement), 0.41-0.60 (moderate agreement), 0.61-0.80 (substantial agreement) 0.81-1.00 (almost perfect agreement) bVariables derived from land use Checklist items  |

N= 3075 Adults aged 70-79 years enrolled in the Health, Aging, and Body Composition Study

(1997-1998)

N= 1527 Pittsburgh enrollees

N= 795 Completed Year 10 follow up (2007)

N=745 Living in the Pittsburgh area with valid address

N= 1548 enrolled at Memphis field center

N= 732 lost to follow up or deceased

N=49 residing outside the greater Pittsburgh region

N=1 address could not be located

N=50 Final analytical sample for inter-rater reliability analysis

Figure 1: Flow chart for final analytical sample of Health, Aging, and Body Composition Study participants with valid addresses included in the inter-rater reliablity analysis

Bibliography

1. Webber, S.C., M.M. Porter, and V.H. Menec, *Mobility in older adults: a comprehensive framework.* Gerontologist, 2010. **50**(4): p. 443-50.

2. Hardy, S.E., et al., *Ability to walk 1/4 mile predicts subsequent disability, mortality, and health care costs.* J Gen Intern Med, 2011. **26**(2): p. 130-5.

3. Rosso, A.L., et al., *Mobility, disability, and social engagement in older adults.* J Aging Health, 2013. **25**(4): p. 617-37.

4. Sheppard, K.D., et al., *Life-space mobility predicts nursing home admission over 6 years.* J Aging Health, 2013. **25**(6): p. 907-20.

5. Simonsick, E.M., et al., *Just get out the door! Importance of walking outside the home for maintaining mobility: findings from the women's health and aging study.* J Am Geriatr Soc, 2005. **53**(2): p. 198-203.

6. Newman, A.B., et al., *Association of long-distance corridor walk performance with mortality, cardiovascular disease, mobility limitation, and disability.* Jama, 2006. **295**(17): p. 2018-2026.

7. Peterson, M.J., et al., *Walking in old age and development of metabolic syndrome: the health, aging, and body composition study.* Metabolic syndrome and related disorders, 2010. **8**(4): p. 317-322.

8. Rosso, A.L., A.H. Auchincloss, and Y.L. Michael, *The urban built environment and mobility in older adults: a comprehensive review.* J Aging Res, 2011. **2011**: p. 816106.

9. Shumway-Cook, A., et al., *Environmental components of mobility disability in community-living older persons.* J Am Geriatr Soc, 2003. **51**(3): p. 393-8.

10. Clarke, P., et al., *Mobility disability and the urban built environment.* Am J Epidemiol, 2008. **168**(5): p. 506-13.

11. Frank, L.D., et al., *Many pathways from land use to health: associations between neighborhood walkability and active transportation, body mass index, and air quality.* Journal of the American planning Association, 2006. **72**(1): p. 75-87.

12. Humpel, N., N. Owen, and E. Leslie, *Environmental factors associated with adults’ participation in physical activity: a review.* American journal of preventive medicine, 2002. **22**(3): p. 188-199.

13. Mackenbach, J.D., et al., *Obesogenic environments: a systematic review of the association between the physical environment and adult weight status, the SPOTLIGHT project.* BMC public health, 2014. **14**(1): p. 233.

14. Clarke, P.J., *The role of the built environment and assistive devices for outdoor mobility in later life.* J Gerontol B Psychol Sci Soc Sci, 2014. **69 Suppl 1**: p. S8-15.

15. Clarke, P. and N.A. Gallagher, *Optimizing mobility in later life: the role of the urban built environment for older adults aging in place.* J Urban Health, 2013. **90**(6): p. 997-1009.

16. Brownson, R.C., et al., *Measuring the built environment for physical activity: state of the science.* Am J Prev Med, 2009. **36**(4 Suppl): p. S99-123 e12.

17. Charreire, H., et al., *Using remote sensing to define environmental characteristics related to physical activity and dietary behaviours: a systematic review (the SPOTLIGHT project).* Health Place, 2014. **25**: p. 1-9.

18. Phillips, C.B., et al., *Online versus in-person comparison of Microscale Audit of Pedestrian Streetscapes (MAPS) assessments: reliability of alternate methods.* Int J Health Geogr, 2017. **16**(1): p. 27.

19. Hoehner, C.M., et al., *Active neighborhood checklist: a user-friendly and reliable tool for assessing activity friendliness.* Am J Health Promot, 2007. **21**(6): p. 534-7.

20. Clarke, P., et al., *Using Google Earth to conduct a neighborhood audit: reliability of a virtual audit instrument.* Health Place, 2010. **16**(6): p. 1224-9.

21. Raman, A. *Cheers to Street View’s 10th Birthday!* The Keyword 2017; Available from: <https://www.blog.google/products/maps/cheers-street-views-10th-birthday/>.

22. Wilson, J.S., et al., *Assessing the built environment using omnidirectional imagery.* Am J Prev Med, 2012. **42**(2): p. 193-9.

23. Badland, H.M., *Can virtual streetscape audits reliably replace physical streetscape audits?* J Urban Health, 2010. **87**(6): p. 1007-16.

24. Kelly, C.M., et al., *Using Google Street View to audit the built environment: inter-rater reliability results.* Ann Behav Med, 2013. **45 Suppl 1**: p. S108-12.

25. United States Census Bureau. *Geocoder*. Available from: <https://geocoding.geo.census.gov>.

26. United States Census Bureau. *American FactFinder*. Available from: <https://factfinder.census.gov>.

27. United States Department of Agriculture Economic Research Service, *Rural-Urban Commuting Area Codes*. 2016.

28. Cunningham, M., *More than just the kappa coefficient: a program to fully characterize inter-rater reliablity between two raters*, in *SAS global forum*. 2009. p. 242.

29. Landis, J.R. and G.G. Koch, *The measurement of observer agreement for categorical data.* Biometrics, 1977. **33**(1): p. 159-74.