



Using Google Earth to conduct a neighborhood audit: Reliability of a virtual audit instrument

Philippa Clarke*, Jennifer Ailshire¹, Robert Melendez, Michael Bader², Jeffrey Morenoff

Institute for Social Research, University of Michigan, Ann Arbor, MI, USA

ARTICLE INFO

Article history:

Received 19 March 2010

Received in revised form

23 July 2010

Accepted 4 August 2010

Keywords:

Neighborhoods

Audit

Google Earth

Reliability

ABSTRACT

Over the last two decades, the impact of community characteristics on the physical and mental health of residents has emerged as an important frontier of research in population health and health disparities. However, the development and evaluation of measures to capture community characteristics is still at a relatively early stage. The purpose of this work was to assess the reliability of a neighborhood audit instrument administered in the city of Chicago using Google Street View by comparing these “virtual” data to those obtained from an identical instrument administered “in-person”. We find that a virtual audit instrument can provide reliable indicators of recreational facilities, the local food environment, and general land use. However, caution should be exercised when trying to gather more finely detailed observations. Using the internet to conduct a neighborhood audit has the potential to significantly reduce the costs of collecting data objectively and unobtrusively.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

1.1. Characterizing neighborhoods in health research

Over the last two decades, the impact of community characteristics on the physical and mental health of residents has emerged as an important frontier of research in population health and health disparities (Diez Roux, 2001; 2004; O'Campo, 2003; Sampson et al., 2002). The measurement of community characteristics is evolving, but strategies typically fall under one of three categories of measurement: secondary analysis of archival data sources, perceived (self-reported) responses in a community survey, and objective audit instruments (Brownson et al., 2009). Using secondary data from administrative sources (e.g. decennial census), both to define neighborhoods and as an aggregate measure of neighborhood characteristics, researchers have examined the relationship between various health outcomes and factors such as population density (Lopez, 2004), land use diversity (Clarke and George, 2005; Cervero and Duncan, 2003), and block size (Boer et al., 2007). These archival data are often enhanced using geographic information systems (GIS) to incorporate data on characteristics such as traffic volume (Tonne et al., 2007),

street connectivity (McGinn et al., 2007), the availability of food (Bader et al., 2010), and recreational facilities (Diez Roux, Evenson et al., 2007) within local neighborhoods.

Tapping individuals' perceptions of their environments is another common measurement strategy (e.g. Moore et al., 2008), particularly in the research on physical activity and the built environment (Brownson et al., 2009). However, subjective reports from respondents are subject to same-source bias (e.g. those in poor health are more likely to report poorer neighborhood conditions) (Echeverría et al., 2008), and conflicting findings can arise when using both subjective and archival measures (McGinn et al., 2007). As an alternative, direct observation of neighborhood characteristics using an audit instrument relies on more objective measurement to capture many of the comprehensive and detailed environmental characteristics relevant for health (Clifton et al., 2007; Clarke et al., 2008; Schaefer-McDaniel et al., 2010). While driving or walking through small-area respondent-centered neighborhoods, researchers observe and document neighborhood features using a standardized instrument (e.g. Pikora et al., 2002). The direct observational method known as systematic social observation (SSO) is a measurement strategy used in the social sciences (Reiss, 1971; Raudenbush and Sampson, 1999; Sampson and Raudenbush, 1999) whereby survey interviewers or raters systematically rate each respondent's neighborhood block (e.g. condition of the street, presence of litter, and heavy traffic) during the survey period. However, these in-person audits are highly resource intensive and costly, making them prohibitive for many studies.

The development and evaluation of measures to capture community characteristics are still at a relatively early stage (Brownson et al., 2009; Sallis, 2009), and only a few studies have

* Corresponding author. Tel.: +1 734 647 9611; fax: +1 734 936 0548.

E-mail address: pjclarke@umich.edu (P. Clarke).

¹ Present address: Center on Biodemography and Population Health, University of Southern California, Los Angeles, CA, USA.

² Present address: Robert Wood Johnson Foundation Health and Society Scholars Program and Leonard Davis Institute for Health Economics, University of Pennsylvania, Philadelphia, PA, USA.

explicitly compared measurement properties across different strategies (e.g. Bader et al., 2010). The purpose of this work was to assess the reliability of a neighborhood audit instrument administered using the internet by comparing these “virtual” data to those obtained from an identical instrument administered “in-person”. Using the internet to conduct a neighborhood audit has the potential to significantly reduce the costs of collecting data “objectively and unobtrusively” (Brownson et al., 2009). Our objective in this work is to ascertain the reliability of this method by capitalizing on existing data that were collected as part of a study on neighborhoods and health in the city of Chicago.

1.2. Using the internet for a neighborhood audit

Recently, there has been dramatic growth in internet capacities for observing and characterizing small area neighborhoods. Google Earth (Google Inc., 2005) is a free, internet-based software that displays satellite images of the earth's surface at a resolution of 15 m or higher. Google Street View is a relatively new technology featured in Google Earth that provides 360° horizontal and 290° vertical panoramic views at the street level (based on images taken at approximately 10 or 20 m intervals) from a height of about 2.5 m. Thus, Google Street View gives the viewer the feeling of virtually being on the street and the capacity to virtually walk down that street. Street View was launched on May 25, 2007, in several major US cities, and has been expanding to include coverage throughout the world.

The highly detailed imagery available in Google Street View raises the possibility of conducting a “virtual” neighborhood audit. Despite the widespread availability of visual data on community and built environments, few studies have utilized such electronic images on the internet to characterize neighborhood environments (Curtis et al., 2010; Doyle et al., 1998). In this paper we assess the level of agreement between street level characteristics documented by trained raters using SSO as part of a community-based survey in the city of Chicago, and data collected with an identical instrument using Google Street View. This is a case study that draws on existing data collected “in-person” in 2002, and collects comparable data using Google Street View when it became available 4–5 years later. While we would ideally like to have had more contemporaneous measurement occasions, cost considerations prohibited the collection of data solely for this purpose. Rather, this is an opportunistic study that draws on existing data to conduct a case study in Chicago, offering initial insight into the reliability of a virtual method. We hope this is a first step in considering the utility of this method and that other researchers will replicate such analyses in other settings with better temporal alignment of data.

2. Methods

2.1. Data

Data obtained from the Chicago Community Adult Health Study (CCAHS), which was conducted in 2002 through face-to-face interviews with a multi-stage probability sample of 3105 adults aged 18 and over, living in the city of Chicago, and stratified into 343 neighborhood clusters previously defined by the Project on Human Development in Chicago Neighborhoods (Sampson et al., 1997). CCAHS was specifically designed to examine the effects of neighborhoods on health and observational data were collected on the block around each sampled residence through the method of systematic social observation. Corresponding with each face-to-face interview, survey raters completed

a standardized instrument for rating the block where the respondent lived. On the cover page of the instrument is a diagram of a typical city block on which the rater fills in the names of the streets s/he is coding (Fig. 1). Each side of one of these streets is referred to as a block face, and a typical city block contains eight block faces. Each rater walked around the entire block two times while recording observations—the first time walking along the “inside” block faces and the second along the “outside” block faces. Inter-rater reliability of this method was demonstrated using a subsample of 80 blocks in a pilot study conducted in 2001 where two raters made separate, independent observations of the same block at the same time. Observed agreement ranged from 0.78 to 1.00 ($\kappa=0.27$ –0.91). Agreement tended to be higher for objective indicators (e.g. presence of high-rise housing; $\kappa=0.84$) and lower for observations requiring a qualitative judgment (e.g. quality of street conditions; $\kappa=0.27$).

Using this standardized instrument, observational data were collected on multiple neighborhood characteristics that have been shown to be related to health (see Table 1), including land use (e.g. housing type, commercial, institutional, industrial), recreational facilities (e.g. parks, playgrounds), food environment (e.g. supermarkets, fast food, restaurants, liquor stores), neighborhood physical and social disorder (e.g. garbage, litter, broken glass, graffiti, signs advertising alcohol), as well as built environment characteristics (e.g. presence of trees, quality of street conditions). Some questions are asked at the level of the block face, meaning that the rater must code each side of the same street separately (e.g. presence of graffiti on buildings, signs or walls). Other questions were asked at the street level where one observation was made for the entire street (e.g. condition of the street). For our purposes we focus on characteristics at the street level, aggregating the block face characteristics up to the street level where necessary.

For comparison, we used an identical instrument on a subset of 60 of these residential blocks (244 streets) to conduct a virtual SSO using Google Earth. These blocks were selected from a random sample of all blocks in the study and were spatially distributed throughout the city of Chicago (Fig. 2), with somewhat greater density on the north side of the city. Using the Street View images for the city of Chicago, a trained rater did a virtual walk around the block where respondents lived and documented observed characteristics using the identical standardized SSO instrument. Google Street View images for the city of Chicago were dated around 2007 (about four to five years after the in-person SSO data were collected).

2.2. Analyses

We examine the inter-source reliability of street-level characteristics observed in the virtual compared to the in-person neighborhood audit. Agreement between observed characteristics using the in-person SSO and the virtual SSO was assessed using the Kappa coefficient (Cohen, 1960). The Kappa statistic adjusts for the amount of agreement that could be expected to occur by chance alone (Landis and Koch, 1977), and ranges from 1.0 (representing perfect agreement) to 0 (representing agreement corresponding to that expected by chance). However, due to the sensitivity of the Kappa statistic to the underlying prevalence of the characteristic (Feinstein and Cicchetti, 1990), we also report the observed agreement between the in-person and virtual SSO data. All analyses were conducted in SAS Version 9.2 for Windows.

3. Results

Observed agreement and Kappa statistics (with 95% confidence intervals) for the SSO data are presented in Table 1. Levels of

Systematic Social Observation Coding Sheet

Block ID # _____ Mode of Transportation: _____
 Observer Name: _____ Observer ID: _____
 Date: _____ Start Time: _____ am pm End Time: _____ am pm

VERY IMPORTANT!!!

On the diagram below, please circle the street numbers and write the street names to indicate the starting point of your observations. All of the block faces on the inside of the block will be coded “a” and all block faces on the outside will be coded “b.” If this diagram in no way resembles the block under observation, use the space at the bottom of the page to sketch a diagram, identifying streets with both a number and a name.

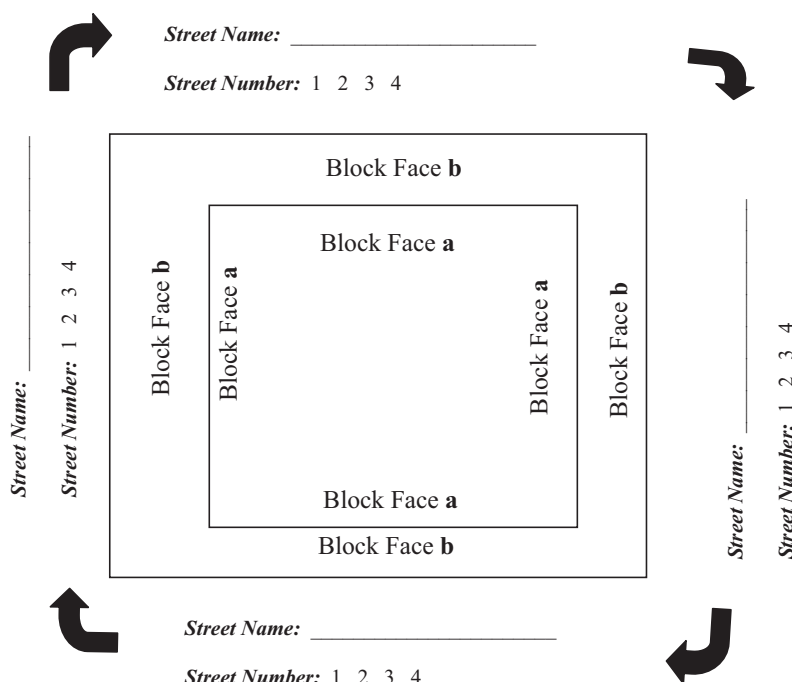


Fig. 1. Chicago Community Adult Health Study.

observed agreement for the presence of recreational facilities and characteristics of the local food environment were high (> 0.90), indicating a high reliability between these types of observational data collected in-person and using Google Street View. Corresponding Kappa coefficients tended to be lower ($\kappa=0.06$ – 0.57), especially for aspects of the environment observed less commonly in residential areas (e.g. supermarkets). Observed agreement for indicators of general and commercial land use ranged from 0.73 (low-rise private housing) to 0.99 (check cashing services), with lower Kappa statistics obtained for less prevalent characteristics such as drug stores or pharmacies ($\kappa=0.15$).

Similarly, indicators of the built environment and neighborhood social and physical disorder were assessed reliably using Google Street View, particularly for objectively observed conditions such as signs advertising alcohol (observed agreement= 0.92 , $\kappa=0.34$) or the presence of trees lining the street (observed agreement= 0.94 , $\kappa=0.49$). However, indicators requiring a finer level of observation (e.g. the presence of garbage, litter, or broken glass) were less reliably assessed using Google Earth (observed agreement= 0.35 ,

$\kappa=0.04$), as were those that were likely to have changed substantially over the five years between the in-person and virtual audit, such as the condition of streets and residential housing (observed agreement= 0.60 – 0.64 , $\kappa=0.03$ – 0.21).

Observed levels of agreement between characteristics collected using the in-person and virtual SSO instruments were comparable to the inter-rater reliability of the in-person audit conducted as part of the Chicago Community Adult Health Study (data not shown). For example, the inter-rater reliability for the presence of liquor stores on a street (observed agreement= 0.97 , $\kappa=0.36$) was similar to the level of agreement across the in-person and virtual audit instrument (observed agreement= 0.96 , $\kappa=0.38$).

4. Discussion

Compared to direct observational data collected as part of a face-to-face interview in the city of Chicago, we demonstrate in this case study that many neighborhood characteristics can be assessed

Table 1Street-level agreement in observed Chicago neighborhood characteristics ($N=244$ streets): systematic social observation in-person (2002) vs. Google Street View (2007).

	Observed agreement	Kappa	95% confidence interval
Recreational facilities			
Any recreational facilities (e.g. park, playground, sports fields)	0.923	0.499	(0.29, 0.70)
Any park	0.953	0.397	(0.11, 0.68)
Any playground	0.966	0.320	(−0.02, 0.66)
Any sports fields, playing fields, courts	0.970	0.573	(0.29, 0.86)
Food Environment			
Any convenience store	0.924	0.064	(−0.12, 0.25)
Any supermarket/grocery store	0.941	0.099	(−0.13, 0.32)
Any fast food/take out	0.924	0.268	(0.03, 0.51)
Any restaurant/other eating place	0.903	0.412	(0.22, 0.61)
Any bar/cocktail lounge	0.932	0.252	(0.01, 0.49)
Any liquor store	0.962	0.381	(0.06, 0.70)
General Land Use			
Any high-rise housing	0.973	0.713	(0.49, 0.93)
Any low-rise private housing	0.730	0.453	(0.35, 0.56)
Any detached single family houses	0.833	0.604	(0.49, 0.72)
Any commercial/industrial unit	0.775	0.475	(0.35, 0.60)
Any institutional land use (e.g. schools)	0.843	0.305	(0.13, 0.47)
Any church/religious center	0.907	0.449	(0.25, 0.64)
Any parking lots	0.818	0.360	(0.21, 0.51)
Commercial Land Use			
Any bank	0.962	0.288	(−0.04, 0.62)
Any check cashing service	0.987	0.394	(−0.15, 0.94)
Any drug store/pharmacy	0.958	0.145	(−0.15, 0.43)
Indicators of Neighborhood Social and Physical Disorder			
Any abandoned, burned out, or boarded up housing	0.924	0.147	(−0.07, 0.37)
Any garbage, litter, or broken glass in the street or on sidewalks	0.347	0.041	(0.01, 0.08)
Any vacant lots or open space	0.843	0.273	(0.11, 0.44)
Condition of residences (well kept vs. moderate/fair condition)	0.637	0.211	(0.10, 0.32)
No visible graffiti	0.797	0.095	(−0.05, 0.24)
Any second hand shop or pawn shop	0.927	0.160	(−0.07, 0.39)
Signs advertising alcohol	0.915	0.339	(0.12, 0.55)
Built Environment Characteristics			
Street condition (poor/fair vs. good)	0.598	0.032	(−0.10, 0.17)
Trees lining the street (trees on all, most or some of the street vs. none)	0.941	0.487	(0.25, 0.73)

reliably using a virtual audit instrument with the Street View feature of Google Earth. We found that the presence of recreational facilities and aspects of the local food environment are reliably captured using a virtual walk with a standardized instrument. Observed agreement in the presence of parks, playgrounds, and sports fields was over 92%, while observed agreement in characteristics of the local food environment (e.g. presence of fast food restaurants, bars, convenience stores) was over 90%.

While observed agreement in objectively rated characteristics was consistently high (>0.70), corresponding Kappa coefficients tended to be lower, likely due to the low prevalence of many of these characteristics in a small subset of an urban residential area (Feinstein and Cicchetti, 1990). In our data the overall prevalence of bars or signs advertising alcohol was around 3%, while the prevalence of abandoned buildings or graffiti was 10%. Because the expected chance agreement is inflated for these rare characteristics, the denominator of the Kappa statistic is minimized, resulting in a low Kappa value (Feinstein and Cicchetti, 1990). Nevertheless, for most characteristics the level of agreement was not due to chance (Kappa statistic significantly different from zero), and many of the Kappas indicated fair ($\kappa=0.20$ – 0.39) to moderate ($\kappa=0.40$ – 0.59) or substantial ($\kappa=0.60$ – 0.79) agreement between the two rating methods (Landis and Koch, 1977).

Consistent with other work on the reliability of in-person audit instruments (Clifton et al., 2007), agreement tended to be lower for characteristics requiring a qualitative judgment, such as the quality

of street conditions, and also for those requiring highly detailed observations at the street level (e.g. presence of garbage, litter, or broken glass) that may be less obvious using Google Street View images. Given the five year interval in our study between the in-person and virtual audit, reliability was also lower for more fluid neighborhood characteristics that are likely to change considerably over time (e.g. the presence of graffiti or the condition of residential housing). Given the general comparability between the observed agreement across the in-person and virtual audit and the inter-rater reliability of the in-person audit (0.78 – 1.00 ; $\kappa=0.27$ – 0.91), some of the variability in characteristics observed across modes of observation may in fact be due to inter-rater reliability or test-retest reliability over the five years between observations.

4.1. Limitations

Currently, coverage in Google Street View is not complete. While it tends to be more comprehensive in urban rather than rural areas, not all cities have Street View available for all streets (especially smaller streets). Moreover, ethical issues and controversies have been raised surrounding the use of Street View, and not all users have access to these data. In addition, the dates of the images in Google Street View are not always readily apparent. Using Google Street View for a virtual neighborhood audit is contingent upon a temporal alignment between the Street View images and the individual data to which researchers may

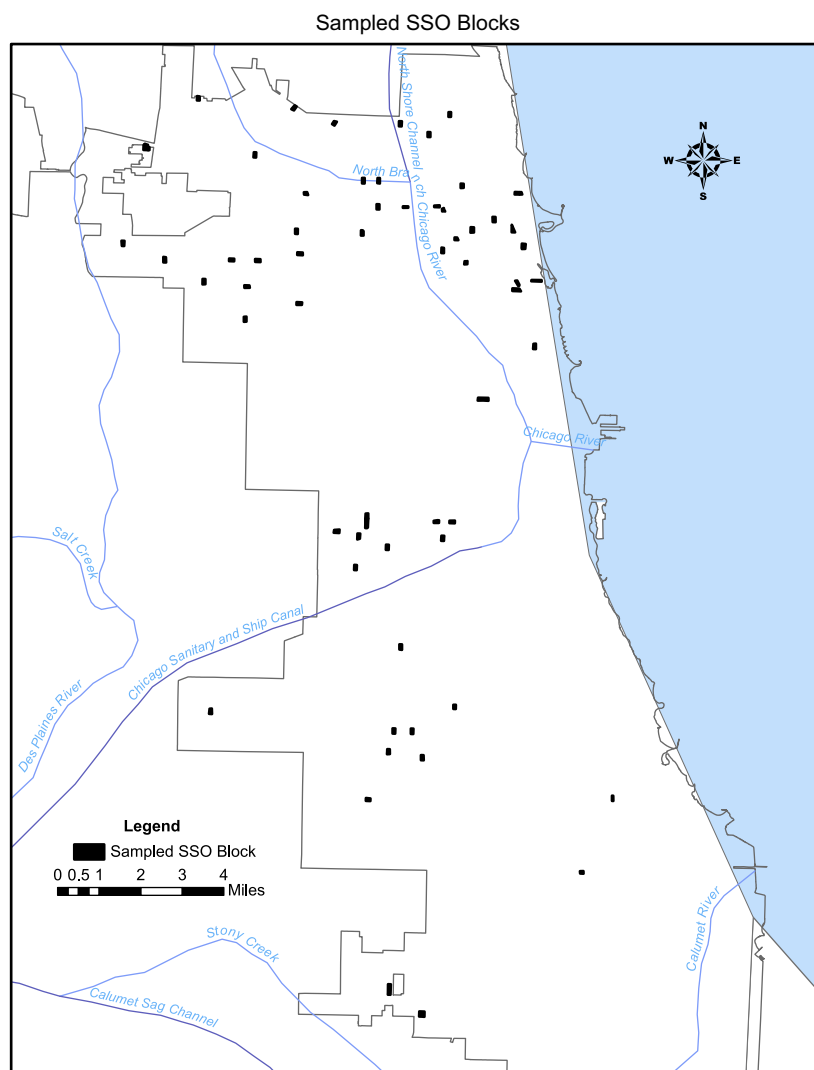


Fig. 2. Map Showing 60 Chicago neighborhood blocks (244 street segments).

wish to link them. Our study was limited by a five year lag between in-person and virtual assessments. Further studies are needed to replicate these analyses in other settings with more contemporaneous timing between the in-person and virtual audit. This line of research would also benefit from an assessment of inter-rater reliability in the virtual audit (using more than one rater for the Street View assessments).

However, our results indicate that a virtual audit instrument using Google Street View can provide reliable indicators of recreational facilities, the local food environment, and general land use at a fraction of the cost of an in-person neighborhood audit. Objective indicators of the built environment and neighborhood social and physical disorder are also reliably assessed. Caution should be exercised when conducting more qualitative observations (e.g. quality of street conditions or residential housing) or when trying to gather more finely detailed observations (e.g. garbage, litter, or broken glass) that benefit from direct observation in the field. Researchers should also be aware that strong agreement between measurements is not necessarily indicative of valid measurement. As for all data collection methods, rigorous and standardized training of raters is important for the quality of both the in-person and neighborhood audit. However, there are opportunities for considerable cost savings

with a virtual data collection instrument because raters are not required to travel to different locations to perform the neighborhood audit. The use of a virtual audit also allows researchers greater flexibility in the data collection phase. Similar to going back to a stored blood spot for biological markers on a respondent, it is possible to return to the Street View images at a later date (provided they have not been updated) if it becomes apparent that other aspects of the environment need to be documented.

Our study contributes to the growing literature on the development and evaluation of measures to capture community characteristics (Brownson et al., 2009; Sallis, 2009). We empirically demonstrate the reliability of using internet-based resources to conduct a neighborhood audit, providing evidence for researchers to consider during the design stages of a project when weighing issues such as the number of raters to employ or the number of city blocks to observe. Despite the widespread availability of visual data on community and built environments, few studies have utilized such electronic images on the internet to characterize neighborhood environments (Curtis et al., 2010; Doyle et al., 1998). Our hope is that future research will continue to examine the utility of the internet for conducting a neighborhood audit across other settings.

References

- Bader, M.D., Ailshire, J.A., Morenoff, J.D., House, J.S., 2010. Measurement of the local food environment: a comparison of existing data sources. *American Journal of Epidemiology* 171, 609–617.
- Boer, R., Zheng, Y., et al., 2007. Neighborhood design and walking trips in ten U.S. metropolitan areas. *American Journal of Preventive Medicine* 32, 298–304.
- Brownson, R.C., Hoehner, C.M., Day, K., Forsyth, A., Sallis, J.F., 2009. Measuring the built environment for physical activity: state of the science. *American Journal of Preventive Medicine* 36 (S99–S123), e12.
- Cervero, R., Duncan, M., 2003. Walking, bicycling, and urban landscapes: evidence from the San Francisco Bay area. *American Journal of Public Health* 93, 1478–1483.
- Clarke, P., Ailshire, J.A., et al., 2008. Mobility disability and the urban built environment. *American Journal of Epidemiology* 168 (5), 506–513.
- Clarke, P., George, L.K., 2005. The role of the built environment in the disablement process. *American Journal of Public Health* 95 (11), 1933–1939.
- Clifton, Kelly J., Livi Smith, A.D., Rodriguez, D., 2007. The development and testing of an audit for the pedestrian environment. *Landscape and Urban Planning* 80, 95–110.
- Cohen, J., 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 20, 37–46.
- Curtis, A., Duval-Diop, D., Novak, J., 2010. Identifying spatial patterns of recovery and abandonment in the post-Katrina Holy Cross neighborhood of New Orleans. *Cartography and Geographic Information Science* 37 (45), 12.
- Diez Roux, A.V., 2001. Investigating neighborhood and area effects on health: commentary. *American Journal of Public Health* 91 (11), 1783–1789.
- Diez Roux, A.V., 2004. The study of group-level factors in epidemiology: rethinking variables, study designs, and analytic approaches. *Epidemiologic Reviews* 26, 104–111.
- Diez Roux, A.V., Evenson, K.R., et al., 2007. Availability of recreational resources and physical activity in a sample of adults. *American Journal of Public Health* 97 (3), 493–499.
- Doyle, S., Dodge, M., Smith, A., 1998. The potential of web-based mapping and virtual reality technologies for modelling urban environments. *Computers, Environment and Urban Systems* 22, 137–155.
- Echeverría, S., Diez-Roux, A.V., Shea, S., Borrell, L.N., Jackson, S., 2008. Associations of neighborhood problems and neighborhood social cohesion with mental health and health behaviors: the multi-ethnic study of atherosclerosis. *Health and Place* 14, 853–865.
- Feinstein, A.R., Cicchetti, D.V., 1990. High agreement but low Kappa: I. The problems of two paradoxes. *Journal of Clinical Epidemiology* 43 (6), 543–549.
- Google Inc. (2005). Google Earth. Mountain View, CA.
- Landis, J.R., Koch, G.G., 1977. The measurement of observer agreement for categorical data. *Biometrics* 33, 159–174.
- Lopez, R., 2004. Urban sprawl and risk for being overweight or obese. *American Journal of Public Health* 94, 1574–1579.
- McGinn, A.P., Evenson, K.R., et al., 2007. The relationship between leisure, walking, and transportation activity with the natural environment. *Health and Place* 13, 588–602.
- Moore, L.V., Diez Roux, A.V., Nettleton, J.A., Jacobs Jr., D.R., 2008. Associations of the local food environment with diet quality—a comparison of assessments based on surveys and geographic information systems: the multi-ethnic study of atherosclerosis. *American Journal of Epidemiology* 167, 917–924.
- O'Campo, P., 2003. Invited commentary: advancing theory and methods for multilevel models of residential neighborhoods and health. *American Journal of Epidemiology* 157, 9–13.
- Pikora, T.J., Bull, F.C.L., et al., 2002. Developing a reliable audit instrument to measure the physical environment for physical activity. *American Journal of Preventive Medicine* 23, 187–194.
- Raudenbush, S.W., Sampson, R.J., 1999. Ecometrics: toward a science of assessing ecological settings, with application to the systematic social observation of neighborhoods. *Sociological Methodology* 29, 1–41.
- Reiss, A.J., 1971. Systematic observation of natural social phenomena. *Sociological Methodology* 3, 3–33.
- Sallis, J.F., 2009. Measuring physical activity environments: a brief history. *American Journal of Preventive Medicine* 36, S86–S92.
- Sampson, R.J., Morenoff, J.D., et al., 2002. Assessing neighborhood effects: Social processes and new directions in research. *Annual Review of Sociology* 28, 443–478.
- Sampson, R.J., Raudenbush, S.W., 1999. Systematic social observation of public spaces: a new look at disorder in urban neighborhoods. *American Journal of Sociology* 105, 603–651.
- Sampson, R.J., Raudenbush, S.W., et al., 1997. Neighborhoods and violent crime: a multilevel study of collective efficacy. *Science* 277, 918–924.
- Schaefer-McDaniel, N., O'Brien Caughy, M., O'Campo, P., Gearey, W., 2010. Examining methodological details of neighbourhood observation and the relationship to health: a literature review. *Social Science and Medicine* 70, 277–292.
- Tonne, C., Melly, S., et al., 2007. A case-control study of exposure to traffic and acute myocardial infarction. *Environmental Health Perspectives* 115, 53–57.