Quantifying spatial misclassification in exposure to noise complaints among low-income housing residents across New York City neighborhoods: a Global Positioning System (GPS) study

Dustin T Duncan, Kosuke Tamura, Seann D Regan, Jessica Athens, Brian Elbel, Julie Meline, Yazan A Al-Ajlouni, Basile Chaix

HHS Public Access
Author manuscript
Ann Epidemiol. Author manuscript; available in PMC 2018 January 01.

Published in final edited form as:
Ann Epidemiol. 2017 January ; 27(1): 67–75. doi:10.1016/j.annepidem.2016.09.017.

Quantifying Spatial Misclassification in Exposure to Noise Complaints Among Low-Income Housing Residents Across New York City Neighborhoods: A Global Positioning System (GPS) Study

Dustin T. Duncan, ScD[1], **Kosuke Tamura, PhD**[1], **Seann D. Regan, MA**[1], **Jessica Athens, PhD**[1], **Brian Elbel, PhD, MPH**[1,2], **Julie Meline, PhD**[3,4], **Yazan A. Al-Ajlouni**[1], and **Basile Chaix, PhD**[3,4]

[1] Department of Population Health, New York University School of Medicine, New York, NY USA

[2] Wagner Graduate School of Public Service, New York University, New York, NY USA

[3] Sorbonne Universités, UPMC Univ Paris 06, UMR_S 1136, Pierre Louis Institute of Epidemiology and Public Health, Paris, France

[4] Inserm, UMR_S 1136, Pierre Louis Institute of Epidemiology and Public Health, Paris, France

Abstract

Purpose—To examine if there was spatial misclassification in exposure to neighborhood noise complaints among a sample of low-income housing residents in New York City, comparing home-based spatial buffers and Global Positioning Systems (GPS) daily path buffers.

Methods—Data came from the community-based NYC Low-Income Housing, Neighborhoods and Health Study, where GPS tracking of the sample was conducted for a week (analytic n=102). We created a GPS daily path buffer (a buffering zone drawn around GPS tracks) of 200-meters and 400-meters. We also used home-based buffers of 200-meters and 400-meters. Using these "neighborhoods" (or exposure areas) we calculated neighborhood exposure to noisy events from 311 complaints data (analytic n=143,967). Friedman tests (to compare overall differences in neighborhood definitions) were applied.

Results—There were differences in neighborhood noise complaints according to the selected neighborhood definitions (p<0.05). For example, the mean neighborhood noise complaint count was 1196 per square kilometer for the 400-meter home-based and 812 per square kilometer for the 400-meter activity space buffer, illustrating how neighborhood definition influences the estimates of exposure to neighborhood noise complaints.

Corresponding author: Dustin T. Duncan, ScD, New York University School of Medicine, Department of Population Health, Spatial Epidemiology Lab, 227 East 30th Street, 6th Floor, Room 621, New York, NY 10016. Tel: 646-501-2674. Fax: 646-501-2706. Dustin.Duncan@nyumc.org.

Conclusions—These analyses suggest that, whenever appropriate, GPS neighborhood definitions can be used in spatial epidemiology research in spatially mobile populations to understand people's lived experience.

Keywords

spatial epidemiology; spatial misclassification; neighborhoods; geographic information systems; global positioning systems; low-income housing residents; noise complaint exposure

Introduction

In spatial epidemiology, most studies rely on crude neighborhood definitions (i.e. geographically-defined administrative boundaries and spatial buffers around a geographic location), which can result in spatial misclassification (i.e. incorrectly characterizing an environmental exposure) [1]. These administrative boundaries neighborhood definitions seem to be applied in studies with little theoretical reasoning. Although administrative boundaries (including ZIP codes and census tracts) are not used as much nowadays in spatial epidemiology research spatial buffers are still quite common, perhaps due to increased Geographic Information System (GIS) capacity. These spatial buffers are static and egocentric, meaning that they are fixed (not dynamic) and focused on a single location, which is a major limitation. Additionally, although most research has focused solely on residential neighborhoods [2], emerging research demonstrates that people are exposed to multiple (e.g. residential, work, social) neighborhood environments in their daily lives (termed “spatial polygamy”) [3-11]. While (some) research has focused on other salient neighborhood contexts (e.g. school neighborhoods for children and work neighborhoods for adults) [2], studies rarely examined more than one neighborhood context in the same study. Consequently, the range of neighborhood contexts one experiences is often missed in spatial epidemiology research. Real-time geospatial methods, including the use of Global Positioning System (GPS) technology are the cutting-edge, best suited method that can overcome these limitations because they better capture neighborhood contexts corresponding to individual lived experiences (known as “activity space neighborhoods”) [12-15]. Thus, there is strong theoretical basis related to the measurement of activity spaces with GPS receivers as a basis for exposure assessment. Previous research though demonstrated that when home-based spatial buffers were compared with GPS-defined “activity space neighborhoods” they shared at most only 12% of the variance in the neighborhood characteristics studied [16]. This suggests that residential neighborhoods are a very poor proxy for people's daily neighborhood exposures because most people's day-to-day activities are conducted outside of their residential neighborhood. However, little research has empirically examined the extent of spatial misclassification including comparing home-based spatial buffers and GPS activity space buffers including daily path buffers, which reflects an individual's daily movement patterns.

The purpose of this methodological study was to examine if there was spatial misclassification in neighborhood noise complaint exposure among a sample of low-income housing residents in New York City, using GPS data we collected. We focus on neighborhood noise complaints, because neighborhood-level exposure to noisy events can be related to health (e.g. blood pressure [17-23] and sleep [24-27], and it might be an especially salient neighborhood exposure to low-income housing residents in urban epicenters such as New York City [28, 29]. Moreover, noisy events that socially perceived as disturbing might be highly relevant to neighborhood disorder, which could be relevant to the studied population.

Materials and Methods

Data used in this study come from the NYC Low-Income Housing, Neighborhoods and Health Study, a pilot study that demonstrates the feasibility of Global Positioning Systems (GPS) data in a sample of low-income housing residents, which is among the first GPS studies to be conducted among a sample of low-income adults. Study details were previously described in detail and are briefly summarized here [30, 31]. Recruitment of 120 low-income housing residents in New York City was conducted through community-based outreach, which included handing out flyers outside of public housing developments in four different New York City neighborhoods, as well as through flyers posted and circulated by community-based organizations that work with low-income individuals (especially public housing residents), flyers posted in community locations (e.g. local stores) and through word of mouth (social networks). Adults were considered eligible for participation in the study if they self-reported living in low-income housing (e.g. public housing) in New York City; were 18 years of age or older; could speak and read English; self-reported not being pregnant; self-reported no difficulty in walking or climbing stairs; and were willing to wear a GPS device (on their person; e.g. in their pocket) for one week. The vast majority (80%) of the participants reported living in public housing (versus other low-income housing) and all participants reported being low-income. These data were collected between June and July 2014. Informed consent was obtained from all participants prior to data collection. The New York University School of Medicine Institutional Review Board reviewed and approved the research protocol.

Neighborhood Noise Complaints

We used the density of noise complaints as our measure of neighborhood exposure to noisy events [32]. In 2010, New York City started a sampling platform (called "NYC 311") operated by the Department of Environmental Protection that residents in New York City can make a call to 311 to report a complaint in their neighborhoods. The types of complaints include home, lost and found, vehicles and parking, transportation, streets and sidewalks, public health and safety, and noise [32]. One of the top three classifications among all types of complaints in the 311 data is noise. In this study, noise complaints came from 311 data during 1/1/2014 – 12/31/2014 (n=145,067). We specifically selected the year as the time frame for the analysis because the participant data were collected in 2014 and because registered complaints could be one-off events, not representative of average conditions across time. We anticipated that our measure is stable and consistent across time because it is a long time period. The 311 noise complaint data contain a time stamp and location of a noise report (i.e., case), address, streets, city, borough, as well as a latitude and longitude of each incidence. Some noise reports were removed – in particular, 1,100 noise reports were deleted due to missing location information (i.e., longitude and latitude of reports). The total analytic noise report cases were 143,967. These noise data included different types of noise, including noise from specific sources, e.g. commercial, helicopter, house of worship, park, streets, and vehicle. The top ten noise complaints from the 311 data in 2014 accounted for approximately 93% of the total complaint reports (See Table 1). The top three noise complaints include loud music/party (37%), construction before/after hour (17%), and loud talking (13%). The 311 noise data can be considered to be pollution indicators for the location of noise incident from residents in New York City [33]. Previous research has used this noise complaint data [33, 34] and we defined neighborhood noise complaints as the density of noise complaints in the different neighborhood definitions (described below), and more specifically the per square kilometer density and kernel density approach. Kernel density estimation is a data smoothing method where inferences about the population are made based on a sample data in this context the amount of noise complaints exposure based on distance from point source.

Address Geocoding

Participants provided their residential address. We geocoded (converted addresses to coordinates) the address following procedures used in our previous work, including cleaning the addresses prior to geocoding which involved standardizing the spelling to the USPS format (e.g. changing “street” to “St”, “avenue” to “Ave”, and “circle” to “Cir”) [35, 36]. These addresses were geocoded in mid-August 2014, using ArcGIS. Addresses were matched using a minimum match score of 65, spelling sensitivity of 60, and side offset of 10 feet. The ArcGIS minimum match score required was 80. We then conducted interactive re-matching in ArcGIS, where addresses can be reviewed and corrected on a case-by-case basis as necessary and for addresses with a match score of >80 that had ties. In the final step, we used Google Earth Pro to geocode the addresses with match scores below 80 and those that ArcGIS (Environmental System Research Institute, Redlands, CA) were unable to geocode. For addresses that were not geocoded in Google Earth Pro, we cleaned them and geocoded them in ArcGIS and Google Earth Pro (following the procedures previously articulated).

Geographic Information System (GIS) Buffers

Studies have used different sizes and zones for home-based spatial buffers. In this study, we used polygon-based street-network derived residential buffers of 200-meters and 400-meters —as have been used in previous neighborhood noise research in New York City [29, 37] and because we thought large buildings can block out noise, making network buffers most appropriate. Polygon-based street network buffers follow the street network for the sizes selected and then are connected the outer buffer edges to form a polygon. We use street network buffers as opposed to Euclidian distance here because these are meant to be comparable to past research [16]. Of note, 200-meters and 400-meters equates approximately to a 1/8 and 1/4 mile, respectively, around a location. The mean for the 200- and 400-meter spatial buffers across participants (n=102) were 0.04 (SD=0.01) square kilometers and 0.20 (SD=0.04)) square kilometers, respectively. In this study sample, many participants spend >50% of time in their residential neighborhood (i.e. 400-meter network buffer) [38].

Global Positioning System (GPS) Protocol and GPS Data Cleaning

The GPS device was set to log in 30-second intervals for location prior to distribution. Consistent with other studies [16, 39-48], GPS tracking of the sample was conducted for a week. During the study orientation and baseline assessment, participants were instructed to place the small QStarz BT-Q1000XT GPS device (Qstarz International Co., Ltd., Taipei, Taiwan) on their belt (using the manufacturer-provided case) or in their pocket and to complete a travel diary [30, 39]. Participants were asked to wear the GPS devices at all times (except when sleeping, swimming or showering). Consisting of a series of checkboxes, the travel diary asked the participant, “Did you charge the GPS monitor today?” and “Did you carry the GPS monitor with you today?” and was meant to help the participant remember to charge the device and carry it with him or her throughout the week. Consequently, the diary include did not include information about specific activities done in a day or any other time-period. Of note, we asked participants to complete their travel diary at home and at nighttime. However, we did not know when or where participants completed their travel diary, which we did not collect and so the diary was not used to inform the GPS data cleaning process. The GPS device was given to participants in a large plastic zipper storage bag, which also contained a mini USB charging cord for the GPS device, a USB wall adapter for charging, a manufacturer-provided GPS belt holder (if requested), a pamphlet containing background information on GPS, and the travel diary. Upon completion of the one-week GPS protocol (i.e. carrying the unit for all journeys, charging the unit daily, and completing the travel diary), researchers went to easily accessible community locations (i.e. coffee shop, library) in the participant's neighborhood to obtain the GPS devices, which is in line with our community-based approach. Participants also returned to the project office to give back the GPS devices, depending on which option was most convenient for him or her. As demonstrated previously, participants were compliant following the study's GPS protocol [30, 39].

GPS data were downloaded from Qstarz GPS devices in .gpx format then stored on a secured server and converted into ESRI .shp files and transferred into a geodatabase for processing, further analysis, map creation, and storage. GPS data in 30 seconds epochs were processed using a script built in python and executed in ArcGIS which eliminated duplicate timestamps, dates outside the range of the study, and removed GPS data points that were isolated spatially as these were likely data errors and not characteristic of typical mobility. GPS data drift is typically more pronounced while GPS receivers are stationary and these issues may be exacerbated in urban environments, however in past testing the GPS drift ranges from between 10 to 30 meters, and while a significant issue these common issues with GPS data did greatly influence out activity space variable calculations.

Of 120 participants enrolled in the study, six participants had no GPS data either due to user error, battery issues, or failing to return the device. Due to mismatched data a further five participants survey data could not be linked to GPS data and were omitted, one participant successfully returned the GPS device but had insufficient data, one participant's ID was found to be a duplicate in post-processing. In addition, five participants were removed because they spent a majority of their week with the GPS device outside New York City, since these data may not reflect typical mobility patterns and we were concerned with typical daily mobility. These restrictions resulted in a final participation and completion rate of 85.8% (*n*=102).

GPS Buffers

There are various ways to define an activity space [16, 49, 50]. In brief, the different ways to define activity spaces make different assumptions about mobility and therefore draw different boundaries around the GPS points. In this study, we used the commonly-used daily path area (a buffering zone drawn around the GPS tracks), which is a method in behavioral geography research to understand where participants spend the majority of their time and exposure to environment [16, 49]. Consequently, the daily path area includes places where the participant actually goes. We created 200-meter and 400-meter GPS buffers in this study. We selected these buffer sizes to rely on comparable buffer radiuses around home and GPS points for the sake of comparability. Of note, we buffered all GPS points for the two GPS buffer sizes and dissolved these separate features into a single feature, or space to create an "activity space" for each participant. The activity space size for the GPS-based daily path buffers was expressed in square kilometers. The mean for the 200- and 400-meter activity buffers across participants (n=102) were 11.40 (SD=9.76) square kilometers and 18.39 (SD=14.20) square kilometers, respectively. GPS activity space buffers for daily paths were created using ArcGIS version 10 (ESRI, Redlands, CA).

Statistical Analysis

First, we computed descriptive statistics for the neighborhood noise complaints for each of the different neighborhood definitions (i.e., 200-meter and 400-meter home-based buffers, and the 200-meter and 400-meter GPS activity space buffers). We applied Friedman test to compare differences in neighborhood definitions in neighborhood noise complaints. Post hoc analysis for the Friedman's test was performed when the null hypothesis was rejected. This allowed us to discover which of the groups (i.e., neighborhood definitions) were responsible for the reason that the null hypothesis was rejected. Analyses were performed for our various neighborhood definitions and for the different density measures. For example, we assessed whether the mean neighborhood noise complaints counts for the 400-meter home-based buffer and for the 400-meter activity space buffer are different.

Results

Figure 1 shows the location of an individual participants address, the various home-based spatial buffers used in this study for that address, and the participants' daily path buffers. In this map we show a kernel density estimation of noise complaints across all neighborhoods to give context for the density of noise across the study area.

The mean difference between the 200-meter home-basd buffer and the GPS activity space buffer for the sampled low-income adults was 833 (SD = 1673), and the mean difference between the 400-meter home-based buffer and the GPS activity space buffer for the sampled low-income adults was 384 (SD = 764) for the count variables. Minimal median differences were found for the kernel density estimates across neighborhood definitions.

Table 2 shows descriptive statistics on noise complaints for the different neighborhood definitions. Differences were stark for the count variables. For example, the mean neighborhood noise complaint count was 1696 per square kilometer for the 200-meter home-based buffer and 863 per square kilometer for the 200-meter activity space buffer. The mean neighborhood noise complaint count was 1196 per square kilometer for the 400-meter home-based buffer and 812 per square kilometer for the 400-meter activity space buffer.

Table 3 shows the results from the overall Friedman test. Overall, the estimates from level of neighborhood noise complaints varied for each neighborhood definition (all $P \leq 0.005$). Models comparing all measures of noise complaints across neighborhood definitions detected 11 specific differences including statistically significant differences in noise event count for the 400-meter home-based and 400-meter activity space buffer. In models comparing the specific measures of noise complaints (i.e., count and kernel density) differences remained including statistically significant differences in kernel density estimated noise for the 200-meter home-based and 200-meter activity space buffer.

We also found significant differences in neighborhood definitions studied when we completed the Friedman tests for each noise complaint metric separately: count and kernel density (data not shown).

Discussion

In this study, we examined if there was spatial misclassification among neighborhood noise complaint exposure among a sample of low-income housing residents in New York City. In particular, this study examined spatial misclassification comparing spaces defined by place of residence (home-based spatial buffers) to spaces defined by daily movement (GPS activity space buffers), which is a methodological novelty of this study. There were differences in metrics of neighborhood noise complaints according to the selected neighborhood definitions, illustrating how neighborhood definition influences the metrics of neighborhood noise complaints. However, the degree of spatial misclassification was less than we anticipated and less substantially than prior work as discussed below. This somewhat reduced level of spatial misclassification might be due to using two localized neighborhood definitions and in light of spatial autocorrelation. Overall, therefore, this study demonstrates that the neighborhood definition matters and if possible GPS buffers can be appropriate for spatially mobile populations. We also note that in some cases evaluating residential environment may still be useful, so home-based buffers can have utility. However, we argue that use of spatial buffers are likely less relevant when the research study needs to consider people's complex lived experience.

This study provides a meaningful contribution to the literature as very few studies have empirically examined spatial misclassification. In a prior study, though, we showed there was substantial spatial misclassification in youths' access to tobacco retailers using administrative neighborhood definitions and home-based spatial buffers [1]. One other study that we are aware of examined neighborhood noise and spatial misclassification [37]. In this study, the researchers used different neighborhood definitions, especially home-based buffers and administrative definitions, demonstrating neighborhood-levels of noise across different metrics. However, this study did not use GPS data. Even less research has examined spatial misclassification comparing home-based spatial buffers and activity spaces, especially GPS-based activity spaces such as daily path buffers. However, a recent study [51], that analyzed exposure assessment in aspects of neighborhood walkability using different neighborhood definitions (e.g. self-reported activity spaces and home-based spatial buffers) found differences across neighborhood definitions. In addition, another study found that when home-based spatial buffers were compared with GPS-defined activity space buffers they shared at most only 12% of the variance in the neighborhood characteristics studied [16]. This study examined environmental features such as fast food outlet density and park-land use and is most comparable to our study.

Future Research

Future research should continue to examine spatial misclassification including comparing home-based spatial buffers and GPS activity space buffers, including different variants of activity space buffers, across geographies and across exposures. Studies are needed in rural geographies and in varying geographic regions to see if these issues of spatial misclassification hold true to non-urban areas and differing regional contexts. The city of New York is a robust study environmental features, however it may be unique in regards to issues of spatial misclassification and future research should examine if the findings observed in the current study are comparable to other contexts, not only geographically but including varying sociodemographic groups. With this said, future research can furthermore examine other neighborhood factors (e.g. tobacco retailers, supermarkets) as well as they relate to spatial misclassification, including using larger samples.

In addition to evaluating issues of spatial misclassification in other regions and among varying populations, the current research opens the door for evaluating spatial mismatch as related to various health outcomes, including blood pressure, sleep, anxiety and substance use. While we hypothesize that noise in one's residential neighborhood could influence health, we also hypothesize that the activity spaces in addition to residential level noise would have more pronounced health effects. However, future research is needed to examine these hypotheses. From a mobility perspective, we do believe that activity spaces are the most salient neighborhood definition for mobile populations. We note that the exposure to noise to consider is likely dependent on the studied outcome(s).

Studies can collect GPS data via GPS-enabled smartphones for determining activity space neighborhoods, which could increase compliance to a GPS protocol, as smartphone are common nowadays. Although we believe current smartphones are of limited interest for GPS assessment due to limited battery life, they may be of great interest in the future, including because of their ability to combine GPS assessment and Global System for Mobile Communications (GSM) triangulation. Indeed it should be noted that future research utilizing GPS data might benefit from data triangulation techniques (utilization of data from cell towers) in conjunction with GPS to obtain even more realistic activity spaces. In addition to overall noise using GIS datasets, future studies can analyze different types of noise—which may have different impacts on health. Future studies can also examine other measures of noise exposure (not examined in the current study), e.g., annual average daily traffic noise and noise by time of day and by season. In the future research, the occurrence of the time and location can be adjusted in the analysis of noise. There are different ways to measure noise. In addition to self-report of noise, other sources lampposts of noise, or other more novel sources such as geolocated Twitter data about noise. Personal noise exposure could be measured as well via noise monitors—not just using GIS datasets of noise. Noise sound pressure data and noise complaint data should be compared in the future. Finally, future research should also examine individuals' mobility preferences, including as it relates to noise. For example, people may move because of noise level or other reasons. People could move to a more pleasing and quiet environment if they are noise sensitive or annoyed while out and about.

Study Strengths and Limitations

This study has a number of strengths. This is one of few studies empirically examining spatial misclassification, and one of very few studies empirically examining spatial misclassification using GPS data. In addition, we examined spatial misclassification using several different metrics, including density measures and kernel density measures of neighborhood-level exposures. Despite these strengths, this study is subject to some limitations.

Participants may have changed their spatial patterns given our distribution of GPS devices leading to potential reactivity bias and selective daily mobility bias. However, our past work suggests that these issues are minimal [30, 39]. Second, this study was conducted in a single geographic location among a relatively small non-probability sample of low-income housing residents. While 102 participants can be viewed as a relatively small sample size for spatial epidemiologic research, given that many recent GPS studies have fewer than 100 participants, our sample size of 102 is on par with the sample sizes of most GPS-based research. In addition, while our findings might only be generalizable to similar adult samples in similar urban areas, given that 20% of Americans now live in the 100 largest cities, and more than 70% of Americans living in urbanized areas with urbanization still on the rise, the relevance is large and growing [52]. We also note that GPS data accuracy in urban locations may suffer from error due to urban canyon effects and multipath reflectance [53]. While the urban environment, especially one as dense as New York City, presents issues for deriving subsequent data from GPS, such as mode of transportation, speed, and so on, these issues are less pronounced when examining data in aggregate as we have done here. Furthermore, this study relies on assumptions for the definition of exposure areas. In particular, the selected buffer size around the GPS points and residential locations could have influenced the findings. Because participants were tracked for only one week (which is currently the standard in spatial epidemiology literature), it is unknown whether spatial patterns were representative of ones' typical travel patterns and there may be seasonal variation in people's travel patterns (e.g. people may be more spatially monogamous in the winter months). Two weeks or even longer may represent someone's typical travel patterns in addition to GPS data over time. Furthermore, we were not able to consider in-door home noise exposure in the current study.

In addition, noise complaint data has certain limitations, which are important to note. More specifically, for example, the intensity of noise varies depending on the types of noise and the duration of noise complaints, ranging from noise from loud music/party and loud talking (e.g., a few hours) to noise from construction before/after hour (e.g., weeks). In this study, each noise complaint was counted as one case in the analysis and we did not disaggregate by noise type. The differences in noise type and duration could influence the magnitude of noise exposure for each participant. Additionally, depending on a location of noise complaint, whether noise occurs in mid-town New York, or a residential area in Bronx, for example, construction noise may diffuse differently. Another limitation is that the data are based on complaint reports from residents in New York City. Often complaints concentrated in the morning and evening since the residents are at home or come back from work. Thus, density of noise tends to sparse across city and time. Residents do not report ambient noise as complaints at a given time and location, since they are not present near the location and time of the occurrence. In contrast, even if there is no noise complaint at a certain time and location, this does not indicate noise do not exist. If an individual happens to live near a frequent caller to 311, this can bias his/her exposure. Residents simply may not be present at the time and location. Additionally, there may be differential reporting of noise (which is subjective) by certain characteristics including individual and neighborhood-level factors, such as age and socio-economic status [54]. In this study, we did not have resources to give noise monitors to participants. A complementary approach may be to replicate the present study using modeled noise (sound pressure) data. Finally, we assume that people are exposed to noise, however people can wear noise-canceling devices (e.g. listening to music on a smartphone) in their neighborhoods.

Conclusions

Researchers must be careful in selection of the neighborhood definition, in light of spatial misclassification. Like other exposures, many studies of neighborhood noise only use home-based spatial buffers to focus on the residential environment [17, 21-23]. Our analyses considering the exposure to reported noisy events suggest that, whenever possible, activity space neighborhood definitions can be used in spatial epidemiology research in spatially mobile populations to understand people's lived experience. The use of home-based neighborhood definitions can bias exposure estimates.

Acknowledgements

At the time of the study, Dr. Dustin Duncan was supported by his New York University School of Medicine Start-Up Research Fund to work on this project. The NYC Low-Income Housing, Neighborhoods and Health Study was supported by the NYU-HHC Clinical and Translational Science Institute (CTSI) Pilot Project Awards Program (Dr. Dustin Duncan, Principal Investigator). The NYU-HHC CTSI is supported in part by grant UL1TR000038 (Dr. Bruce Cronstein, Principal Investigator and Dr. Judith Hochman, co-Principal Investigator) from the National Center for Advancing Translational Sciences of the National Institutes of Health. We thank the research assistants for this project: Maliyhah Al-Bayan; Shilpa Dutta; William Goedel; Brittany Gozlan; Kenneth Pass; James Williams; and Abebayehu Yilma. We thank Jeff Blossom for geocoding the participants' addresses. We also thank the participants for engaging in this research. Finally, we thank Dr. Keith Goldfield for statistical advisement.

Glossary of Terms (Appendix)

Administrative Boundary

A boundary set by an administrative organization (e.g. police department, United States Postal Service, US Census Bureau) and examples include police districts, ZIP codes and census tracts

Daily Path Buffer

A type of activity space neighborhood based on GPS technology. This neighborhood calculates a buffering zone drawn around GPS tracks

GPS Defined Activity Space Buffer

A set of spatial locations visited by an individual over a given period, corresponding to his/her exhaustive spatial footprint; the regular activity space is the subset of locations regularly visited over that period. One way to define activity space buffers is via Global Positioning Systems (GPS) technology

Geographic Information System (GIS)

A system that is concerned with capturing, storing, analyzing and managing all types of spatial or geographical data

Global Positioning System (GPS)

A space-based navigation system that provides location and time information in all conditions

Spatial Buffer

Defines a neighborhood as a radius around a particular location, calculated using GIS technology. Many types of spatial buffers exist depending on the buffer size and what the location that the radius is set around represents (e.g. home, workplace, etc.). These spatial buffers are static (meaning that they are fixed and not dynamic) and egocentric (meaning that focused on a single location)

Spatial Misclassification

To classify an environmental exposure incorrectly

References

1. Duncan DT, et al. Examination of how neighborhood definition influences measurements of youths' access to tobacco retailers: a methodological note on spatial misclassification. American journal of epidemiology. 2014; 179(3):373–381. [PubMed: 24148710]
2. Leal C, Chaix B. The influence of geographic life environments on cardiometabolic risk factors: a systematic review, a methodological assessment and a research agenda. Obesity Reviews. 2011; 12(3):217–230. [PubMed: 20202135]
3. Matthews SA. Spatial polygamy and the heterogeneity of place: studying people and place via egocentric methods. Communities, neighborhoods, and health. Springer. 2011:35–55.
4. Koblin BA, et al. Methods to measure the impact of home, social, and sexual neighborhoods of urban gay, bisexual, and other men who have sex with men. PloS one. 2013; 8(10):e75878. [PubMed: 24146785]
5. Dragowski EA, et al. Social and sexual contexts explain sexual risk taking in young gay, bisexual, and other young men who have sex with men, ages 13–29 years. Journal of HIV/AIDS & Social Services. 2013; 12(2):236–255.

6. Egan JE, et al. Migration, neighborhoods, and networks: approaches to understanding how urban environmental conditions affect syndemic adverse health outcomes among gay, bisexual and other men who have sex with men. AIDS and Behavior. 2011; 15(1):35–50.

7. Sharp G, Denney JT, Kimbro RT. Multiple contexts of exposure: Activity spaces, residential neighborhoods, and self-rated health. Social Science & Medicine. 2015; 146:204–213. [PubMed: 26519605]

8. Duncan DT, Kapadia F, Halkitis PN. Examination of spatial polygamy among young gay, bisexual, and other men who have sex with men in New York City: the p18 cohort study. International journal of environmental research and public health. 2014; 11(9):8962–8983. [PubMed: 25170685]

9. Duncan, D. Quantifying Spatial Polygamy Among Low-Income Housing Residents in NYC: A GPS Study. 143rd APHA Annual Meeting and Exposition; October 31-November 4, 2015; APHA; 2015.

10. Tobin KE, Latkin CA, Curriero FC. An examination of places where African American men who have sex with men (MSM) use drugs/drink alcohol: A focus on social and spatial characteristics. International Journal of Drug Policy. 2014; 25(3):591–597. [PubMed: 24484732]

11. Tobin K, et al. Social geographies of African American men who have sex with men (MSM): A qualitative exploration of the social, spatial and temporal context of HIV risk in Baltimore, Maryland. Health & place. 2013; 22:1–6. [PubMed: 23518256]

12. Boruff BJ, Nathan A, Nijënstein S. Using GPS technology to (re)-examine operational definitions of ‘neighbourhood’in place-based health research. International journal of health geographics. 2012; 11(1):1. [PubMed: 22230476]

13. Chaix B, et al. GPS tracking in neighborhood and health studies: a step forward for environmental exposure assessment, a step backward for causal inference? Health & Place. 2013; 21:46–51. [PubMed: 23425661]

14. Chaix B, et al. An interactive mapping tool to assess individual mobility patterns in neighborhood studies. American journal of preventive medicine. 2012; 43(4):440–450. [PubMed: 22992364]

15. Chaix B, et al. Neighbourhoods in eco-epidemiologic research: delimiting personal exposure areas. A response to Riva, Gauvin, Apparicio and Brodeur. Social science & medicine. 2009; 69(9): 1306–1310. [PubMed: 19692161]

16. Zenk SN, et al. Activity space environment and dietary and physical activity behaviors: a pilot study. Health & place. 2011; 17(5):1150–1161. [PubMed: 21696995]

17. Katsouyanni K, et al. Hypertension and exposure to noise near airports: the HYENA study. Environmental health perspectives. 2008; 116(3):329. [PubMed: 18335099]

18. Melamed S, Fried Y, Froom P. The interactive effect of chronic exposure to noise and job complexity on changes in blood pressure and job satisfaction: a longitudinal study of industrial employees. Journal of occupational health psychology. 2001; 6(3):182. [PubMed: 11482631]

19. Andren L, et al. Effect of noise on blood pressure and ‘stress’ hormones. Clinical Science. 1982; 62(2):137–141. [PubMed: 7053913]

20. Abdelraziq IR, Ali-Shtayeh MS, Abdelraziq HR. Effects of Noise Pollution on Blood Pressure, Heart Rate and Hearing Threshold in School Children. Pakistan Journal of Applied Science. 2003; 3(10-12):717–723.

21. Méline J, et al. Road, rail, and air transportation noise in residential and workplace neighborhoods and blood pressure (RECORD Study). Noise and Health. 2015; 17(78):308. [PubMed: 26356373]

22. Babisch W, et al. Road traffic noise and hypertension–Accounting for the location of rooms. Environmental research. 2014; 133:380–387. [PubMed: 24952459]

23. Herbold M, Hense H-W, Keil U. Effects of road traffic noise on prevalence of hypertension in men: results of the Luebeck Blood Pressure Study. Sozial-und Präventivmedizin. 1989; 34(1):19–23. [PubMed: 2785319]

24. DeSantis AS, et al. Associations of neighborhood characteristics with sleep timing and quality: the Multi-Ethnic Study of Atherosclerosis. Sleep. 2013; 36(10):1543. [PubMed: 24082314]

25. Hume KI, Brink M, Basner M. Effects of environmental noise on sleep. Noise and Health. 2012; 14(61):297. [PubMed: 23257581]

26. Muzet A. Environmental noise, sleep and health. Sleep medicine reviews. 2007; 11(2):135–142. [PubMed: 17317241]

27. Omlin S, Bauer GF, Brink M. 2011; 13(53):299.

28. Hygiene NYCD. 2014
29. Kheirbek I, et al. Spatial variation in environmental noise and air pollution in New York City. Journal of Urban Health. 2014; 91(3):415–431. [PubMed: 24488652]
30. Duncan DT, et al. Application of global positioning system methods for the study of obesity and hypertension risk among low-income housing residents in New York City: a spatial feasibility study. Geospatial health. 2014; 9(1):57–70. [PubMed: 25545926]
31. Duncan DT, Regan SD. Mapping multi-day GPS data: a cartographic study in NYC. 2015:1–3.
32. Data NO. 2010
33. Zheng, Y., et al. Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing. ACM; 2014. Diagnosing New York city's noises with ubiquitous data..
34. Salamon J, Jacoby C, Bello JP. A dataset and taxonomy for urban sound research. 2014
35. Duncan DT, et al. Validation of Walk Score® for estimating neighborhood walkability: an analysis of four US metropolitan areas. International journal of environmental research and public health. 2011; 8(11):4160–4179. [PubMed: 22163200]
36. Duncan DT, et al. Evaluation of the positional difference between two common geocoding methods. Geospatial Health. 2011; 5(2):265–273. [PubMed: 21590677]
37. Tenailleau QM, et al. Assessing residential exposure to urban noise using environmental models: does the size of the local living neighborhood matter&quest. Journal of Exposure Science and Environmental Epidemiology. 2015; 25(1):89–96. [PubMed: 24866266]
38. Duncan DT, et al. Perceived spatial stigma, body mass index and blood pressure: a global positioning system study among low-income housing residents in New York City. Geospatial health. 2016; 11(2)
39. Duncan DT, et al. Feasibility and Acceptability of Global Positioning System (GPS) Methods to Study the Spatial Contexts of Substance Use and Sexual Risk Behaviors among Young Men Who Have Sex with Men in New York City: A P18 Cohort Sub-Study. PloS one. 2016; 11(2):e0147520. [PubMed: 26918766]
40. Klinker CD, et al. Using accelerometers and global positioning system devices to assess gender and age differences in children's school, transport, leisure and home based physical activity. Int J Behav Nutr Phys Act. 2014; 11(1):8. [PubMed: 24457029]
41. Dessing D, et al. Active transport between home and school assessed with GPS: a cross-sectional study among Dutch elementary school children. BMC public health. 2014; 14(1):227. [PubMed: 24597513]
42. Harrison F, et al. How well do modelled routes to school record the environments children are exposed to?: a cross-sectional comparison of GIS-modelled and GPS-measured routes to school. International journal of health geographics. 2014; 13(1):1. [PubMed: 24383521]
43. McCluskey A, et al. Feasibility and validity of a wearable GPS device for measuring outings after stroke. ISRN Rehabilitation. 2012 2012.
44. Clark RA, et al. A pilot investigation using global positioning systems into the outdoor activity of people with severe traumatic brain injury. Journal of neuroengineering and rehabilitation. 2014; 11(1):1. [PubMed: 24393611]
45. Wiehe SE, et al. Adolescent health-risk behavior and community disorder. PloS one. 2013; 8(11):e77667. [PubMed: 24278107]
46. Yen IH, et al. A pilot study using global positioning systems (GPS) devices and surveys to ascertain older adults' travel patterns. Journal of Applied Gerontology. 2013:0733464813479024.
47. Yan K, et al. Innovation through wearable sensors to collect real-life data among pediatric patients with cardiometabolic risk factors. International journal of pediatrics. 2014 2014.
48. Hurvitz PM, Moudon AV. 2012; 42(4):411–417.
49. Christian WJ. 2012; 3(4):287–295.
50. Hirsch JA, et al. Generating GPS activity spaces that shed light upon the mobility habits of older adults: a descriptive analysis. International journal of health geographics. 2014; 13(1):1. [PubMed: 24383521]

51. Perchoux C, et al. Residential buffer, perceived neighborhood, and individual activity space: New refinements in the definition of exposure areas–The RECORD Cohort Study. Health & place. 2016; 40:116–122. [PubMed: 27261634]
52. U.S. Census Bureau, P.D.. Annual Estimates of the Resident Population for Incorporated Places Over 50,000, Ranked by July 1, 2012 Population: April 1, 2010 to July 1, 2012. American Fact Finder; 2013.
53. Mooney SJ, et al. Quantifying Distance Overestimation From Global Positioning System in Urban Spaces. American journal of public health. 2016; 106(4):651–653. [PubMed: 26890178]
54. Whitfield A. Assessment of noise annoyance in three distinct communities living in close proximity to a UK regional airport. International journal of environmental health research. 2003; 13(4):361–372. [PubMed: 14594702]

Highlights

- The existence of misclassification in neighborhood noise exposure has been understudied.
- We compared noise complaints between neighborhoods defined using home-based buffers and daily path buffers.
- Differences in noise metrics were found based on the neighborhood definition.

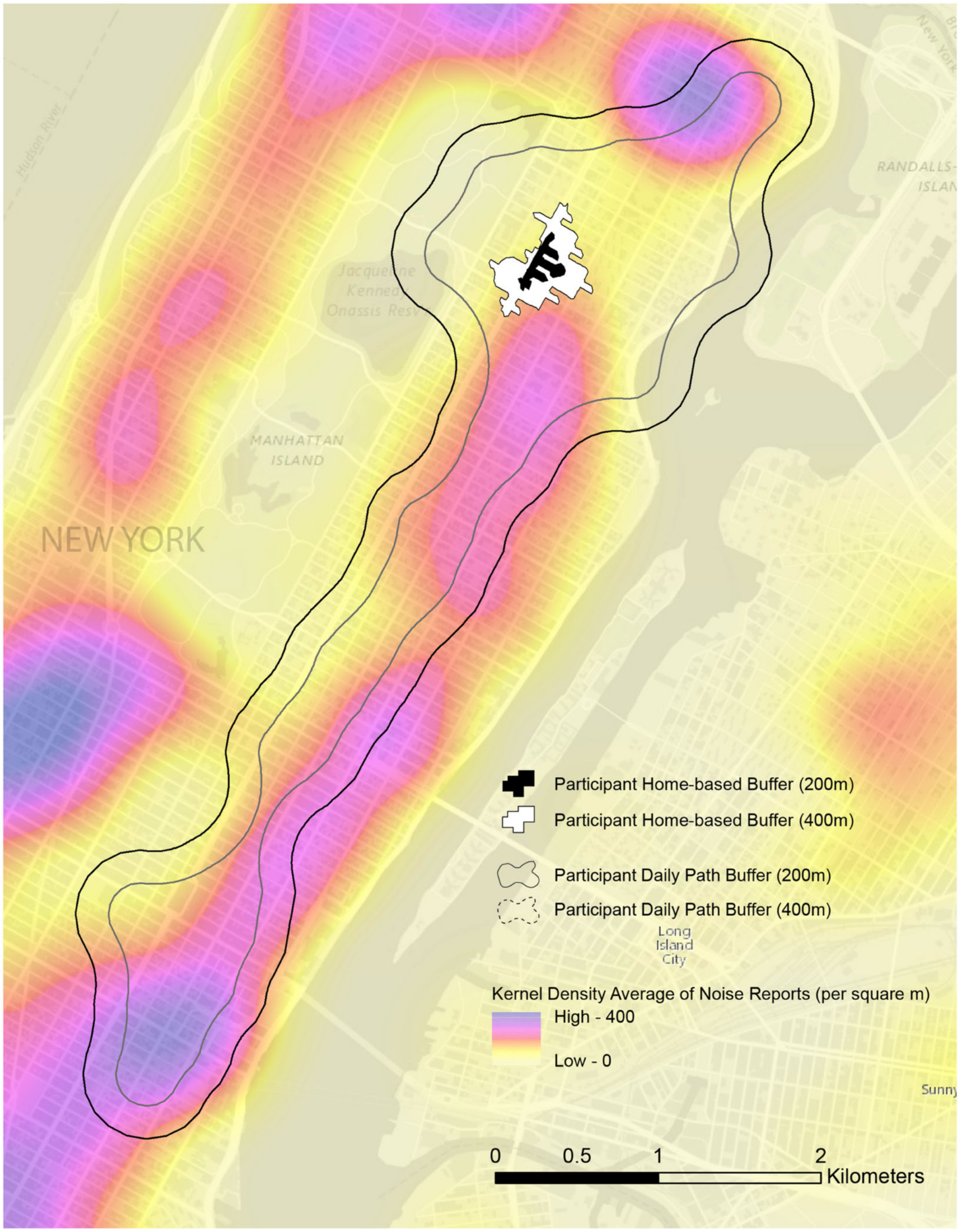

Figure 1.

Comparison of Home-based Buffers and GPS Daily Path Buffers, with Kernel Density Estimates of Neighborhood Noise Complaints.

Table 1

Description of geo-coded noise complaint type, count, and percentage (n=143,967)

Common complaint types	Count	Percent
1. Loud music/party	54,004	37
2. Construction before/after hours	24,293	17
3. Loud talking	18,696	13
4. Car/truck music	9,184	6
5. Barking dog	7,528	5
6. Construction equipment	6,036	4
7. Air condition/ventilation equipment	4,157	3
8. Engine idling	4,059	3
9. Banging/pounding	3,184	2
10. Alarms	3,048	2
11. Others	9,778	7

Table 2

Density of noise reports from the 311 data by neighborhood definition during the year of 2014 (*n*=143,967)

Neighborhood definition	No. noise report per square kilometers		
	Mean (SD)	**Median (IQR)**	**Range**
Count			
200m home-based buffer	1,696 (1,804)	920 (1,955)	5,108
200m activity space buffer	862 (384)	853 (631)	1,447
400m home-based buffer	1,196 (859)	837 (1,452)	2,994
400m activity space buffer	812 (332)	799 (508)	1,437
Kernel Density			
200m home-based buffer	838 (630)	669 (1334)	1,858
200m activity space buffer	835 (364)	826 (626)	1,405
400m home-based buffer	839 (638)	636 (1265)	1,843
400m activity space buffer	791 (318)	784 (484)	1,277

Note: SD= Standard Deviation, IQR= Interquartile Range

Table 3

Friedman test comparing neighborhood noise across neighborhood definitions, noise (count and kernel density)

Differences Tested	*P*-value
200m activity space (kernel density) - 200m activity space (count)	
200m activity space (kernel density) - 200m home-based (count)	
200m activity space (kernel density) - 400m activity space (count)	
200m activity space (kernel density) - 400m home-based (count)	*
200m home-based (count) - 200m activity space (count)	
200m home-based (count) - 400m activity space (count)	
200m home-based (kernel density) - 200m activity space (count)	§
200m home-based (kernel density) - 200m activity space (kernel density)	^
200m home-based (kernel density) - 200m home-based (count)	§
200m home-based (kernel density) - 400m activity space (count)	
200m home-based (kernel density) - 400m activity space (kernel density)	
200m home-based (kernel density) - 400m home-based (count)	§
400m activity space (count) - 200m activity space (count)	
400m activity space (kernel density) - 200m activity space (count)	§
400m activity space (kernel density) - 200m activity space (kernel density)	
400m activity space (kernel density) - 200m home-based (count)	§
400m activity space (kernel density) - 400m activity space (count)	
400m activity space (kernel density) - 400m home-based (count)	§
400m home-based (count) - 200m activity space (count)	
400m home-based (count) - 200m home-based (count)	
400m home-based (count) - 400m activity space (count)	**
400m home-based (kernel density) - 200m activity space (count)	**
400m home-based (kernel density) - 200m activity space (kernel density)	
400m home-based (kernel density) - 200m home-based (count)	**
400m home-based (kernel density) - 200m home-based (kernel density)	
400m home-based (kernel density) - 400m activity space (count)	
400m home-based (kernel density) - 400m activity space (kernel density)	
400m home-based (kernel density) - 400m home-based (count)	§

P-value for test of asymptotic general independence < 0.001

^ p. < 0.1

* p. < 0.05

** p. < 0.01

§ p. < 0.001