HHS Public Access

Author manuscript

Alcohol Clin Exp Res. Author manuscript; available in PMC 2020 August 01.

Published in final edited form as:

Alcohol Clin Exp Res. 2019 August ; 43(8): 1714–1726. doi:10.1111/acer.14119.

Methods for Evaluating the Association between Alcohol Outlet Density and Violent Crime

PAMELA J. TRANGENSTEIN, PHD MPH[1,2,*], **FRANK C. CURRIERO, PHD**[3], **JACKY M. JENNINGS, PHD**[4], **DANIEL WEBSTER, PHD**[5], **CARL LATKIN, PHD**[6], **RAIMEE H. ECK, PHD MPH, MPA**[7], **DAVID H. JERNIGAN, PHD**[2]

[1]Alcohol Research Group, Emeryville, USA

[2]Boston University School of Public Health, Department of Health Law, Policy and Management, Boston, USA

[3]Johns Hopkins Bloomberg School of Public Health, Department of Epidemiology, Baltimore, USA

[4]Johns Hopkins School of Medicine, Baltimore, USA

[5]Johns Hopkins Bloomberg School of Public Health, Department of Health Policy and Management, Baltimore, USA

[6]Johns Hopkins Bloomberg School of Public Health, Department of Health, Behavior and Society, Baltimore, USA

[7]National Cancer Institute, 9609 Medical Center Drive, Bethesda, MD 20892

Abstract

Background: The objective of this analysis is to compare measurement methods – counts, proximity, mean distance, and spatial access – of calculating alcohol outlet density and violent crime using data from Baltimore, Maryland.

Methods: Violent crime data (n=11,815) were obtained from the Baltimore City Police Department and included homicides, aggravated assaults, rapes, and robberies in 2016. We calculated alcohol outlet density and violent crime at the census block (CB) level (n=13,016). We then weighted these CB-level measures to the census tract level (n=197) and conducted a series of regressions. Negative binomial regression was used for count outcomes and linear regression for proximity and spatial access outcomes. Choropleth maps, partial R^2, Akaike's Information Criterion, and root mean squared error guided determination of which models yielded lower error and better fit.

Results: The inference depended on the measurement methods used. Eight models that used a count of alcohol outlets and/or violent crimes failed to detect an association between outlets and crime, and three other count-based models detected an association in the opposite direction. Proximity, mean distance, and spatial access methods consistently detected an association between outlets and crime and produced comparable model fits.

*Corresponding Author: 6001 Shellmound St., Suite 450, Emeryville, CA 94608, ptrangenstein@arg.org, (510) 898-5839.

The authors all declare no conflict of interest.

Conclusion: Proximity, mean distance, and spatial access methods yielded the best model fits and had the lowest levels of error in this urban setting. Spatial access methods may offer conceptual strengths over proximity and mean distance. Conflicting findings in the field may be in part due to error in the way that researchers measure alcohol outlet density.

Keywords

Alcohol; alcohol outlet density; violent crime; spatial access measures

Introduction

Social disorganization theory proposes that the structural components of neighborhoods may increase violent crime through their effect on informal control (Stark, 1987). In this theory, collective efficacy (i.e., “social cohesion among neighbors combined with their willingness to intervene on behalf of the common good”) mediates the association between neighborhood characteristics and violent crime (Sampson et al., 1997). This means that higher levels of neighborhood disorganization can inhibit informal social control (i.e., the ability of residents to realize and enforce shared goals in a way that regulates individuals’ behaviors), which in turn may fuel violent crime (Sampson et al., 1997). Supporting this theory, there is a rich literature connecting the spatial distribution of alcohol outlets with the level of violent crime (Scribner et al., 1999, Scribner et al., 2000, Zhang et al., 2015, Jennings et al., 2014, Gruenewald et al., 2006, Gorman et al., 2017, Franklin et al., 2010, Yu et al., 2008, Yu et al., 2009, Livingston, 2008b, Livingston, 2008a).

Researchers recently challenged the alcohol outlet density field to increase the rigor of their measurement methods (Gmel et al., 2016, Holmes et al., 2014), which sparked an interesting discussion about how to advance measurement in this area (Morrison et al., 2016, Fry et al., 2018, Lu et al., 2018). At a high level, all measurement methods aim to achieve a similar goal: to describe the spatial configuration of alcohol outlets. To help researchers identify the appropriate type of methods for alcohol outlet density research, the Centers for Disease Control and Prevention (CDC) recently published a guide on methods for measuring alcohol outlet density (Centers for Disease Control and Prevention, 2017b), referred to here as the CDC guide. However, providing evidence about the statistical performance of the available measurement techniques was beyond the scope of the guide.

The measurement methods described in the CDC guide are commonly used in alcohol outlet density research (Holmes et al., 2014) and more generally (Apparicio et al., 2008, Talen and Anselin, 1998, Handy and Niemeier, 1997). These measures can be broadly categorized as counts, proximity, mean distance, and spatial access. Counts are the simplest measure, and spatial access are the most complex. Counts capture the number of customers’ options (i.e., alcohol outlets) available in a given area (Handy and Niemeier, 1997). Counts weight all outlets equally, regardless of their location. Proximity measures (sometimes called “minimum distance” or “nearest neighbor”) are based on the minimum effort principle (Zipf, 1949), and they are calculated as the distance (or another metric of opportunity cost such as time) between a reference point and the closest outlet. Mean distance measures use descriptive statistics (e.g., mean, median, mode) to summarize the distance to a given set of outlets. Finally, spatial access measures (also called "gravity-based" measures) sum the inverse distances to a given set of outlets. By using inverse distance, spatial access measures assign higher weights to outlets that are closer to the reference point (Handy and Niemeier, 1997). Each of these methods are described in greater detail in the methods section.

It is unlikely that any one method will perform the best in every setting. However, the measurement method used can ultimately determine the inference of study results, which underscores the importance of understanding how different methods introduce or remove error in specific settings (Apparicio et al., 2008). As described in the methods section, there are four key considerations that can help researchers with this determination: 1) the unit of analysis, 2) the aggregation method, 3) the measure of accessibility, and 4) the type of distance (Apparicio et al., 2008, Handy and Niemeier, 1997).

Two studies provide initial evidence to help evaluate alcohol outlet density measurement methods in an urban setting (Seattle, WA) (Grubesic et al., 2016, Groff, 2014). Both studies conclude that spatial access methods may have statistical advantages over counting the number of outlets (Grubesic et al., 2016, Groff, 2014). Groff was the first to compare spatial access and count methods to alcohol outlet density (Groff, 2013). Using street block data and a container-based approach with three buffer sizes (800 feet, 1,200 feet, and 2,800 feet), she concluded that a basic spatial access measure (including only the sum of the inverse distances to outlets) explained the greatest amount of variation when compared to a count and a more advanced spatial access measure that added sales information as a proxy for outlet attractiveness (Groff, 2013). Later, Grubesic, Wei, Miller & Pridemore compared several gravity models (a spatial access method that measures the interaction between two objects, in this case alcohol outlets and census tracts) to count methods using kappa statistics to measure agreement. They concluded that spatial access methods were more sensitive than counts (Grubesic et al., 2016). While informative, these existing analyses only tested a subset of measurement methods; they did not test proximity, mean distance, or spatial access methods that use a "choice set" approach.

This study statistically compares count, proximity, mean distance, and spatial access methods for measuring the alcohol outlet density and violent crime in Baltimore City, Maryland. Our regression models test the association between alcohol outlet density and the level of violent crime after adjusting for alcohol outlet clusters, markers of social disorganization (i.e., drug arrests, vacant housing, percent Black, and median annual household income), and general neighborhood context (i.e., percent aged 18–35 years old, population density). We will assess error and model fit to evaluate the statistical performance of these measures in Baltimore City.

Materials and Methods

Background

Setting.—With a 2017 population of 614,000 residents, Baltimore is the largest city in Maryland and the 29th largest city in the United States (US Census Bureau, 2017). In 2016, Baltimore had 1,780 violent crimes per 100,000 residents, (United States Department of Justice, 2017).

Data Sources

Alcohol Outlets.—Liquor license information, including license type and address, was obtained for 1,218 alcohol outlets from the Board of Liquor License Commissioners for 2016. Liquor license information was current as of June 4, 2016. Fourteen (1.1%) license types with atypical restrictions on locations, days/hours of sales and types of products that may be sold were excluded, including arenas (n=7), municipal (n=5), Pimlico Race Track (n=1), and the Baltimore Zoo (n=1). The addresses for the remaining 1,204 outlets were geocoded using ArcGIS10.6.1.

Violent Crimes.—Victim-based violent crime incident data for 2016 were obtained from the Baltimore Police Department (BPD), including type of crime and location. Violent crimes included homicide, aggravated assault (including non-fatal shootings), rape, and robbery (Federal Bureau of Investigation, 2016). These crimes were selected because police reports of serious crimes such as these are reliable indicators of the real crime rate (Gove et al., 1985). In 2016, there were 11,815 violent crimes (318 homicides [2.7%], 5,711 aggravated assaults [48.3%], 285 rapes [2.4%], and 5,501 robberies [46.6%]). BPD publishes these data monthly and provide coordinates for each crime incident. BPD excludes crimes for which they were unable to geocode the incident location; the proportion of crimes that BPD was able to geocode is unknown.

Covariates.—We obtained most of our covariates from the 2016 American Community Survey (ACS) five-year estimates, which averaged data from 2012–2016. The ACS is an annual national survey that collects vital household information from nearly 2 million addresses each year (US Census Bureau, 2009).We also obtained 2016 drug arrest data from BPD and 2016 vacant housing locations from the Housing Authority of Baltimore via an online data sharing portal called OpenBaltimore.

Measures

This section outlines four considerations to explain the methods used in this study: 1) geographic unit, 2) method of aggregation, 3) type of accessibility measure, and 4) type of distance.

Geographic Units.—By definition, ecologic analyses like the present study compare populations or communities with a geographic unit of analysis. Sometimes called “containers,” these units can be either administrative/geopolitical boundaries (e.g., census tract [CT], ZIP Code) or user-defined areas (e.g., a buffer zone around a respondent’s house). The two primary concerns that guide determination of the most appropriate unit of analysis are aggregation bias (Hewko et al., 2002) and data availability. In urban settings like Baltimore City, accessibility can vary widely over short distances, and bias arises when measures average across heterogeneous areas (Hewko et al., 2002, Waller and Gotway, 2004). The potential for aggregation bias increases as the size of the geographic unit gets larger (Hewko et al., 2002). To minimize aggregation bias, we calculated alcohol outlet density, violent crime, drug arrests, and vacant housing at the census block (CB) level (n=13,488). ArcGIS was unable to calculate spatial measures for 472 CBs, resulting in an analytic sample of 13,016 CBs. Of the 472 excluded CBs, 393 did not contain any violent crimes, 256 had a population of zero, and 59 CBs did not contain roads.

ACS does not publish unrestricted data at the CB level, so we downloaded demographic covariates using a larger unit of analysis (CTs) and aggregated our CB-level measures to the CT level. To do this, we calculated area-weighted averages of CBs within CTs: $\sum_{1}^{n} \frac{A_i}{A_j} X_i$

where A_i is the area of CB_i in square miles, A_j is the area of CTj in square miles, X_i is the CB-level measure (e.g., alcohol outlet density) in CB_i, and n is the number of CBs in CT_j. We used CTs over census block groups (CBGs) for the unit of analysis for the ACS data, because there is a higher margin of error in ACS data at the CBG level. There are 200 CTs in Baltimore City, and 197 of these CTs had complete data and were able to be included in the analysis.

Methods of Aggregation.—There are two primary methods of aggregation: 1) counting numbers of outlets inside the unit of analysis and 2) measuring distance(s) to a set of outlets from reference points in each geographic unit (Apparicio et al., 2008, Handy and Niemeier, 1997). In this analysis, we calculate and compare four count-based variables and four distance-based variables (referred to as proximity, mean distance, and spatial access measures below).

Counts.: As the name implies, counts sum the number of outlets in a given area. Counts are often weighted by a measure of space or reach (e.g., population, square miles). These are among the easiest measures to calculate, do not require street-level data, are intuitive, and permit comparisons across communities (Centers for Disease Control and Prevention, 2017a). However, aggregating point-level data by counting the number of points make an implicit assumption that points (outlets) are uniformly distributed throughout the container. If outlets cluster together, this method will not be able to detect it (Centers for Disease Control and Prevention, 2017a).

In the present study, we summed the number of alcohol outlets and violent crimes located in each CB. For the alcohol outlet density measures, we created four count variables using different denominators: 1) no denominator, 2) population denominator, 3) area denominator (measured in square miles), and 4) roadway miles denominator. For violent crime, we calculated one count variable, which was the total number of violent crimes per CB.

Reference Points and Distances.: One common alternative to counting outlets combines reference points and distances to a set out outlets. Commonly, these measures use centroids as the reference points, and there are two main types of centroids: geometric and population-weighted. Geometric centroids are the middle of the geographic unit. They are straight-forward to calculate using geographic information systems, but can introduce bias if people don’t live or interact across the entire geographic unit (Hewko et al., 2002). Population-weighted centroids can reduce this bias, but they require information about the distribution of populations. It is also important to note that a limitation of these methods is that the results that researchers obtain using methods with reference points may depend on the type of reference point (e.g., centroid, mean center) used (Waller and Gotway, 2004).When using a design with reference points and distances, there is one additional design consideration that will determine the type(s) and level(s) of measurement error. This is whether a design is "container-based" or smoothed.

Container-Based Designs.: Container-based designs only measure accessibility within the unit of analysis. For example, counts are container-based designs, because they define an area in which to count. One limitation of container-based designs is that the results of analyses that use them will depend on the size and shape of the container used, which is a statistical challenge known as the "modifiable areal unit problem" (MAUP) (Waller and Gotway, 2004). Container-based methods are also prone to edge effects, which means that alcohol outlets across a container boundary may influence the level of violent crime inside the container but the measures wouldn't capture it because they cannot reach across the unit borders (Centers for Disease Control and Prevention, 2017a). Lastly, when using administrative containers, researchers should consider whether the containers are related to the association of interest. In particular, measuring variables using administrative boundaries that are unrelated to the association of interest can change or obscure patterns in the data (Talen and Anselin, 1998, Carlos et al., 2017). In addition to the container-based count variables, this analysis includes one additional container-based variable for both alcohol outlets and violent crime that summed the inverse distances to all outlets/crimes located within a 0.25-mile buffer from the CB centroid.

Smoothed Estimates.: The alternative to container-based design is a smoothed design. These measures are not confined to geographic units; rather, they pool information located within and/or beyond unit boundaries. The advantage of this approach is that drawing information from surrounding areas may help produce more stable estimates in sparse areas (Waller and Gotway, 2004). Spatial smoothing can also help avoid limitations of container-based designs, but it is not without its own drawbacks. Researchers must determine how many outlets/crimes to include for each unit of analysis, which can be subjective. This analysis includes three types of smoothed estimates: the distance to the nearest outlet/crime, the mean distance to the seven nearest outlets/crimes, and a spatial access measure that also uses distances to the seven nearest outlets/crimes (as described below).

Measures of Accessibility.—Only designs that combine distances with reference points need to consider measures of accessibility, because counting is both a method of aggregation and accessibility. This analysis compares three measures of accessibility: proximity, mean distance, and "gravity-based" spatial access methods (also called spatial accessibility indices [SAIs]).

Proximity.: Proximity methods measure the distance between a reference point and the closest outlet/crime, which is a smoothed design so the reference point and the outlet/crime do not need to be in the same geographic unit. One appeal of these measures is that they only require two data points. However, proximity measures cannot account for "spatial polygamy" or the cumulative influence of a cluster of outlets/crimes on a given location (Grubesic et al., 2016). In this analysis, we calculated one proximity variable for both alcohol outlets and violent crimes.

Mean Distance.: Mean distance measures are calculated as the arithmetic mean (though other summary measures like total, median, or mode are also possible) of the distances from a reference point to a set of outlets (Talen and Anselin, 1998, Handy and Niemeier, 1997). One appeal of these measures is that the unit is standard (e.g., miles, feet) and easy to understand. Mean distance methods may be container-based or smoothed.

In this study, we calculated one smoothed mean distance variable with a set of alcohol outlets/crimes (called a "choice set"). We defined the size of the choice set using literature about consumer decision making. Consumers consider seven plus or minus two options when making choices or evaluating settings (see Zhang, Lu, & Holt, 2011, for discussion). Thus, our mean distance variable calculated as the average distance from the CB centroid to the seven nearest outlets.

Spatial Access Measures.: Spatial access methods are derived from gravity-based models, which have been commonly used to determine locations for retail stores (Reilly, 1931) and understand population dynamics (Stewart, 1941). In general, gravity models follow this format: $\sum_{j=1}^{n} \frac{1}{d_{ij}^{\beta}}$ where d_{ij} is the distance between reference point i and j, and β is a friction parameter that summarizes how that distance decays over time/distance. The appeal of spatial access measures is that they use inverse distance. This is a strength because it weights outlets/crimes that are closer to the reference point more highly than those that are further away. In other words, inverse distances discount measures of alcohol outlet/violent crime accessibility for distance (Groff, 2013). However, spatial access measures that use inverse distance weighting cannot include coincident reference points and outlets/crimes, because the inverse of zero is undefined. One hundred and forty-four (144, 12.0%) outlets and eight crimes (8, <0.1%) were coincident with CBs centroids in our study, and we excluded these outlets/crimes from the spatial access measures.

Spatial access methods may be either container-based or smoothed. Container-based spatial access measures sum the inverse distances between a reference point and a set of alcohol outlets/crimes that fall inside a container (e.g., a 0.25-mile buffer around a CB centroid). In this study, we calculated a container-based spatial access measure for both alcohol outlets and crime that summed the inverse distances from the CB centroid to all outlets/crimes located inside a 0.25-mile buffer. In contrast, the smoothed spatial access design defined a set of alcohol outlets/crimes using the choice set approach described in the mean distance section, which summed of inverse distances from the CB centroid to the seven nearest outlets/crimes.

Type of Distance.—Finally, there are three types of distance to consider when calculating the distance between centroids and points: 1) Euclidian (straight-line or "as the crow flies"), 2) Shortest network (road-based) distance, and 3) Shortest network time. Based on Zipf's Principle of Least Effort (Zipf, 1949), we use the shortest network distance for our distance-based measures.

Control Variables.—Neighborhood contextual factors were included as covariates in the regression models, and were selected based on the social disorganization theory (Stark, 1987, Sampson et al., 1997). Demographic covariates included percent Black, median annual household income, population density, and percent of population aged 18–35 years, all of which came from the ACS. We adjusted for percent Black because Blacks tend to drink less than whites (Substance Abuse and Mental Health Administration, 2017), and the Baltimore population is predominately Black (US Census Bureau, 2017).

We also included three environmental covariates: alcohol outlet clusters, drug arrests, and vacant buildings. We adjusted for alcohol outlet clusters, because the association between alcohol outlet density and violent crime may differ in these pockets of high density. Seventy-nine percent of alcohol outlets were located less than 0.1 miles from the nearest alcohol outlet. Following the methods used in Zhang et al. (2015), we created and merged 0.1-mile buffers around each alcohol outlet and defined sets of overlapping buffers that included 50 or more alcohol outlets as high-density clusters. This approach classified 44 CTs (22.3%) as high-density areas. We created a binary variable to identify these high-density CTs.

We selected vacant housing and drug arrests as measures of disorganization, because Baltimore has high levels of vacant houses and drug use. For both drug arrests and vacant buildings, we summed the number of arrests/buildings per CB, constructed a weighted average to the CT level (following the same measures as those used for the alcohol outlet density variables), and performed a natural log transformation of the area-weighted counts.

Statistical Analysis

Regression was used to determine the association between different permutations of the alcohol outlet density and violent crime variables. These analyses were conducted using Stata version 14 (StataCorp, 2015). We used a natural log-transformation for the measures of alcohol outlet density to reduce the positive skew and to mitigate the influence of outliers. Count methods added 0.0001 to the variables before applying the natural log transformation because there were CBGs with no outlets. The rest of the statistical analyses depended on the way we measured violent crime, and a total of 32 models were tested. These models are summarized in Table 1. All regressions used a q-value estimated with the Simes-Benjamini-Hochberg correction for multiple testing (Newson, 2010). Akaike's Information Criterion (AIC) and the partial R^2 guided determination of which model yielded the best fit for the data, where smaller AICs and larger R^2 values indicated better fit. We also used a 10-fold cross-validation process to compare actual and predicted levels of crime using root mean squared error (RMSE).

Count Outcomes.—Negative binomial regression was used for models 1–8. Deviance goodness of fit analyses confirmed that Poisson regressions did not provide an adequate fit to the data. The negative binomial regressions used the natural log of the 2016 population as the offset.

Proximity & Spatial Access Outcomes.—Linear regression was used for models 9–32. Both the dependent and independent variables were log-transformed, so the regression coefficients can be interpreted as elasticities, which measure the change in a variable Y that is associated with a given change in variable X.

Spatial Analyses.—All spatial analyses were performed in R. Moran's Index (Moran's I) was calculated on the measures of violent crime and regression standardized residuals using a first order Queen adjacency matrix requiring at least two adjacent sides to determine spatial dependence. A Monte Carlo estimation process was used for the proximity and spatial access measurements. The unadjusted regression models should be approximately accurate, because the initial regressions accounted for more than 50% of the spatial dependence, the remaining residual spatial variation is small (Moran's I = 0.04–0.19, see Table 3), and the negative binomial regression accounts for overdispersion. We calculated spatial lags as the mean of that variable in the neighboring CBGs (Waller and Gotway, 2004). The lagged terms for the alcohol outlet density variables and covariates did not account for any additional spatial dependence, but many of these terms were significant, so they were included in the final models.

Collinearity between the covariates was not a problem, as all variance inflation factors were less than two (Sheather, 2009). We scaled the covariates to aid interpretation, so a one-unit increase represented a 10% increase in percent Black and percent aged 18–35 years, 100 houses for count of vacant housing, and $10,000 for median annual household income.

Results

On average, CBs were 0.01 square miles (range: <0.01–0.57 square miles) and contained 47 residents (range: 0–3,369 residents) (see Table 2). The percent of the residents in CTs that were Black had a bimodal distribution (mean 62.6%, range 0.4–99.5%), suggesting trends of racial segregation. The median annual household income was $46,744 (range: $12,279-$202,813). Baltimore CBs contained between 0 and 14 alcohol outlets with an average of less than one outlet in each CB. Many CBs had no or low counts of alcohol outlets.

Figure 1 shows the geographic distribution of the alcohol outlets and alcohol outlet density variables. Overall, the different measurement methods produced similar trends: alcohol outlet density was highest in the city center (downtown) and adjacent to the Inner Harbor (which is an entertainment area just south of downtown) as well as roughly two miles east and west from the city center. All methods captured the pockets of high density in east and west Baltimore, though these areas appeared smallest with the count variables and largest with mean distance and the SAI using a choice set. The starkest difference between the measurement methods is that counts produced maps that were more monochromatic than the other methods. This is visual evidence that counts classified almost all CBs as low density (light colors), whereas the other methods classified a minority of CBs (~20%) this way. The methods also conflicted in how they characterized the density in larger CBs, which is evident in downtown and southeast Baltimore. Counts using an area-based or roadway miles denominator had lower density levels downtown, which CBs are larger and have more businesses but fewer residences. Similarly, the raw counts and the counts with a population denominator categorized the large CBs in southeast Baltimore as high availability while the proximity and spatial access methods categorized them as low.

The inference depended on the method with which alcohol outlet density and violent crime were calculated (see Table 3). Twelve models concluded that there was no association between alcohol outlet density and violent crime, and all of these models used a count to measure at least one variable. None of the models that used a raw count (models 1, 9, 17, and 25) or a count divided by total roadway miles (models 4, 12, 20, and 28) found a significant association between alcohol outlet density and violent crime. Three models that measured either the dependent (model 5) or independent variable (models 18 and 26) found an association in the opposite direction -- that is that higher alcohol outlet density is associated with *lower* violent crime.

Seventeen models found an association between alcohol outlet density and violent crime in the expected direction. Fourteen (82.4%) of these models used an aggregation method that combined centroids and distances for the independent variable, and 15 (88.2%) used this type of method to measure the dependent variable.

Table 4 shows the model fit statistics, including partial R^2 for the alcohol outlet density variables as well as the RMSE and AIC values for each model fit. Measures of alcohol outlet density that used centroids and distances consistently explained the most variation in the outcomes. Among the models that used the count of violent crime as the outcome (models 1–8), the models with SAIs for alcohol outlet density explained the most variation in the outcome. However, among the models that used the proximity (models 9–16) or container-based SAI to measure violent crime (models 25–32), the proximity and mean distance variables for alcohol outlet density explained the most variation in the outcome.

Comparing the RMSE using cross-validation methods showed that the count-based models (models 1–8) had the highest absolute error between actual and expected outcomes. The RMSE was below 1 for all the other models. Within the models that used proximity, mean distance or spatial access methods to measure violent crime (models 9–32), proximity and mean distance tended to have the lowest RMSE values. However, the RMSE values for the SAIs are very similar. Finally, among models 1–8, model 3 with the count of alcohol outlets divided by area had the lowest absolute error.

We conducted a sensitivity analysis using the variables from model 23 (a choice set SAI for both alcohol outlet density and violent crime) to test statistical advantages of different choice set sizes. To do this, we fixed the choice set size for violent crime (seven crimes) and systematically varied the choice set size for the alcohol outlet density variable. Across these regressions, the AIC decreased by about one unit until the set included 25 outlets, where it stabilized.

Discussion

The measurement method determined whether or not models detected an association between alcohol outlets and violent crime. Some counts of alcohol outlets explained almost no variation in violent crime (R^2<0.1), and models that used counts were the least likely to detect an association. When they did, it was often in the opposite direction. Proximity and spatial access methods consistently detected the underlying association and yielded models with better fit. Accessibility methods that combine reference points and distances – proximity, mean distance, and spatial access methods – offer statistical advantages when quantifying the alcohol environment in an urban setting. Although their statistical performance was similar, spatial access methods also have conceptual strengths; they can detect clustering, they are more intuitive because density increases when spatial access measures increase, and they provide a gravity-based design that can integrate measures of outlet attraction (e.g., size, sales). For these reasons, we conclude that spatial access methods were most appropriate for this urban setting.

Counts mischaracterized large CBs because they were unable to detect the distribution of the outlets. These CBs often had several outlets located along the CB boundary. Counts treated these outlets as if they were distributed evenly across the CB, which led the method to conclude there was high availability when the majority of the CB had relatively low access. In addition, proximity methods showed evidence of random error in pockets of the city with greater numbers of outlets. This appeared as a peppering effect on the choropleth maps, and is likely the result of the simplicity of proximity methods. These methods only use the distance to one alcohol outlet, and determining the nearest outlet is a somewhat random process.

The two SAIs had different strengths and weaknesses. The SAI that used the choice set approach performed better in areas with low alcohol outlet density, while the container-based spatial access measure had advantages in high-density areas. Because it used spatial smoothing, the choice set SAI drew data from adjacent areas to measure density in these areas. The container-based SAI couldn’t reach beyond geographic boundaries, so it assigned zeroes to areas with no outlets/crimes within 0.25 miles. This created a variable with a bimodal distribution that had one peak for CBs with no alcohol outlets within the buffer and another peak for CBs that had at least one alcohol outlet within this range.

The container-based SAI had a wider range, which provided a more detailed summary of areas with high outlet/crime density. The choice set SAIs characterized some of high-density areas as having lower density than their container-based alternatives. This difference likely arises because the container-based approach accounts for both the number and the distribution of the alcohol outlets in the area, while the number is fixed in the choice set SAIs.

Prior studies have debated about how to best refine count methods through a reasoned choice of a denominator (Livingston, 2008a, Scribner et al., 1999, Yu et al., 2008, Romley et al., 2007, Hay et al., 2009, Badland et al., 2016, Kavanagh et al., 2011, Milam et al., 2013). This analysis compared the most common denominators for counts of alcohol outlet density. For the majority of models, the area-based denominator tended to have the highest R^2. This pattern was reversed in the AICs, suggesting that the area-based denominator performed the best in this setting. However, counts cannot detect clustering of alcohol outlets and systematically inflate the density in large CBs. The results from these analyses suggest that this debate misses the larger issue of how researchers can integrate measures of accessibility into their research if they have street-level data.

There are conflicting findings in the alcohol outlet density literature that have not yet been reconciled. For example, it is unclear whether on-premise (e.g., bars, restaurants) (Lipton and Gruenewald, 2002, Livingston, 2008b, Toomey et al., 2012, Gruenewald et al., 2006) or off-premise outlets (e.g., liquor stores) (Snowden, 2016, Livingston, 2008b, Branas et al., 2009, Pridemore and Grubesic, 2013, Gorman et al., 2005, Liang and Chikritzhs, 2011, Livingston, 2011) have stronger associations with the levels of harms. There is also disagreement about whether alcohol consumption mediates the association between alcohol outlet density and related harms, where some authors conclude there is (Scribner et al., 2008) and others conclude there is not (Iritani et al., 2013, Waller et al., 2013). Given that the majority of these studies use a count to measure alcohol outlet density (Holmes et al., 2014), measurement bias and error may contribute to these conflicts.

To date, the authors are unaware of any guidance for the optimal number of outlets to use to define a choice set for SAIs. We found that the SAIs with larger choice set sizes characterized high-density areas more accurately. While this provides statistical evidence of some benefit for larger numbers of observations in SAIs, this benefit may evaporate in the face of real-world conditions. The average CB had less than one outlet, which means a choice set of seven approximately averaged across seven CBs and a choice set of 25 outlets smoothed across 25 CBs. This yielded a statistically stable estimate but measured a meso-level effect instead of a local, micro-level effect. This means researchers may want to make context-specific decisions that account for the number of alcohol outlets and the unit of analysis regarding choice set size.

This analysis has several limitations. First, it only assessed total alcohol outlet density and did not disaggregate by outlet type. While this facilitated comparisons, it is possible that the statistical advantages of the respective methods could depend on the types of outlet and different methods more accurately capture dynamics of subtypes of outlets. We were also unable to determine whether all alcohol outlets were still open at the time of the analysis, as it is possible that some outlets closed in the 16 months between data generation and analysis. Also, the BPD data only include crimes that were reported to the police, so it is possible that there is underreporting. Future research may want to consider using population-weighted centroids, as this may capture access more accurately (Waller and Gotway, 2004). Lastly, these analyses began by defining Baltimore as a container, and therefore may suffer from edge effects. It is possible that the level of violent crime in the CBGs located along the Baltimore City boundaries may be associated with the access to alcohol outlets located in Baltimore County, which is a separate jurisdiction that surrounds Baltimore City.

Finally, it is possible that the relative advantages and disadvantages of these methods depend on the context. Findings described here are specific to Baltimore, which has unique demographics and history. Baltimore's population has steadily fallen since its peak of 950,000 residents in the 1950s (Bureau of the Census, 1950, Bureau of the Census, 2018). During this population decline, the number of alcohol outlets remained fairly constant, leading Baltimore residents to have high exposure to alcohol outlets. Departure of numerous residents also led to large swaths of vacant homes, which contributed to social disorganization that has been exacerbated by active drug markets and high poverty rates. Baltimore is also a city comprised of a patchwork of neighborhoods, which can cause demographics to shift substantially across small geographic areas. Consequently, Baltimore has substantial health disparities, which may or may not map accurately to individual CBs; life expectancy differs by as much as 20 years across neighborhoods (Baltimore Neighborhood Indicators Alliance, ND). Detecting an association between alcohol outlet density and violent crime in a city with high crime, poverty, and a range of social issues suggests that the association between alcohol outlets and violence is robust. While researchers reached similar conclusions to those arrived at here using data from Seattle, Washington, (Grubesic et al., 2016), future research could be conducted to determine whether spatial access methods are superior analytically in cities with different demographics.

Measuring the number and location of alcohol outlets are critical for understanding and predicting the potential negative impact of those outlets on surrounding communities. Effective, evidence-based policy begins with accurate measurement of the alcohol environment. This paper confirms that the way researchers measure alcohol outlet density and related outcomes matters in order to accurately describe the relationship between alcohol outlets and associated harms. In particular, these study findings are consistent with Grubesic et al.'s findings that spatial access methods offer statistical advantages over alternatives. The advantages over count or proximity methods appear substantial for both measures of alcohol outlet density and violent crime, at least in this urban setting. SAIs with a choice set appear to be the most versatile tool, capturing variability in both high and low-density areas. However, proximity methods appear to be a reasonable alternative that are easier to calculate and may offer statistical advantages in sparse areas. Similarly, container-based spatial access measures may be more accurate in dense areas. In the end, the findings from this study may provide additional support for researchers' decisions about which methods to use when characterizing the number and locations of alcohol outlets in other jurisdictions.

Acknowledgements

Funding

This report was supported by Cooperative Agreement Number 5U48DP005045 from the Centers for Disease Control and Prevention. Its contents are solely the responsibility of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention or the Department of Health and Human Services.

The project described was supported by Award Numbers T32AA007240, Graduate Research Training in Alcohol Problems: Alcohol-related Disparities and P50AA005595, Epidemiology of Alcohol Problems: Alcohol-Related Disparities from the National Institute on Alcohol Abuse and Alcoholism. The content is solely the responsibility of the authors and does not necessarily represent the official view of the National Institute on Alcohol Abuse and Alcoholism or the National Institutes of Health.

References

APPARICIO P, ABDELMAJID M, RIVA M & SHEARMUR R 2008 Comparing alternative approaches to measuring the geographical accessibility of urban health services: Distance types and aggregation-error issues. International journal of health geographics, 7, 7. [PubMed: 18282284]

BADLAND H, MAVOA S, LIVINGSTON M, DAVID S & GILES-CORTI B 2016 Testing spatial measures of alcohol outlet density with self-rated health in the Australian context: Implications for policy and practice. Drug Alcohol Rev, 35, 298–306. [PubMed: 26424454]

BALTIMORE NEIGHBORHOOD INDICATORS ALLIANCE. ND Children and Family Health: Life Expectancy [Online] Baltimore, MD Available: http://bniajfi.org/indicator/Children%20and%20Family%20Health/ [Accessed].

BRANAS CC, ELLIOTT MR, RICHMOND TS, CULHANE DP & WIEBE DJ 2009 Alcohol consumption, alcohol outlets, and the risk of being assaulted with a gun. Alcoholism: Clinical and Experimental Research, 33, 906–915.

BUREAU OF THE CENSUS 1950 1950 Census of Population: Preliminary Counts Washington, DC: US Department of Commerce, Bureau of the Census.

BUREAU OF THE CENSUS. 2018 Quick Facts: Baltimore City, Maryland [Online] Washington, DC Available: https://www.census.gov/quickfacts/fact/table/baltimorecitymaryland,US/PST045217 [Accessed August 2, 2018].

CARLOS HA, GABRIELLI J & SARGENT JD 2017 Validation of commercial business lists as a proxy for licensed alcohol outlets. BMC public health, 17, 480. [PubMed: 28526005]

CENTERS FOR DISEASE CONTROL AND PREVENTION 2017a CDC guide for measuing alcohol outlet density Atlanta, GA: Centers for Disease Control and Surveillance.

CENTERS FOR DISEASE CONTROL AND PREVENTION 2017b Guide for Measuring Alcohol Outlet Density Atlanta, GA: Centers for Disease Control and Prevention, US Dept of Health and Human Services.

FEDERAL BUREAU OF INVESTIGATION. 2016 Violent Crime [Online]. Available: https://ucr.fbi.gov/crime-in-the-u.s/2015/crime-in-the-u.s.−2015/offenses-known-to-law-enforcement/violent-crime [Accessed February 21, 2017].

FRANKLIN FA II, T. A. L, WEBSTER DW & PAN WK 2010 Alcohol outlets and violent crime in Washington DC. Western Journal of Emergency Medicine, 11, 283. [PubMed: 20882151]

FRY R, ORFORD S, RODGERS S, MORGAN J & FONE D 2018 A best practice framework to measure spatial variation in alcohol availability. Environment and Planning B: Urban Analytics and City Science, 2399808318773761.

GMEL G, HOLMES J & STUDER J 2016 Are alcohol outlet densities strongly associated with alcohol-related outcomes? A critical review of recent evidence. Drug and alcohol review, 35, 40–54. [PubMed: 26120778]

GORMAN DM, GORMAN D, ZHU L, GORMAN D, ZHU L, HOREL S, GORMAN D, ZHU L & HOREL S 2005 Drug ‘hot-spots’, alcohol availability and violence. Drug and alcohol review, 24, 507–513. [PubMed: 16361207]

GORMAN DM, PONICKI WR, ZHENG Q, HAN D, GRUENEWALD PJ & GAIDUS AJ 2017 Violent crime redistribution in a city following a substantial increase in the number of off-sale alcohol outlets: A Bayesian analysis. Drug Alcohol Rev

GOVE WR, HUGHES M & GEERKEN M 1985 Are uniform crime reports a valid indicator of the index crimes? An affirmative answer with minor qualifications. Criminology, 23, 451–502.

GROFF ER 2013 Measuring a Place’s Exposure to Facilities Using Geoprocessing Models: An Illustration Using Drinking Places and Crime. In: LEITNER M (ed.) Crime Modeling and Mapping Using Geospatial Technologies New York, New York: Springer.

GROFF ER 2014 Quantifying the exposure of street segments to drinking places nearby. Journal of Quantitative Criminology, 30, 527–548.

GRUBESIC TH, WEI R, MURRAY AT & PRIDEMORE WA 2016 Comparative approaches for assessing access to alcohol outlets: exploring the utility of a gravity potential approach. Population health metrics, 14, 25. [PubMed: 27486385]

GRUENEWALD PJ, FREISTHLER B, REMER L, LASCALA EA & TRENO A 2006 Ecological models of alcohol outlets and violent assaults: crime potentials and geospatial analysis. Addiction, 101, 666–677. [PubMed: 16669900]

HANDY SL & NIEMEIER DA 1997 Measuring accessibility: an exploration of issues and alternatives. Environment and planning A, 29, 1175–1194.

HAY GC, WHIGHAM PA, KYPRI K & LANGLEY JD 2009 Neighbourhood deprivation and access to alcohol outlets: a national study. Health Place, 15, 1086–93. [PubMed: 19540790]

HEWKO J, SMOYER-TOMIC KE & HODGSON MJ 2002 Measuring neighbourhood spatial accessibility to urban amenities: does aggregation error matter? Environment and Planning A, 34, 1185–1206.

HOLMES J, GUO Y, MAHESWARAN R, NICHOLLS J, MEIER PS & BRENNAN A 2014 The impact of spatial and temporal availability of alcohol on its consumption and related harms: a critical review in the context of UK licensing policies. Drug and alcohol review, 33, 515–525. [PubMed: 25186193]

IRITANI BJ, WALLER MW, HALPERN CT, MORACCO KE, CHRIST SL & FLEWELLING RL 2013 Alcohol outlet density and young women's perpetration of violence toward male intimate partners. Journal of Family Violence, 28, 459–470. [PubMed: 23914050]

JENNINGS JM, MILAM AJ, GREINER A, FURR-HOLDEN CDM, CURRIERO FC & THORNTON RJ 2014 Neighborhood alcohol outlets and the association with violent crime in one Mid-Atlantic city: the implications for zoning policy. Journal of Urban Health, 91, 62–71. [PubMed: 24002723]

KAVANAGH AM, KELLY MT, KRNJACKI L, THORNTON L, JOLLEY D, SUBRAMANIAN SV, TURRELL G & BENTLEY RJ 2011 Access to alcohol outlets and harmful alcohol consumption: a multi-level study in Melbourne, Australia. Addiction, 106, 1772–9. [PubMed: 21615583]

LIANG W & CHIKRITZHS T 2011 Revealing the link between licensed outlets and violence: counting venues versus measuring alcohol availability. Drug Alcohol Rev, 30, 524–35. [PubMed: 21896075]

LIPTON R & GRUENEWALD P 2002 The spatial dynamics of violence and alcohol outlets. Journal of studies on alcohol, 63, 187–195. [PubMed: 12033695]

LIVINGSTON M 2008a Alcohol outlet density and assault: a spatial analysis. Addiction, 103, 619–628. [PubMed: 18339106]

LIVINGSTON M 2008b A longitudinal analysis of alcohol outlet density and assault. Alcohol Clin Exp Res, 32, 1074–9. [PubMed: 18445114]

LIVINGSTON M 2011 Alcohol outlet density and harm: comparing the impacts on violence and chronic harms. Drug Alcohol Rev, 30, 515–23. [PubMed: 21896074]

LU H, ZHANG X, HOLT JB, KANNY D & CROFT JB 2018 Quantifying spatial accessibility in public health practice and research: an application to on-premise alcohol outlets, United States, 2013. International journal of health geographics, 17, 23. [PubMed: 29945619]

MILAM AJ, FURR-HOLDEN CD, HARRELL P, IALONGO N & LEAF PJ 2013 Off-Premise Alcohol Outlets and Substance Use in Young and Emerging Adults. Subst Use Misuse

MORRISON C, CERDÁ M, GORMAN DM, GRUENEWALD PJ, MAIR CF, NAIMI TS, SCRIBNER R, STOCKWELL T, TOOMEY TL & WIECZOREK WF 2016 Commentary on G mel et al. (2015): Are alcohol outlet densities strongly associated with alcohol-related outcomes? A critical review of recent evidence. Drug and alcohol review, 35, 55–57. [PubMed: 26450352]

NEWSON RB 2010 Frequentist q-values for multiple-test procedures. Stata journal, 10, 568.

PRIDEMORE W & GRUBESIC T 2013 Alcohol outlets and community levels of interpersonal violence: spatial density, outlet type, and seriousness of assault. Journal of Research in Crime and Delinquency, 50, 132–159.

REILLY WJ 1931 The Law of Retail Gravitation New York: WJ Reilly. Inc.

ROMLEY JA, COHEN D, RINGEL J & STURM R 2007 Alcohol and environmental justice: the density of liquor stores and bars in urban neighborhoods in the United States. Journal of studies on alcohol and drugs, 68, 48–55. [PubMed: 17149517]

SAMPSON RJ, RAUDENBUSH SW & EARLS F 1997 Neighborhoods and violent crime: A multilevel study of collective efficacy. Science, 277, 918–924. [PubMed: 9252316]

SCRIBNER R, COHEN D, KAPLAN S & ALLEN SH 1999 Alcohol availability and homicide in New Orleans: conceptual considerations for small area analysis of the effect of alcohol outlet density. Journal of studies on alcohol, 60, 310–316. [PubMed: 10371257]

SCRIBNER R, MASON K, THEALL K, SIMONSEN N, SCHNEIDER SK, TOWVIM LG & DEJONG W 2008 The contextual role of alcohol outlet density in college drinking. Journal of Studies on Alcohol and Drugs, 69, 112–120. [PubMed: 18080071]

SCRIBNER RA, COHEN DA & FISHER W 2000 Evidence of a structural effect for alcohol outlet density: a multilevel analysis. Alcoholism: Clinical and Experimental Research, 24, 188–195.

SHEATHER S 2009 A modern approach to regression with R, New York, NY, Springer.

SNOWDEN AJ 2016 Alcohol Outlet Density and Intimate Partner Violence in a Nonmetropolitan College Town: Accounting for Neighborhood Characteristics and Alcohol Outlet Types. Violence Vict, 31, 111–23. [PubMed: 26646575]

STARK R 1987 Deviant places: A theory of the ecology of crime. Criminology, 25, 893–910.

STATACORP 2015 Stata Statistical Software: Release 14 College Station, TX: StataCorp LP.

STEWART JQ 1941 An inverse distance variation for certain social influences. Science, 93, 89–90. [PubMed: 17729640]

SUBSTANCE ABUSE AND MENTAL HEALTH ADMINISTRATION 2017 Results from the 2016 National Survey on Drug Use and Health: Detailed Tables Rockville, MD.

TALEN E & ANSELIN L 1998 Assessing spatial equity: an evaluation of measures of accessibility to public playgrounds. Environment and planning A, 30, 595–613.

TOOMEY TL, ERICKSON DJ, CARLIN BP, LENK KM, QUICK HS, JONES AM & HARWOOD EM 2012 The association between density of alcohol establishments and violent crime within urban neighborhoods. Alcoholism: Clinical and Experimental Research, 36, 1468–1473.

UNITED STATES DEPARTMENT OF JUSTICE, F. B. O. I. 2017 2016 Crime in the U.S

US CENSUS BUREAU 2009 A compass for understanding and using American Community Survey data: What PUMS data users need to know Washington, DC.

US CENSUS BUREAU. 2017 Quick Facts: Baltimore City, Maryland [Online]. Available: https://www.census.gov/quickfacts/fact/table/baltimorecitymarylandcounty/HSG445215 [Accessed June 23, 2017].

WALLER L & GOTWAY C 2004 Applied spatial statistics for public health data, Hoboken, NJ, John Wiley & Sons, Inc.

WALLER MW, IRITANI BJ, CHRIST SL, HALPERN CT, MORACCO KE & FLEWELLING RL 2013 Perpetration of intimate partner violence by young adult males: the association with alcohol outlet density and drinking behavior. Health & place, 21, 10–19. [PubMed: 23395919]

YU Q, LI B & SCRIBNER RA 2009 Hierarchical additive modeling of nonlinear association with spatial correlations--an application to relate alcohol outlet density and neighborhood assault rates. Stat Med, 28, 1896–912. [PubMed: 19402025]

YU Q, SCRIBNER R, CARLIN B, THEALL K, SIMONSEN N, GHOSH-DASTIDAR B, COHEN D & MASON K 2008 Multilevel spatio-temporal dual changepoint models for relating alcohol outlet destruction and changes in neighbourhood rates of assaultive violence. Geospatial health, 2, 161. [PubMed: 18686265]

ZHANG X, HATCHER B, CLARKSON L, HOLT J, BAGCHI S, KANNY D & BREWER RD 2015 Changes in density of on-premises alcohol outlets and impact on violent crime, Atlanta, Georgia, 1997–2007. Prev Chronic Dis, 12, E84. [PubMed: 26020548]

ZIPF GK 1949 Human behavior and the principle of least effort

Author Manuscript
Author Manuscript
Author Manuscript

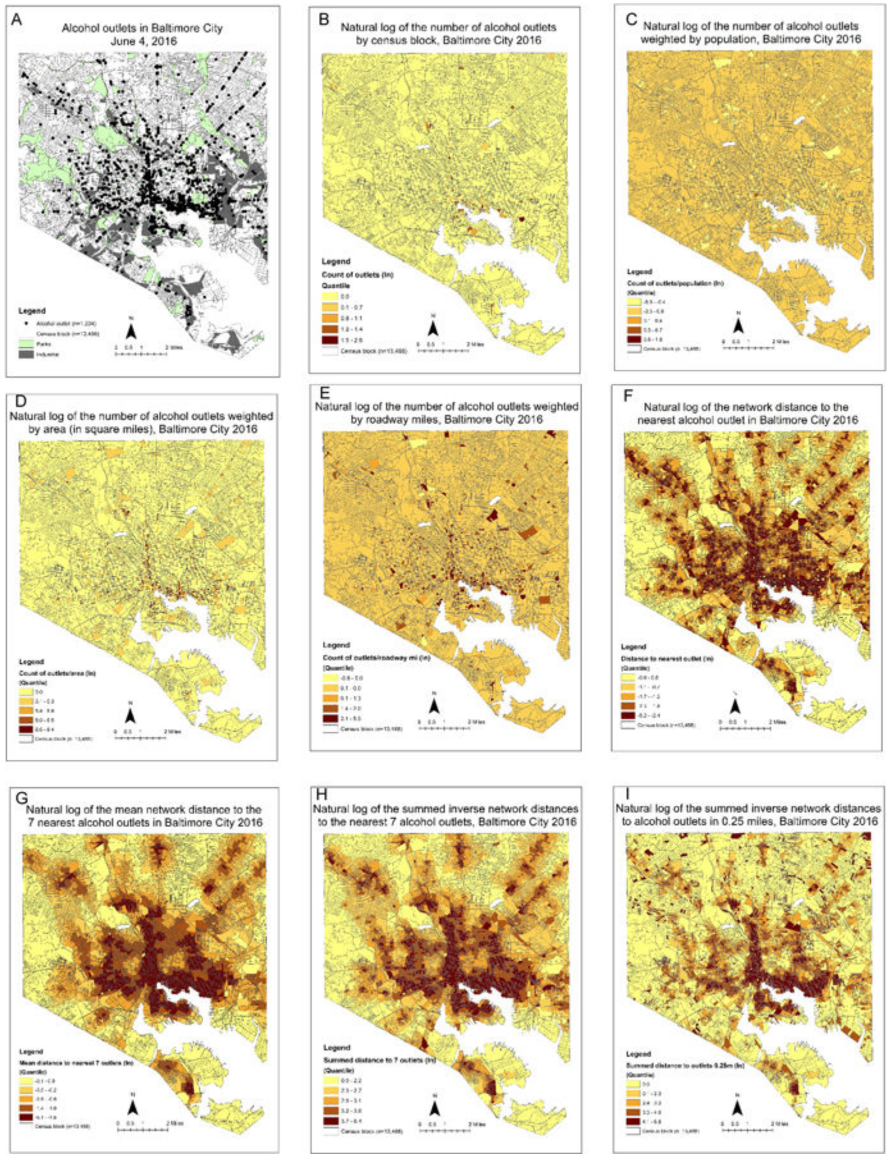

Figure 1.

Distribution of Alcohol Outlets and Alcohol outlet density Variables

Table 1.

Overview of the 32 Models, Listed by Model Number

Measure classification						Violent crime			
						Count	Proximity	Spatial access	
	Method of aggregation	Measure of accessibility	Subtype	Type of distance	Variable type	Raw count	Distance to nearest crime[a]	SAI with 7 nearest crimes[b]	SAI with 0.25-mile buffer[c]
Alcohol outlets	Count	Count	—	—	Raw count	#1	#9	#17	#25
			Weighted		Count weighted by population	#2	#10	#18	#26
					Count weighted by area	#3	#11	#19	#27
					Count weighted by roadway miles	#4	#12	#20	#28
	Distances & centroids	Proximity	—	Network	Distance to nearest outlet[d]	#5	#13	#21	#29
		Average distance	Choice set	Network	Mean distance to 7 nearest outlets[e]	#6	#14	#22	#30
		Spatial access		Network	SAI with 7 nearest outlets[f]	#7	#15	#23	#31
			Container-based	Network	SAI with 0.25-mile buffer[g]	#8	#16	#24	#32

SAI Spatial accessibility index

[a]Calculated as the minimum network (road-based) distance from the census block centroid to the closest violent crime.

[b]Calculated by summing the inverse network (road-based) distances from the census block centroid to each of the seven closest violent crimes, regardless of whether they are in the same census block as the centroid.

[c]Calculated by summing the inverse network (road-based) distance from the census block centroid to each violent crime located within a 0.25-mile buffer.

[d]Calculated as the minimum network (road-based) distance from the census block centroid to the closest alcohol outlet.

[e]Calculated as the average network (road-based) distance from the census block centroid to the seven closest alcohol outlets, regardless of whether they are in the same census block as the centroid.

[f]Calculated by summing the inverse network (road-based) distances from the census block centroid to each of the seven closest alcohol outlets, regardless of whether they are in the same census block as the centroid.

[g] Calculated by summing the inverse network (road-based) distance from the census block centroid to each alcohol outlet located within a 0.25-mile buffer.

Table 2.

Descriptive Characteristics of Baltimore City, 2016

Variable	Mean	SD	Min	Max
Census blocks (n=13,016)				
Total population in 2010	46.66	78.52	0.00	3,369.00
Total area (square miles)	0.01	0.01	0.00	0.57
Total roadway miles	0.30	0.30	0.00	5.25
Count of drug arrests	0.58	2.28	0.00	70.00
Count of vacant houses	1.40	4.38	0.00	102.00
Count of violent crime[b]	3.85	7.38	0.00	162.00
Minimum distance to violent crime[c]	0.11	0.17	0.00	1.78
SAI of violent crime – seven[d]	216.37	2,360.34	3.73	151,397.80
SAI of violent crime – 0.25 mi[e]	371.79	2,474.40	4.00	151,895.3
Count of alcohol outlets	0.09	0.44	0.00	14.00
Outlets per 1,000 residents	<0.01	0.08	0.00	6.00
Outlets per square mile	38.54	264.95	0.00	11,764.71
Outlets per roadway mile	0.52	3.81	0.00	247.87
Proximity (miles)[f]	0.30	0.27	0.00	2.30
Mean distance to seven nearest outlets[g]	0.52	0.36	0.01	2.51
SAI for seven nearest outlets[h]	26.59	27.30	2.78	597.57
SAI with 0.25-mile buffer[i]	225.19	47.90	0.00	895.44
Census tracts (n=197)				
Total population in 2016	3,121.74	1,367.94	616.00	6,962.00
Percent Black	62.62%	33.85%	0.37%	99.47%
Count vacant houses	2.89	3,29	0.00	19.42
Median annual household income	$46,744	$26,025	$12,279	$202,813
Percent of population aged 18–34 years	26.79%	11.73%	5.60%	70.50%

Variable	Mean	SD	Min	Max
Count of violent crime	7.16	5.21	0.25	30.98
Minimum distance to violent crime	0.09	0.07	0.02	0.56
SAI of violent crime – seven	249.99	704.47	7.73	8,601.03
SAI of violent crime – 0.25 mi	382.33	756.07	4.75	8,756.68
Count of alcohol outlets	0.13	0.20	0.00	1.31
Outlets per 1,000 residents	5.35	14.54	0.00	100.21
Outlets per square mile	27.92	51.21	0.00	368.92
Outlets per roadway mile	0.67	1.18	0.00	8.32
Proximity (miles)[f]	0.31	0.25	0.04	1.75
Mean distance to seven nearest outlets	0.52	0.33	0.10	2.04
SAI for seven nearest outlets	25.99	19.24	3.42	110.22
SAI with 0.25-mile buffer	20.82	34.53	0.00	216.44

SD=standard deviation; Min=minimum; Max=maximum; SAI=spatial accessibility index

[a]Median annual household incomes greater than $250,000 are censored.

[b]Violent crime incident data are from 2016 and include homicide, aggravated assault, rape, and robbery.

[c]Calculated as the minimum network (road-based) distance from the census group centroid to the closest violent crime.

[d]Calculated by summing the inverse network (road-based) distances from the census centroid to each of the seven closest violent crimes.

[e]Calculated by summing the inverse network (road-based) distance from the census centroid to each violent crime located within a 0.25-mile buffer.

[f]Calculated as the minimum network (road-based) distance from the census block centroid to the closest alcohol outlet.

[g]Calculated as the average network (road-based) distance from the census block centroid to the seven closest alcohol outlets.

[h]Calculated by summing the inverse network (road-based) distances from the census block centroid to each of the seven closest alcohol outlets.

[i]Calculated by summing the inverse network (road-based) distance from the census block centroid to each alcohol outlet located within a 0.25-mile buffer.

Table 3.

Results of Regression Analyses of Alcohol Outlet Density on Violent Crime Exposure, Baltimore City 2016

	Count of Violent Crimes[a]		Proximity to Nearest Violent Crime[b]		SAI for Seven Nearest Violent Crimes[c]		SAI for Violent Crimes in 0.25-Mile Buffer[d]	
	IRR	95% CI	β	95% CI	β	95% CI	β	95% CI
	Model #1		Model #9		Model #17		Model; #25	
Raw count of alcohol outlets	0.98	0.93, 1.03	>−0.01	−0.04, 0.04	−0.07	−0.15, 0.01	−0.08	−0.16, 0.01
Raw count of alcohol outlets in adjacent tract	0.96	0.87, 1.07	−0.01	−0.10, 0.08	0.07	−0.09, 0.22	0.09	−0.07, 0.25
Drug arrests[e]	1.30***	1.21, 1.40	***	−0.21, −0.07	0.23***	0.11, 0.35	0.34***	0.22, 0.47
Percent Black[f]	1.00	0.97, 1.04	−0.08***	−0.11, −0.05	0.10***	0.05, 0.15	0.13***	0.08, 0.19
Vacant housing[e]	1.06	0.99, 1.14	0.02	−0.04, 0.08	−0.04	−0.15, 0.06	−0.03	−0.15, 0.09
Median annual household income[g]	0.92***	0.88, 0.97	0.01	−0.03, 0.05	−0.04	−0.10, 0.03	−0.06	−0.13, 0.01
Percent population aged 18–35[f]	1.01*	0.82, 1.22	−0.14***	−0.21, −0.06	0.18**	0.05, 0.31	0.32***	0.18, 0.45
Population density	1.00	0.99, 1.00	0.07	−0.11, 0.24	0.25	−0.06, 0.56	0.27	−0.06, 0.60
High-density cluster[h]	1.86***	1.51, 2.30	−0.50***	−0.01, −0.30	0.79***	0.45, 1.13	1.11***	0.74, 1.47
Moran's I	0.15*		0.18***		0.10**		0.17***	
	Model #2		Model # 10		Model #18		Model #26	
Count of alcohol outlets divided by total population	0.98	0.96, 1.01	0.02	−0.01, 0.04	−0.06**	−0.09, −0.02	−0.07**	−0.11, −0.03
Drug arrests	0.95*	0.91, 0.99	0.06**	0.02, 0.10	−0.04	−0.11, 0.03	−0.06	−0.14, 0.01
Count of alcohol outlets divided by population for adjacent tracts	1.29***	1.20, 1.39	−0.13***	−0.20, −0.07	0.22***	0.10, 0.33	0.33***	0.21, 0.45
Percent Black	1.01	0.97, 1.04	−0.08***	−0.10, −0.05	0.10***	0.05, 0.15	0.13***	0.08, 0.18
Vacant housing	1.04	0.97, 1.11	0.04	−0.02, −0.07	−0.08	−0.18, 0.03	−0.07	−0.19, 0.04
Median annual household income	0.92***	0.89, 0.96	<0.01	−0.03, −0.04	−0.04	−0.11, 0.02	−0.06	−0.13, 0.01
Percent population aged 18–35	1.09*	1.01, 1.18	−0.12**	−0.19, −0.05	0.15*	0.02, 0.28	0.28***	0.14, 0.41
Population density	1.01	0.83, 1.22	0.05	−0.12, −0.22	0.28	−0.02, 0.58	0.31	−>0.01, 0.63

	Count of Violent Crimes[a]		Proximity to Nearest Violent Crime[b]		SAI for Seven Nearest Violent Crimes[c]		SAI for Violent Crimes in 0.25-Mile Buffer[d]	
	IRR	95% CI	β	95% CI	β	95% CI	β	95% CI
High-density cluster	1.79***	1.46, 2.19	−0.46***	−0.65, −0.27	0.74***	0.40, 1.07	1.04***	0.68, 1.39
Moran's I	0.14***		0.14***		0.08*		0.13**	
	Model #3		Model # 11		Model #19		Model #27	
Count of alcohol outlets divided by area	1.05	0.98, 1.12	−0.14***	−0.20, −0.08	0.30***	0.19, 0.40	0.34***	0.23, 0.44
Count of alcohol outlets divided by area for adjacent tracts	1.17**	1.05, 1.30	−0.07	−0.17, 0.02	0.04	−0.12, 0.21	0.16*	0.01, 0.32
Drug arrests	1.26***	1.17, 1.35	−0.10**	−0.16, −0.04	0.14*	0.03, 0.24	0.22***	0.12, 0.33
Percent Black	1.02	0.98, 1.05	−0.09***	−0.11, −0.06	0.12***	0.08, 0.17	0.16***	0.12, 0.21
Vacant housing	1.03	0.96, 1.10	0.06*	<0.01, 0.11	−0.12*	−0.22, −0.02	−0.13*	−0.23, −0.03
Median annual household income	0.92***	0.89, 0.96	0.01	−0.03, 0.04	−0.04	−0.10, 0.01	−0.07*	−0.12, −0.01
Percent population aged 18–35	1.03	0.95, 1.12	−0.05	−0.12, 0.02	0.02	−0.11, 0.14	0.10	−0.02, 0.22
Population density	1.03	0.86, 1.24	<0.01	−>0.01, <0.01	0.35*	−0.08, 0.63	0.41**	0.14, 0.68
High-density cluster	1.43**	1.14, 1.79	−0.25*	−0.44, −0.05	0.40*	0.06, 0.74	0.53**	0.20, 0.87
Moran's I	0.15***		0.16***		0.08*		0.10**	
	Model #4		Model #12		Model #20		Model #28	
Count of alcohol outlets divided by total roadway miles	1.03	0.97, 1.08	−0.04	−0.09, 0.01	−>0.01	−0.09, 0.09	0.02	−0.07, 0.12
Count of alcohol outlets divided by roadway miles for adjacent tracts	1.09	0.99, 1.22	−0.07	−0.17, 0.02	0.20*	0.03, 0.36	0.27**	0.10, 0.45
Drug arrests	1.28***	1.19, 1.38	−0.12***	−0.19, −0.06	0.20***	0.09, 0.32	0.31***	0.18, 0.43
Percent Black	1.01	0.98, 1.05	−0.09***	−0.12, −0.06	0.12***	0.07, 0.17	0.16***	0.10, 0.21
Vacant housing	1.05	0.98, 1.13	0.03	−0.03, 0.09	−0.06	−0.17, 0.04	−0.06	−0.17, 0.06
Median annual household income	0.92***	0.88, 0.96	0.01	−0.03, 0.04	−0.04	−0.10, 0.03	−0.06	−0.13, 0.01
Percent population aged 18–35	1.08	0.99, 1.18	−0.12**	−0.19, −0.05	0.14*	0.01, 0.27	0.27***	0.13, 0.41
Population density	1.00	0.82, 1.22	0.06	−0.11, 0.24	0.26	−0.05, 0.56	0.28	−0.05, 0.60
High-density cluster	1.69***	1.36, 2.11	−0.43***	−0.63, −0.23	0.69***	0.34, 1.04	0.96***	0.58, 1.33

	Count of Violent Crimes[a]		Proximity to Nearest Violent Crime[b]		SAI for Seven Nearest Violent Crimes[c]		SAI for Violent Crimes in 0.25-Mile Buffer[d]	
	IRR	**95% CI**	**β**	**95% CI**	**β**	**95% CI**	**β**	**95% CI**
Moran's I	0.15**		0.19***		0.10*		0.17**	
	Model #5		Model #13		Model #21		Model #29	
Proximity to nearest outlet[i]	0.84*	0.72, 0.99	0.54***	0.41, 0.68	−0.74***	−0.99, −0.49	−0.97***	−1.21, −0.73
Proximity to nearest outlet, adjacent tracts	0.78*	0.62, 0.98	−0.07	−0.25, 0.11	−0.01	−0.35, 0.32	−0.12	−0.44, 0.20
Drug arrests	1.24***	1.15, 1.33	−0.07*	−0.12, −0.01	0.11*	<0.01, 0.21	0.18***	0.08, 0.28
Percent Black	1.01	0.98, 1.04	−0.08***	−0.10, −0.05	0.10***	0.06, 0.15	0.13***	0.09, 0.18
Vacant housing	1.04	0.97, 1.11	0.04	−0.01, 0.09	−0.08	−0.18, 0.01	−0.08	−0.17, 0.01
Median annual household income	0.93***	0.89, 0.96	0.01	−0.02, 0.04	−0.04	−0.10, 0.02	−0.06*	−0.12, −0.01
Percent population aged 18–35	1.03	0.95, 1.12	−0.03	−0.09, 0.03	0.01	−0.11, 0.13	0.08	−0.03, 0.20
Population density	1.04	0.87, 1.25	−0.02	−0.16, 0.12	0.40**	0.13, 0.67	0.47***	0.22, 0.73
High-density cluster	1.40**	1.11, 1.76	−0.18*	−0.36, >−0.01	0.30	−0.05, 0.64	0.39*	0.07, 0.72
Moran's I	0.15***		0.18***		0.07		0.09*	
	Model #6		Model #14		Model #22		Model #30	
Mean distance to seven nearest outlets[j]	0.80	0.64, 1.00	0.71***	0.54, 0.88	−0.87***	−1.21, −0.53	−1.20***	−1.52, −0.88
Mean distance to seven nearest outlets, adjacent tract	0.85	0.64, 1.13	−0.19	−0.40, 0.03	0.10	−0.32, 0.53	0.07	−0.34, 0.47
Drug arrests	1.24***	1.15, 1.34	−0.06*	−0.11, −0.01	0.11	−>0.01, 0.22	0.17***	0.07, 0.28
Percent Black	1.02	0.98, 1.05	−0.09***	−0.11, −0.06	0.12***	0.07, 0.16	0.16***	0.11, 0.20
Vacant housing	1.03	0.96, 1.10	0.06*	0.01, 0.11	−0.11*	−0.21, −0.01	−0.12*	−0.21, −0.03
Median annual household income	0.92***	0.89, 0.96	0.01	−0.02, 0.04	−0.05	−0.11, 0.01	−0.07*	−0.13, −0.02
Percent population aged 18–35	1.04	0.96, 1.13	−0.02	−0.09, 0.04	0.02	−0.11, 0.14	0.09	−0.03, 0.21
Population density	1.03	0.85, 1.24	−0.01	−0.15, 0.13	0.38**	0.10, 0.65	0.44***	0.18, 0.71
High-density cluster	1.46**	1.15, 1.85	−0.18*	−0.36, −0.01	0.34	−0.01, 0.69	0.44*	0.10, 0.77
Moran's I	0.14**		***		0.08*		0.12**	

	Count of Violent Crimes[a]		Proximity to Nearest Violent Crime[b]		SAI for Seven Nearest Violent Crimes[c]		SAI for Violent Crimes in 0.25-Mile Buffer[d]	
	IRR	95% CI	β	95% CI	β	95% CI	β	95% CI
	Model #7		Model #15		Model #23		Model #31	
SAI for seven nearest outlets[k]	1.26*	1.03, 1.53	−0.52***	−0.69, −0.35	0.84***	0.53, 1.16	1.07***	0.76, 1.37
SAI for seven nearest outlets, adjacent tracts	1.09	0.84, 1.42	0.08	-0.15, 0.30	−0.21	−0.62, 0.20	−0.12	−0.52, 0.29
Drug arrests	1.25***	1.16, 1.34	−0.08**	−0.14, −0.02	0.13*	0.02, 0.23	0.21***	0.10, 0.31
Percent Black	1.01	0.98, 1.05	−0.09***	−0.11, −0.06	0.12***	0.08, 0.17	0.16***	0.11, 0.20
Vacant housing	1.03	0.96, 1.10	0.06*	<0.01, 0.11	−0.11*	−0.20, −0.01	−0.11*	−0.21, −0.02
Median annual household income	0.92***	0.88, 0.96	0.01	−0.02, 0.04	−0.05	−0.11, 0.01	−0.08*	−0.13, −0.02
Percent population aged 18–35	1.05	0.97, 1.14	−0.05	−0.11, 0.02	0.04	−0.08, 0.16	0.12*	<0.01, 0.24
Population density	1.03	0.85, 1.24	0.02	−0.13, 0.16	0.35*	0.08, 0.63	0.41**	0.14, 0.67
High-density cluster	1.50***	1.19, 1.89	−0.21*	−0.40, −0.02	0.39*	0.04, 0.74	0.51**	0.17, 0.84
Moran's I	0.14**		0.18***		0.09*		0.14**	
	Model #8		Model #16		Model #24		Model #32	
SAI for outlets within 0.25 mile[l]	1.08*	1.01, 1.15	−0.13***	−0.18, −0.07	0.20***	0.10, 0.30	0.26***	0.16, 0.36
SAI for outlets within 0.25 miles, adjacent tracts	1.09	0.99, 1.20	−0.07	−0.15, 0.01	0.11	−0.05, 0.26	0.18*	0.02, 0.33
Drug arrests	1.26***	1.17, 1.35	−0.10**	−0.16, −0.04	0.15**	0.04, 0.26	0.24***	0.13, 0.35
Percent Black	1.02	0.99, 1.05	−0.09***	−0.12, −0.07	0.13***	0.08, 0.17	0.17***	0.12, 0.22
Vacant housing	1.04	0.97, 1.11	0.04	−0.01, 0.10	−0.08	−0.19, 0.02	−0.08	−0.19, 0.02
Median annual household income	0.93***	0.90, 0.97	−0.01	−0.04, 0.03	−0.02	−0.09, 0.04	−0.01	−0.10, 0.02
Percent population aged 18–35	1.04	0.96, 1.13	−0.05	−0.12, 0.02	0.04	−0.09, 0.17	0.13	>−0.01, 0.26
Population density	1.02	0.85, 1.23	0.05	−0.11, 0.20	0.31*	0.03, 0.59	0.35*	0.06, 0.63
High-density cluster	1.48***	1.18, 1.86	−0.25*	−0.44, −0.05	0.41*	0.06, 0.76	0.57*	0.21, 0.92
Moran's I	0.16***		0.16***		0.07		0.11**	

[*] $p<0.05$,

** $p<0.01$,

*** $p<0.001$

IRR= incidence rate ratio; SAI=Spatial accessibility index; CI=Confidence Interval

[a]Calculated as the number of violent crimes (i.e., homicides, aggravated assaults, rapes, and robberies) in the census block in 2016.

[b]Calculated as the minimum network (road-based) distance from the census block centroid to the closest violent crime.

[c]Calculated by summing the inverse network (road-based) distances from the census block centroid to each of the seven closest violent crimes.

[d]Calculated by summing the inverse network (road-based) distance from the census block centroid to each violent crime located within a 0.25-mile buffer.

[e]Number of drug arrests/vacant houses inside CB. Variables were transformed using the natural logarithm.

[f]A one-unit increase equals a 10% increase in percent Black and percent aged 18–35 years.

[g]Median annual household incomes greater than $250,000 are censored, and a one-unit increase is $10,000.

[h]Clusters were defined as groups of 50 or more alcohol outlets all located within 0.1 miles from the nearest outlet.

[i]Calculated as the minimum network (road-based) distance from the census block centroid to the closest alcohol outlet.

[j]Calculated as the average network (road-based) distance from the census block centroid to the seven closest alcohol outlets.

[k]Calculated by summing the inverse network (road-based) distances from the census block centroid to each of the seven closest alcohol outlets.

[l]Calculated by summing the inverse network (road-based) distance from the census block centroid to each alcohol outlet located within a 0.25-mile buffer.

Table 4.

Akaike's Information Criterion for Regression Analyses of Alcohol Outlet Density on Violent Crime Exposure, Baltimore City 2016

Alcohol outlet density	Count of Violent Crimes[a]			Proximity to Nearest Violent Crime[b]			SAI for Seven Nearest Violent Crimes[c]			SAI for Violent Crimes in 0.25-Mile Buffer[d]		
	RMSE	Partial R^2	AIC	RMSE	Partial R^2	AIC	RMSE	Partial R^2	AIC	RMSE	Partial R^2	AIC
Raw count of alcohol outlets	38.1	0.2	1,864	0.50	0.0	284	0.84	1.7	508	0.92	1.8	535
Count of alcohol outlets divided by total population	41.4	0.6	1,857	0.48	1.0	271	0.84	4.0	501	0.91	5.1	522
Count of alcohol outlets divided by total area in square miles	35.8	2,6	1,846	0.44	10.0	236	0.75	13.8	462	0.77	17.5	459
Count of alcohol outlets divided by total roadway miles	36.5	1.3	1,862	0.49	1.4	278	0.85	0.0	506	0.91	0.1	529
Proximity to nearest outlet[e]	37.1	2.3	1,844	0.38	26.0	195	0.76	15.2	455	0.72	25.4	433
Mean distance to seven nearest outlets[f]	38.5	2.3	1,849	0.39	26.5	190	0.77	11.9	465	0.72	22.1	444
SAI for seven nearest outlets[g]	38.2	3.1	1,850	0.41	16.6	216	0.76	13.1	464	0.75	20.3	453
SAI for outlets inside 0.25-mile buffer[h]	37.3	3.1	1,847	0.44	9.5	239	0.79	7.4	478	0.82	11.7	479

AIC=Akaike's Information Criterion; SAI=Spatial accessibility index; RMSE=root mean squared error

[a]Calculated as the number of violent crimes (i.e., homicides, aggravated assaults, rapes, and robberies) in the census block in 2016.

[b]Calculated as the minimum network (road-based) distance from the census block centroid to the closest violent crime.

[c]Calculated by summing the inverse network (road-based) distances from the census block centroid to each of the seven closest violent crimes.

[d]Calculated by summing the inverse network (road-based) distance from the census block centroid to each violent crime located within a 0.25-mile buffer.

[e]Calculated as the minimum network (road-based) distance from the census block centroid to the closest alcohol outlet.

[f]Calculated as the average network (road-based) distance from the census block centroid to the seven closest alcohol outlets.

[g]Calculated by summing the inverse network (road-based) distances from the census block centroid to each of the seven closest alcohol outlets.

[h]Calculated by summing the inverse network (road-based) distance from the census block centroid to each alcohol outlet located within a 0.25-mile buffer.