Drug and Alcohol Dependence 178 (2017) 188–193

Contents lists available at ScienceDirect

Drug and Alcohol Dependence

journal homepage: www.elsevier.com/locate/drugalcdep

Full length article

The spatial accuracy of geographic ecological momentary assessment (GEMA): Error and bias due to subject and environmental characteristics

Jeremy Mennis[a,*], Michael Mason[b], Andreea Ambrus[a], Thomas Way[c], Kevin Henry[a]

[a] Temple University, Philadelphia, PA, United States
[b] University of Tennessee, Knoxville, TN, United States
[c] Villanova University, Radnor Township, PA, United States

ARTICLE INFO

Keywords:
Ecological momentary assessment
EMA
Location
Ecology
Environment
Neighborhood
Accuracy
Data quality

ABSTRACT

Background: Geographic ecological momentary assessment (GEMA) combines ecological momentary assessment (EMA) with global positioning systems (GPS) and geographic information systems (GIS). This study evaluates the spatial accuracy of GEMA location data and bias due to subject and environmental data characteristics.

Methods: Using data for 72 subjects enrolled in a study of urban adolescent substance use, we compared the GPS-based location of EMA responses in which the subject indicated they were at home to the geocoded home address. We calculated the percentage of EMA locations within a sixteenth, eighth, quarter, and half miles from the home, and the percentage within the same tract and block group as the home. We investigated if the accuracy measures were associated with subject demographics, substance use, and emotional dysregulation, as well as environmental characteristics of the home neighborhood.

Results: Half of all subjects had more than 88% of their EMA locations within a half mile, 72% within a quarter mile, 55% within an eighth mile, 50% within a sixteenth of a mile, 83% in the correct tract, and 71% in the correct block group. There were no significant associations with subject or environmental characteristics.

Conclusions: Results support the use of GEMA for analyzing subjects' exposures to urban environments. Researchers should be aware of the issue of spatial accuracy inherent in GEMA, and interpret results accordingly. Understanding spatial accuracy is particularly relevant for the development of 'ecological momentary interventions' (EMI), which may depend on accurate location information, though issues of privacy protection remain a concern.

1. Introduction

Ecological momentary assessment (EMA) allows researchers to collect data on subjects' behaviors, moods, and social interactions in real time and in subjects' natural environments, offering a unique ability to investigate ecological and social mechanisms and consequences of substance use (Shiffman et al., 2008). Recently, EMA has been combined with global positioning systems (GPS) technology embedded in mobile phones to encode location at the time of EMA response, or, in rarer cases, as a continuous stream of location data. These location data capture an individual's activity space – the spatial expression of the home location and routine places an individual travels for work, leisure, and other activities (Mennis and Mason, 2011). These rich, spatial EMA data can be integrated with other spatial data describing the natural, built, and socioeconomic environments with which a subject interacts throughout their daily life using geographic information systems (GIS). This geographic EMA (GEMA) approach (Epstein et al., 2014; Kirchner and Shiffman, 2016), facilitates the investigation of contextual effects on health behaviors to a degree that was heretofore impossible prior to recent developments in mobile and geospatial technologies (Stahler et al., 2013). While the past several years has seen substantial growth in GEMA for analyses of tobacco, alcohol, and illicit drug use (Epstein et al., 2014; Freisthler et al., 2014; Kirchner et al., 2013; Mason et al., 2016; Mennis et al., 2016; Schuz et al., 2015), the technique is still in its infancy.

One issue that has yet to be investigated by GEMA researchers concerns the spatial accuracy of the location data collected via GPS-embedded mobile phones. The spatial accuracy of GPS is dependent upon the ability of the GPS receiver to communicate with orbiting GPS satellites, and thus may be mitigated by tall or dense structures (as in urban areas); hilly terrain; indoor use; atmospheric attenuation; or poor geometric configuration of the GPS receiver with the GPS satellites (Garnett and Stewart, 2015; Wing et al., 2005; Zandbergen and Barbeau, 2011). While GEMA researchers have noted challenges to

* Corresponding author.
E-mail address: jmennis@temple.edu (J. Mennis).

http://dx.doi.org/10.1016/j.drugalcdep.2017.05.019
Received 6 March 2017; Received in revised form 9 May 2017; Accepted 10 May 2017
Available online 22 June 2017
0376-8716/ © 2017 Elsevier B.V. All rights reserved.

collecting location data in urban and indoor settings, as well as due to the potential reluctance of subjects to reveal the location of illicit activities such as substance use (Mitchell et al., 2014; Rudolf et al., 2016; Watkins et al., 2014), to the best of our knowledge, no studies have explicitly examined the spatial accuracy of GEMA location data.

The present study seeks to address this gap in the literature by investigating the degree to which the location data collected through GEMA via GPS-embedded mobile phones captures the actual locations at which subjects respond to the EMA prompts. The objectives of this research are two-fold: 1) to assess the spatial accuracy of GEMA location data in an urban environment, and 2) to examine the association of GEMA spatial accuracy with subject and environmental characteristics, which can reveal sources of error and bias in accuracy assessment. By investigating the spatial accuracy of GEMA location data, we aim to contribute to a better and more refined understanding of the analytical potential and limitations of GEMA methods.

2. Data and methods

2.1. GEMA data

GEMA location data for this study comes from the Social-Spatial Adolescent Study, a longitudinal study of neighborhood and social contextual effects on adolescent substance use, based in Richmond, Virginia (Mason et al., 2016). Two hundred and forty-eight subjects were recruited for the study, primarily from an adolescent medicine outpatient clinic, during 2012–2014. The criteria for inclusion were age (13–14 years old), patient status (a registered clinic patient), and residency (Richmond area resident), with written informed consent obtained for all participants and parents/guardians. All participants were given a mobile phone for the period of the study, through which EMA surveys were delivered and within which was embedded an assisted GPS (A-GPS) receiver, common to many mobile phones, which captures location using GPS as augmented by data provided through cellular communications. EMA surveys were administered to each subject via text message with an embedded URL link 3–6 times per day, over a four-day period, every other month, over a period of two years. Each EMA survey takes less than one minute to complete, and subjects were given nine minutes to complete each survey. The EMA responses were gathered via a cloud-based, third-party, text-messaging platform on a secured Linux web server and included each participant's survey answers, geographic coordinate position, and response date and time. We use the term 'EMA location' to refer to the geographic coordinate position captured for an individual EMA survey response.

2.2. Verification data

The challenge in estimating the spatial accuracy of the EMA locations is the absence of verification data, i.e., the true coordinate position of each EMA location. To this end, we restrict our sample to those EMA responses where the subject has indicated that they are at home when responding to the EMA survey, where the first question of each EMA survey asks "Where are you now?" We compare the coordinate positions of the home-based EMA locations to the coordinate position of the subject's geocoded home address. The sample used in this study includes only the subjects who did not move during the entire two-year period of the study. We also restricted the analysis to subjects with more than five EMA responses, in order to ensure an adequate sample of EMA locations for each subject.

2.3. Analysis of EMA location accuracy

Following other studies of GPS accuracy (Schipperjin et al., 2014; Zandbergen and Barbeau, 2011), we calculate the percentage of EMA locations that fall within a series of distance bands from the verification data: a half mile (2640 feet), quarter mile (1320 feet), eighth mile (660 feet), and sixteenth mile (330 feet) of the geocoded home, as measured using the straight-line (Euclidean) distance using ArcGIS software (ESRI, Inc.). We then calculate, for each subject, the percentage of EMA observations that occur within each distance band. We also calculate the percentage of EMA observations that occur in the correct U.S. Census Bureau tract and block group, which has the advantage of evaluating the utility of joining EMA location data with Census units commonly used to capture neighborhood disadvantage and disorder in contextual studies of substance use (Mennis et al., 2016). Thus, for each subject we have a set of six accuracy assessments, capturing displacement from each subject's home, using different distance thresholds (e.g., within an eighth mile, within the same tract). For each of the six spatial accuracy measures we report the mean, standard deviation, and quartile values (minimum, 25th percentile, median, 75th percentile and maximum) over the set of subjects.

2.4. GPS receiver error and sources of bias

We use the term 'GPS receiver error' to refer to the error in the reported geographic coordinate position of the EMA location associated with the attenuation of the GPS receiver and satellite communications. We acknowledge, however, that other factors can bias the accuracy results. The first, 'subject response bias,' occurs when the subject has intentionally or accidentally indicated that they are at home for a particular EMA survey but they are not. We also acknowledge that there is some ambiguity to the concept of 'home;' For instance, a subject's parents may be divorced and the subject considers more than one house their home, or the subject also considers another relative's house their home, or the subject is nearby their home and thus indicates that they are home in the EMA survey.

The second type of bias, 'verification data bias,' concerns the uncertainty associated with the geocoded home location, which can occur because the text address acquired from the subject is incorrect or because geocoding generates a geographic coordinate position that does not perfectly capture the location of the home structure. Notably, street address geocoding estimates the coordinate position of an address along a street segment based on the range of addresses along that segment, whereas address point geocoding assigns the coordinate position of an address using a preset coordinate position associated with an address. In addition, some of our subjects live in large apartment complexes where the street address may not capture the location of the actual apartment. For these reasons, we expect the EMA location to be nearby, but not always perfectly coincident with, the geocoded home location. This is one motivation for reporting subject-level accuracy as a percentage of EMA locations within a certain distance of the geocoded home; it accounts for the uncertainty in verification data location (Zandbergen and Barbeau, 2011).

Two other types of potential verification data bias concern the use of tract and block group data; their boundaries may not be perfectly registered (geographically aligned) with the geocoding reference data, and the precision of the tract and block group boundary data may not be adequate for comparing geocoded and EMA locations. These two types of bias may be particularly problematic for home addresses located on major streets that serve as boundaries of tracts or block groups, where a small difference in registration or precision could cause a home location to be geocoded in the wrong tract or block group.

We perform statistical analyses to investigate the role of environmental, subject, and verification data mechanisms of error and bias in the accuracy results. We employ the well-known Moran's *I* test (Lloyd, 2010), using an inverse distance squared weighting function, to assess whether accuracy is spatially clustered across subjects' geocoded home locations, which would suggest an environmental correlate of accuracy indicating GPS receiver error. Directional bias in the accuracy measures, which would suggest verification data bias associated with registration mismatch between the EMA locations and the geocoded home locations, is assessed by calculating the compass bearing from the

Download English Version:

https://daneshyari.com/en/article/5120348

Download Persian Version:

https://daneshyari.com/article/5120348

Daneshyari.com