American Dietetic Association

RESEARCH

Research and Professional Briefs

Public Directory Data Sources Do Not Accurately Characterize the Food Environment in Two Predominantly Rural States

MEGHAN R. LONGACRE, PhD; BRIAN A. PRIMACK, MD, EdM; PETER M. OWENS, PhD; LUCINDA GIBSON, MS; SANDY BEAUREGARD; TODD A. MACKENZIE, PhD; MADELINE A. DALTON, PhD

ABSTRACT

Communities are being encouraged to develop locally based interventions to address environmental risk factors for obesity. Online public directories represent an affordable and easily accessible mechanism for mapping community food environments, but may have limited utility in rural areas. The primary aim of this study was to evaluate the efficacy of public directories vs rigorous on-site field verification to characterize the community food environment in 32 geographically dispersed towns from two rural states covering 1,237.6 square miles. Eight types of food outlets were assessed in 2007, including food markets and eating establishments, first using two publically available online directories followed by onsite field verification by trained coders. χ^2 and univariate binomial regression were used to determine whether the proportion of outlets accurately listed varied by food outlet type or town population. Among 1,340 identified outlets, only 36.9% were accurately listed through public directories; 29.6% were not listed but were located during field observation. Accuracy varied by outlet type, being most accurate for big box stores and least accurate for farm/produce stands. Overall, public directories accurately identified fewer than half of the food outlets. Accuracy was significantly lower for rural and small towns compared to mid-size and urban towns ($P<0.001$). In this geographic sample, public directories seriously misrepresented the actual distribution of food outlets, particularly for rural and small towns. To inform local obesity-prevention efforts, communities should strongly consider using field verification to characterize the food environment in low-population areas.

J Am Diet Assoc. 2011;111:577-582.

M. R. Longacre is an instructor, Community Health Research Program, Hood Center for Children and Families, Department of Pediatrics, Dartmouth Medical School, Lebanon, NH. B. A. Primack is an assistant professor, Divisions of General Internal Medicine and Adolescent Medicine, Departments of Medicine and Pediatrics, Center for Research on Health Care, University of Pittsburgh School of Medicine, Pittsburgh, PA. P. M. Owens is an urban planner, L. Gibson is principal, and S. Beauregard is an engineer, Smart Mobility, Inc, Norwich, VT. T. A. Mackenzie is an associate professor, Community Health Research Program, Hood Center for Children and Families, Departments of Community and Family Medicine and Medicine, Dartmouth Medical School, Lebanon, NH. M. A. Dalton is a professor, Community Health Research Program, Hood Center for Children and Families, Departments of Pediatrics and Community and Family Medicine, Dartmouth Medical School, Lebanon, NH.

Address correspondence to: Meghan R. Longacre, PhD, Community Health Research Program, Hood Center for Children and Families, Department of Pediatrics, Dartmouth Medical School, HB 7465, One Medical Center Dr, Lebanon, NH 03756-0001. E-mail: Meghan.R.Longacre@dartmouth.edu

Manuscript accepted: September 22, 2010.

0002-8223/$36.00

doi: 10.1016/j.jada.2011.01.008

Rural residence is an important correlate of obesity (1,2). Characteristics of rural environments, including limited access to healthy foods, may influence obesity-related behaviors (3). In response to increasing calls for environmentally based modifications to address obesity (4-6), communities are developing local interventions targeting geographic risk factors (7-9). The Centers for Disease Control and Prevention recommends appropriate measurement of community food environments to inform these obesity-prevention strategies (4). Although on-site field validation is recognized as the gold standard for identifying community food sources, this method is both costly and time-intensive, particularly for rural areas characterized by large expanses of undeveloped land (10,11). Use of secondary data sources, such as those available through commercial databases and public directories, offer local communities an easily accessible and typically no-cost mechanism for mapping their food environment. More research is needed, however, on the validity of secondary data sources for describing food environments in rural areas (12-14).

Several researchers have compared the accuracy of secondary data sources vs field validation in urban communities outside of the United States. These studies report accuracy between 65% and 85% for commercial databases and local government listings, and between 50% and 65% for Internet-based listings (14-16). All three types of secondary data sources are not without limitations. Specifically, commercial databases may exclude information on low-revenue, locally owned food establishments; listings

within governmental databases may have insufficient information to classify food outlet types in detailed categories; and Internet listings may be updated infrequently (13,14).

Few studies have compared the validity of secondary data sources vs field validation in rural areas of the United States. Sharkey found that public lists omitted between 20% and 36% of field validated food markets in six impoverished, remote counties in Central Texas (17). In addition, only one study, conducted in an urban city in the United Kingdom, examined the accuracy of secondary sources by differing food outlet types, such as food markets and restaurants (14). Lake and colleagues (14) demonstrated that restaurants and pubs were most likely to be listed on public data sources but not found in the field. Others have recognized specific challenges in using commercial databases to characterize unique food environments, such as those associated with ethnic minority communities (18). Similarly, commercial data sources may have limited use in rural compared to urban areas because of lower precision geocoding (19,20) and a greater presence of smaller, locally owned establishments for procuring foods (eg, seasonal farm stands, general stores). The purpose of the current study was to evaluate the efficacy of using secondary data sources vs rigorous field validation to characterize the food environment in two predominantly rural states. Specific aims investigated whether accuracy varied by food outlet type or by degree of rurality. Information obtained from two public directory Internet sites was selected for comparison with field validation because it was expected that these data would be most easily and quickly accessible by local communities.

METHODS

Data for the current study were collected as part of a larger study of individual, family, and environmental influences on adolescent obesity in primarily rural and small town geographic areas of Northern New England. The study, titled "Environmental and Family Influences on Adolescent Overweight," was approved by the Committee for the Protection of Human Subjects at Dartmouth College.

Data Collection

In 2007, two public directory Internet sites were used to create an inventory of town-wide food outlets for 32 geographically dispersed towns throughout New Hampshire and Vermont. Food outlet data were first collected via the "places of interest" function on Google Earth (Google Inc; Mountain View, CA), which provides business and geographic location data gathered from a variety of commercial sources (21). The inventory was then augmented using Yahoo! Yellow Pages (Yahoo! Inc; Sunnyvale, CA). Yahoo! Yellow Pages (which was functioning in 2007, but closed as of March 2010 and was replaced by Yahoo! Local) collects business listings through its data provider, InfoUSA (Infogroup Inc; Omaha, NE), one of the largest commercial business databases worldwide (22,23). It was expected that these two sites would maximize the advantages of both commercial and Internet listings.

Towns were selected based on town-of-residence for an ongoing study (24). ArcGIS 9.1 (ESRI, 2004, Redlands, CA) was employed to create an aerial photo map of each town that identified town boundaries, street networks, and inventoried food outlet locations derived from the public directories. Field verification was conducted within 1 month of public directory data collection by two-person coding teams who systematically drove all town street networks, confirmed the presence and location of inventoried food outlets, and identified onsite outlets not included on the inventory. The accuracy of public directories vs field observations was evaluated as follows: outlet identified on Internet and found at expected location (accurately listed); outlet identified on Internet and found at a different location (mislocated); outlet identified on Internet but not found through field observation (not found); and outlet not identified on Internet but found through field observation (not listed). Outlets were considered mislocated if coders could not visually locate the outlet while positioned at the Internet-identified location. Categorization into the accuracy groups was based on the two-person coding decisions during onsite town visits and used geocoded food outlet location data and detailed town maps.

Field coders used a structured Community Food Observation Form and a detailed manual, developed for the current study, to categorize and describe food outlets. The Community Food Observation Form was developed by a team of experienced researchers and geographic experts after a thorough review of the literature and extensive observations in towns of similar size and rurality to the study towns. Before data collection, we pretested the public data download process and the Community Food Observation Form in four nonstudy towns, which allowed us to establish face validity and comprehensiveness of the food outlet categories. During pretesting, we evaluated inter-rater reliability of the coders' field observations, including identification of all food outlets and categorization of food outlet type. We found 100% agreement between the two coding teams for each of these measures.

Coders classified outlets as either food markets, consisting of six specific outlet categories (ie, general store; convenience store; supermarket/grocery store; specialty food store; big box store; seasonal and year-round fixed location farm/produce stand) or eating establishments, consisting of two outlet categories (ie, fast-food restaurants, defined as any food outlet where the patron orders food at a counter or window; and full-service restaurants). General stores are defined as local retailers with a broad selection of merchandise, including grocery items, hardware, and gardening supplies. Big box stores included warehouse membership clubs (eg, BJ's [Westborough, MA], Sam's Club [Bentonville, AR]) and large retail supercenters, provided they contained packaged food/grocery sections. Specialty food stores included food outlets that exclusively sold a specific type of food, such as meat or fish markets. In-store observations were conducted to verify outlet classification. The resulting eight categories represent a modified version of the North American Industry Classification System (25). Food markets housing a fast-food business (n=43) were counted as two distinct outlets if, based on in-store observations, the fast-food section had a separate name or logo, entryway, cash