Applied Geography 51 (2014) 82–89

Contents lists available at ScienceDirect

Applied Geography

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

Does food environment influence food choices? A geographical analysis through "tweets"

Xiang Chen*, Xining Yang

Center for Urban and Regional Analysis and Department of Geography, The Ohio State University, Columbus, OH 43210, United States

Keywords:
Food access
Food environment
Space-time
Twitter
Social media
VGI

A B S T R A C T

Access to nutritious food is imperative to physical well-being and quality of life. Previous food environment studies have revealed a disparity of access to healthful food on various geographical scales. An overlooked facet of this spatial perspective is the impact of the food environment at the individual level. Individuals tend to make diverse food purchasing and dining choices, including where, when, how, and which types of food to acquire. An unexplored avenue for further investigation is measuring the extent to which people's preference for food is elicited by exposure to their immediate food environment. This paper takes an innovative approach to this question by soliciting individual data about food-related activities from social media, or specifically, "tweets" (messages sent on Twitter). With spatiotemporally tagged information, tweets provide an ideal method for measuring the exposure to the food environment in real time. This measure, as a representative of individual food access, is associated with users' particular diet choices conveyed in their tweets. By comparing groups of Twitter users who shop in grocery stores to those who dine at fast food restaurants, we found that the prevalence of grocery stores that stock fresh produce within an individual's neighborhood may significantly influence him or her to make nutritious food choices. This study has a great potential to inform health professionals and stakeholders of the significance of social media in assisting with crowdsourcing human subject data that incorporate spatiotemporal dimensions and to explore individual diets in relation to their perceived food environment, which can positively impact the health of communities.

Introduction

Access to nutritious food is imperative to physical well-being and quality of life. Failing to consume healthful food on a regular basis can lead to a series of adverse health outcomes, including obesity, diabetes and cardiovascular diseases (Shaw, 2006). Researchers have noted that access to nutritious food is influenced by both spatial and non-spatial mediators that result in inequality of access across communities. These barriers to healthful food choices include geographic, economic, informational, and cultural aspects (McEntee & Agyeman, 2010). To date, the majority of the literature simply focuses on geographic access to nutritious food in relation to socio-economic status (SES). These studies have taken a predominantly statistical approach to examining if correlations exist between SES variables and food access though regression models. Not surprisingly, many studies found a positive correlation between low SES and limited access to quality food in selected local regions (Glanz, Sallis, Saelens, & Frank, 2007; Moore & Diez Roux, 2006) as well as in U.S. nationwide studies (Powell, Slater, Mirtcheva, Bao, & Chaloupka, 2007), and few studies failed to identify this correlation (Guy & David, 2004).

An overlooked facet in this spatial disparity of food access is the causality linked to food choices at the individual level. Individuals tend to make diverse food purchasing and dining choices, including where, when, how, and which types of food to acquire. The role of food access in shaping food choices cannot be completely understood from generalized regional studies. What must be assessed is the impact of the quality of an individual's food environment on his or her food purchasing choices. Although studies have previously identified among other factors the availability of nearby grocery stores have played a significant role in influencing food buying practices (Walker, et al. 2011; Walker, Block, & Kawachi, 2012), these studies suffer from the following limitations: (1) the sampling size is very limited due to the time-consuming process of soliciting individual samples and (2) studies of individuals' dietary choices are based on the home locations while overlooking the effects of mobility on procuring food (Kestens, Lebel, Daniel, Thériault, &

* Corresponding author.
E-mail addresses: chen.1616@osu.edu, chenxiangpeter@gmail.com (X. Chen).

http://dx.doi.org/10.1016/j.apgeog.2014.04.003
0143-6228/

Pampalon, 2010; McKinnon, Reedy, Morrissette, Lytle, & Yaroch, 2009). At the individual level, the deficiency in food procurement may be due to a scarcity of healthful food retailers in individuals' vicinity. An unexplored direction is to measure if people's preference for food is elicited by exposure to a particular type of food environment in real time, such as a lack of quality grocery stores nearby that would prevent them from making nutritious food choices. This finding would also offer compelling evidence for shaping strategies to improve the health of communities.

The traditional measure of food access based on location of food outlets has considerable limitations for understanding the causality of food choices (Wrigley, Warm, Margetts, & Whelan, 2002); on the other hand, individual-level analysis has always posed a daunting challenge for social scientists because of the tedious nature of collecting a considerable number of representative samples. This paper takes an innovative approach to soliciting individual data about food-related activities from Volunteered Geographic Information (VGI) via social media, or specifically, "tweets" (messages sent on Twitter). Tweets not only contain self-reported information about life experiences and activities but are also tagged with accurate location and time information. They provide an ideal solution for measuring exposure to the food environment at the exact time when the user tweets. This measure, as a representative of individual food access, is associated with users' diet choices conveyed in their tweets. We then analyzed them to determine if there was a significant association between their surrounding food environment and the quality of their particular food choices. When groups of Twitter users who shop in grocery stores were compared to those who dine at fast food restaurants, it was revealed that the prevalence of nearby fresh produce grocery stores may significantly influence individuals to make healthful food choices. Conversely, density of fast food restaurants in an individual's immediate area may not discourage healthful food choices. To the authors' knowledge, this is the first time social media has been adopted for the evaluation of individuals' food-related activities that are influenced by one's local food environment. This study has great potential for informing health scientists about the importance of using social media to measure individual behavior and for informing community advocates and stakeholders of the significant role quality food retailers can play in changing the dietary practices of area residents.

Measuring access to food: place-based vs. individual-based

The problem of food access can be interpreted in many different ways. It not only involves geographic access to food but also pertains to affordability, availability of culturally appropriate food, and knowledge about nutrition (McEntee & Agyeman, 2010; Shaw, 2006). Beyond these descriptions, the majority of studies approach the problem from a rather geographic perspective, where the metaphor "food desert" is employed to illustrate areas devoid of safe, affordable grocery stores with extensive arrays of quality food items. Such grocery stores which offer fresh fruits and vegetables are termed "green retailers" (Wrigley, Warm, Margetts, & Whelan, 2002). A consequence of a lack of such retailers in urban food deserts is an infiltration of poor-quality food outlets, such as convenience stores and fast food restaurants, where only packaged, processed, and energy-dense food is provided (Drewnowski & Specter, 2004).

Reducing and eliminating food deserts has been the goal of stakeholders and government agencies around the world ever since the term was first coined by the U.K.'s Nutrition Task Force's Low Income Project Team in 1996 (Reisig & Hobbiss, 2000). Since then, it has become a topic of interest in a Congressional report compiled by the U.S. Department of Agriculture (USDA) in 2009 (USDA, 2009). A method for evaluating the formation of food deserts is needed before a solution can be found. To date, there is a lack of consensus on methods used for differentiating high/low access areas and demarcating food deserts. With respect to measuring food access, the basic rationale differs from either a place-based perspective or an individual-based perspective.

A place-based perspective examines the built environment in which people acquire food in multiple contexts, such as food stores, restaurants, worksites, and schools (McKinnon, Reedy, Morrissette, Lytle, & Yaroch, 2009). A typical place-based approach measures the level of geographic access in terms of the diversity, proximity and variety of food outlets in a predefined geographic unit, such as a neighborhood, census tract, census block, zip code zone, county, or state (Apparicio, Cloutier, & Shearmur, 2007). The geographic unit serves as a container for summarizing the spatial attributes of food outlets in terms of total number (Berg & Murdoch, 2008; Powell, Slater, Mirtcheva, Bao, & Chaloupka, 2007), number relative to population (Ball, Timperio, & Crawford, 2009), average number within a distance of the centroid of the unit (Larsen & Gilliland, 2008; Raja, Ma, & Yadav, 2008), and distance to the nearest store from the centroid (Larsen & Gilliland, 2008; Pearce, Witten, & Bartie, 2006). Disparity of access has been identified across geographic units, compounded with regression models for determining if isolation from a quality food environment is correlated with racial composition or with socio-economical deprivation of districts. Another place-based measure, which relies on the emergence of Geographic Information Systems (GIS), takes into account the spatial distribution of food resources by creating a circularly buffered zone or a network-shaped zone around the focal location of food outlets (Algert, Agrawal, & Lewis, 2006; Chen & Clark, 2013; Clarke, Eyre, & Guy, 2002). The creation of the buffer is grounded in the perception that sustainable neighborhoods should have sufficient access to nutritious food within a walkable distance. Areas outside the food access zones are considered underserved and are in need of policy intervention. In spite of the container measure and buffer measure, other place-based measures explore emerging geographic analysis methods by examining the spatial relationship between food supply and demand. Examples of these methods include various accessibility models (Helling & Sawicki, 2003; Paez, Mercado, Farber, Morency, & Roorda, 2010) and spatial interaction models (Clarke, Eyre, & Guy, 2002).

The place-based approach is limited by the fact that it generalizes individual attributes, such as gender, age, ethnicity, and other demographic aspects that vary widely among people (Miller, 2007). Because people's life experiences vary, their preferences for food and knowledge of nutrition are difficult to infer from aggregated census data. In order to capture these distinctions, the individual-based approach can be adopted by focusing directly on the food choices of people grouped by neighborhood, age, gender, and social roles. Importantly the foci of these individual level studies include general households (Wrigley, Warm, Margetts, & Whelan, 2002) or selected groups such as low-income rural residents (Smith & Morton, 2009), women (Ball, Crawford, & Mishra, 2006; Laraia, Siega-Riz, Kaufman, & Jones, 2004), and male adolescents (Jago, Baranowski, & Harris, 2006). Once the population of interest is defined, standardized assessment tools in the form of videotaped interviews or paper-based questionnaires are customized and distributed to the specific population. The proposed survey questions are not only limited to accessibility to food but are more focused on the respondents' financial status, nutritional knowledge, health concerns, food purchasing and consumption behaviors, and the neighborhood food environment that shapes their daily food choices. Geographic access is also explored through various methods for measuring foodscapes at the individual level by referring to the home address of respondents (Ball, Timperio, & Crawford, 2009; Jago, Baranowski, & Harris, 2006; Jeffery, Baxter, McGuire, & Linde, 2006; Laraia, Siega-Riz, Kaufman, & Jones, 2004; Wrigley, Warm, Margetts, & Whelan, 2002).

The majority of previous individual-based studies are flawed in two ways. First, home addresses are fixed locations that only represent a limited range of activity. People's daily movements can extend beyond their households and expose them to other food environments, such as cafeterias at employment sites or schools (Kestens, Lebel, Daniel, Thériault, & Pampalon, 2010). Their mobility not only changes the physical environment that limits the choices of food resources but also introduces food cues, such as advertisements and interpersonal communications that may elicit an increased craving for certain foods (Ferriday & Brunstrom, 2008). In addition, individuals' food procurement is not only spatially but also temporally constrained. As the operating hours of food stores are invariably limited, those whose discretionary time conflicts with store hours will find it difficult to access food, eventually severely limiting the quality of their food choices (Chen & Clark, 2013). Previous food studies, however, have overlooked this temporal inaccessibility to food stores and the contraction of activity space impinged by time.

New instruments for individual-based food studies

The recent trend centering on individuals examines exposure to one's local food environment in terms of availability or density of food resources attainable in one's daily travels. The goal of these studies is to identify the mechanism by which environmental influences translate into individual food-related choices and to highlight the SES indicators that contribute to these choices. Previous studies have relied exclusively on techniques that trace individual movement: travel diary surveys (Kestens, Lebel, Daniel, Thériault, & Pampalon, 2010) and GPS tracking (Chaix, Meline, Duncan, Merrien, Karusisi, Perchoux, et al., 2013). Travel diaries are collected in the form of paper-based or online surveys from individual participants that document their activities over a specified span of time. They include questions to determine an individual's attitudes towards their search for food sources, which aid in behavioral and psychological analysis. However, the design, collection, processing, and analysis of individual travel diaries constitute a tedious and labor intensive process. The level of accuracy and precision of responses regarding the location and time of food-related activities also must be verified. GPS tracking allows for a more automated and accurate data collection process; however, the lack of contextual information poses challenges to categorizing GPS tracks into separate activities and the segmental trips between them (Chaix, Meline, Duncan, Merrien, Karusisi, Perchoux, et al., 2013). The geographically referenced individual data with excessive spatiotemporal details has been criticized for intruding on individual privacy (Kwan, Casas, & Schmitz, 2004). An improved approach is to combine GPS tracking with a follow-up activity survey. Unfortunately, previous studies using this approach invariably limited the time span of the survey to only one day (Chaix, Meline, Duncan, Merrien, Karusisi, Perchoux, et al., 2013).

Recent studies in applied geography have made use of Web 2.0, a concept coined to describe user-generated web contents in a bottom-up approach via social media, such as Twitter, Facebook, Flickr, and Google+ (Arribas-Bel, 2014). These self-reported contents, known as VGI, offer an alternative approach to keeping track of food-related activities from individual users' dietary preferences and shared personal experiences. Social media is becoming a new instrument for analyzing individual awareness of health information and prevalence of diseases (Robillard, Johnson, Hennessey, Beattie, & Illes, 2013). Emerging as the leading platform of social media, Twitter has established an active user base of 5.5 millions (Twitter, 2013), and this number has grown at an unprecedented rate in recent years with the increasing popularity of web-enabled mobile devices. In the spectrum of health studies, Twitter has been used to explore the spatial distribution and temporal variation of health-related subjects, including flu epidemics (Chew & Eysenbach, 2010), cholera outbreaks (Ehrenberg, 2012), and causes of dementia (Robillard, Johnson, Hennessey, Beattie, & Illes, 2013). To the authors' knowledge, this study is the first to engage Twitter for measuring the impact of the food environment on individual eating behavior.

Twitter messages, or tweets, serve as an ideal instrument for examining individual food access for several reasons: (1) tweets are geo-tagged in space and positioned in time, and with this precisely defined spatiotemporal information, the dynamic food environment to which the user is exposed can be retrieved with a high degree of accuracy (Arribas-Bel, 2014); (2) tweets entail user-generated individual activities, and individuals' particular food choices can be reflected by their content, allowing for food themed analyses; (3) tweets are voluntarily disseminated by individual users and are open to the public. This self-reported mechanism overcomes the privacy concerns of acquiring geocoded individual information.

Data collection and exclusion criteria: case study in Columbus, Ohio

Unlike other social media studies that encompass a relatively large geographic area, our case study was demarcated at a regional scale in order to evaluate the food environment in greater detail. Columbus, Ohio's capital city in the Midwest U.S. was adopted as the study region. According to the U.S. 2010 census, Columbus has an estimated population of 787,033, making it the largest city in Ohio. For this reason, it frequently functions as a "test city" for new consumer goods and attracts diverse industries (Hunker, 2000). The city is home to many business headquarters and fast food restaurants, including Wendy's, White Castle, and Bob Evans restaurants. The diversity of the food industry and a burgeoning population in Columbus make it an accurate representation of U.S. cities. We collected two sets of data within Franklin County, the largest county in Columbus: Tweets with verified content that include information about a current or upcoming food choice and food outlets, including quality grocery stores and fast food restaurants in the region. These two datasets were correlated during analysis to examine the relationship between Twitter users' food choices and their exposure to the local food environment.

Raw tweets that were (1) disclosed as publicly available and (2) tagged with the sender's location over the span of five weekdays (from Dec.2 2013 to Dec.6 2013) were streamed directly from Twitter using an open source Python library. These tweets were geocoded using *ESRI ArcMap* 10.2, a leading GIS platform, and were overlaid within the boundary of Franklin County in Central Columbus, Ohio. Eventually, a total of 81,543 tweets within the boundary were collected for further verification. In order to retrieve tweets that contained only food-related activities, they were narrowed down by using key word searches. Key words were meant to capture two very different sets of food sources available in this region: names of grocery store chains and individually owned stores selling fresh fruits and vegetables, known as "green retailers" (Wrigley, Warm, Margetts, & Whelan, 2002), such as Kroger's, Giant Eagle, Walmart, Meijer, Trader Joe's, Save-A-Lot, *etc*., and names of fast food restaurants listed in the local Yellow Pages, such as McDonald's, KFC, Subway, White Castle, Burger King, Taco Bell, Chipotle, *etc*. Of these selected tweets those with erroneous information or themes unrelated to the following four categories were excluded: (1) words that have double meanings, *e.g*., "save a lot" referring to a promotion instead of a store brand; (2) statements and opinions irrelevant to food choices (while tweets expressing desires to make a trip to a grocery or fast food restaurant were kept), e.g., "Meijer Christmas lights commercial this year seriously brings me to tears; " (3) activities that occurred in the past or were scheduled far into the future (while near future activities were kept), *e.g.*, "Just came from Walmart shopping for the kids; " (4) activities occurring in a place other than the food venue or activities not conducted by the Twitter user, e.g., "Someone brought me Subway and Pepsi!" To ensure the validity of the data, tweets with a vague context were judged at the researchers' discretion. After irrelevant tweets were excluded, the selected grocery shopping-themed tweets (termed "healthful tweets", $n = 61$) and fast food dining themed tweets (termed "unhealthful tweets", $n = 296$) were geocoded in the study region using their disclosed latitude and longitude coordinates. Table 1 shows examples of these tweets including the sending time and concurrent coordinates. Fig. 1a shows their geographic distribution within the study region.

The information about green retailers and fast food restaurants came from *InfoUSA*, a company that provides business data and marketing services. Specifically, the raw data containing all types of businesses in the study region were narrowed down by employing the NAICS (North American Industry Classification System) codes of supercenters, supermarkets/convenience stores, and fast food restaurants. Grocery retailers with fewer than five employees were excluded to ensure that all the stores mentioned in the study carried a wide range of fresh produce. Fast food restaurants were further distinguished from full-service restaurants and untraditional food retailers, such as snack shops, bakeries, bars, *etc*. that may also carry unhealthful food items. These selected green retailers ($n = 118$) and fast food restaurants ($n = 2106$) were geocoded in *ArcMap*, as shown in Fig. 1b.

Analyses of exposure to the local food environment

An individual's exposure to a particular food environment can be measured in many ways, including number of nearby stores (Jago, Baranowski, & Harris, 2006; Laraia, Siega-Riz, Kaufman, & Jones, 2004), distance to the nearest store (McEntee & Agyeman, 2010), and store density, which is calculated by a kernel function (Kestens, Lebel, Daniel, Thériault, & Pampalon, 2010). We chose to measure the quality of the food environment in terms of the numbers of green retailers and fast food restaurants within a food access zone, or specifically, within a buffered distance of the Twitter user's geo-tagged location. This measure provides a proxy of accessible healthful and unhealthful food sources within the individual's proximity. As tweets do not include individual differences in modal split, time budgets, and other socioeconomic variables that account for users' mobility for procuring food, we made use of fixed buffer distances (0.5 miles and 1 mile) to delineate food access zones.

Table 2 shows the major findings of the study in the chosen five-weekday period. In a 0.5-mile buffered access zone of reported tweets, the average number of green retailers was smaller than that of fast food restaurants. This result was due to the fact that the overall number of green retailers was smaller than that of fast food restaurants in this particular region. Between-group-comparison showed a significant difference in the average number of green retailers, 1.28 for healthful tweets and 0.54 for unhealthful tweets. This strong difference suggests that the presence of a healthier food environment may inspire a more nutritious way of eating. Interestingly, the difference in the average numbers of fast food restaurants in the two groups was not as prominent, with 19.50 for healthful tweets and 17.71 for unhealthful tweets. The slightly (and unexpectedly) larger number of fast food restaurants around healthful tweets suggests that an unhealthful food environment may not explain individuals' preference for fast food or discourage healthful food choices. Rather, the desire for fast food may be mediated by other non-spatial factors, such as financial constraints, food cues, taste of food, and lack of cooking skills that increase the motivation to patronize fast food restaurants (Driskell, Meckna, & Scales, 2006). These results were further verified by one-way ANOVA (Analysis of Variance) tests between the subgroups of tweets: A significant difference for number of green retailers was found in healthful tweets and unhealthful tweets (p-value < 0.05), while no significant difference for the number of fast food restaurants was found in the two subgroups (p-value $= 0.46$). The sensitivity analysis was conducted by expanding the food access zones to one mile, where similar results were identified.

Another dimension of this evaluation involved examining if a change in the number of fast food restaurants/green retailers would significantly influence individual food choices. Fig. 2 shows the percentages of healthful tweets versus unhealthful tweets in relation to the number of fast food restaurants (2a) and green retailers (2b) in a 0.5-mile access zone. In Fig. 2a, we could not identify an apparent pattern as to how the prevalence of fast food restaurants shapes individuals' eating habits, as suggested by Table 2. However, we found that even with low access to fast food restaurants ($n < 5$), a higher percentage of individuals (>85%) still tended to acquire fast food, suggesting that the choice to eat fast food is not restricted by spatial barriers but may be stimulated by personal preferences. In contrast, Fig. 2b shows that an increase of green retailers in an individual's vicinity increased the percentage of healthy eaters at a significant rate (313.9% from $n = 0$ to $n = 1$, −8.9% from $n = 1$ to $n = 2$, 110.5% from $n = 2$ to $n = 3$) and outnumbered the percentage of fast food eaters at $n = 3$ (52.6% > 47.4%), indicating that spatial clustering of green retailers may help foster a healthful food environment and eventually encourage food procurement in grocery stores where more nutritious and wholesome food items are available. This observation offers corroborating evidence that the presence of green retailers has a "marked effect" on improving the quality of diets for nearby residents (Wrigley, Warm, Margetts, & Whelan, 2002).

These findings take on a more nuanced meaning when the issue of obesity is involved. Examining the association between the local food environment and neighborhood obesity has been recognized

Table 1
Sample tweets indicating food-related activities in the study region.

Date	Time	Content	Latitude	Longitude
12/2/2013	10:43	I'm at @Walmart Supercenter (Whitehall, OH)	39.9543	−82.9004
12/2/2013	11:57	I can't beleive (believe) Giant Eagle has such poor customer service. Only a few registers open around lunch time and the lines are long!	40.0322	−82.9083
12/2/2013	12:50	Someone come(s) with me to Chipotle.	39.8742	−83.0486
12/2/2013	15:42	Sitting here at five guys alone.	39.9886	−83.0023
12/2/2013	18:25	Only thing getting me through studying is the food I'm going to get at Kroger.	39.9974	−83.0091
12/4/2013	18:51	Pulled my hair straightener, toothpaste, Tupperware out of my purse looking for my wallet in the checkout line at Kroger.	40.089	−82.8246
12/5/2013	12:14	Taco Bell for the third day in a row.	40.0318	−83.1417

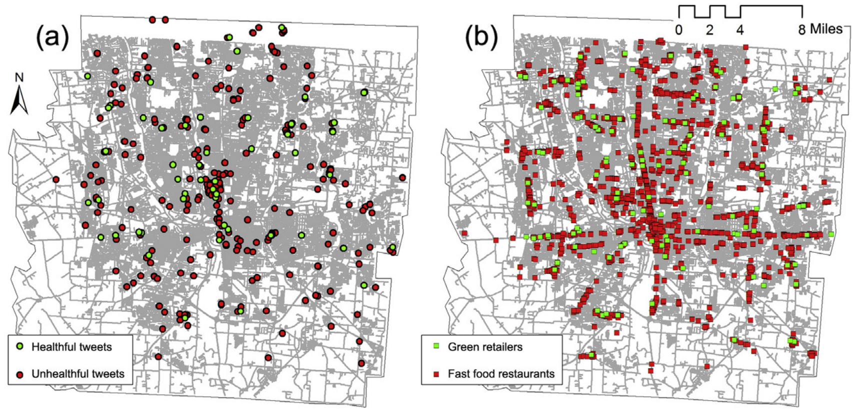

Fig. 1. The distributions of (a) selected tweets and (b) food outlets in the study region: Columbus, Ohio, USA.

as having a far-reaching effect on shaping strategies for improving community health. According to several cross-sectional studies, findings regarding the link between the obesogenic food environment and obesity deviated from our perception: there was no significant relationship between density of fast food restaurants and BMI (body mass index) of the local residents (Burdette & Whitaker, 2004; Jeffery, Baxter, McGuire, & Linde, 2006; Simmons, McKenzie, Eaton, Cox, Khan, Shaw, et al., 2005), and the presence of supermarkets was found to be inversely correlated with an overweight population (Morland, Diez Roux, & Wing, 2006). It is crucial to better understand the correlation between the community food environment and obesity (Holsten, 2008). This study provides a tentative explanation for this correlation, or even causality: an obesogenic food environment does not mediate unhealthy eating, which is considered a significant contributor to obesity. On the other hand, the influence of a healthful food environment is increased by expanding healthful food choices. Recognizing the significant role of green retailers is necessary for creating strategic food initiatives and land-use policies that will aid in community health awareness and will mitigate risk of obesity in the local food system. Instead of an effort to discourage consumption of fast food, public resources should be strategically allocated for improving the food environment, such as increasing the residents' accessibility to supermarkets and providing more healthful diet options that meet nutritional recommendations.

Conclusion

The emergence of social media and the infiltration of VGI have provided new channels for collecting human subject data and identifying their activity patterns. This study extracted individuals' diet choices from the widely popular social media platform, Twitter, and examined how access to the food environment around them influences their specific dietary choices. As tweets contain user-generated content and are retrievable by the public, they provide an effortless data collection process with spatiotemporally tagged information for the investigation of human activities related to food and health research. Although social media have generated a myriad of applications for public health monitoring (Chew & Eysenbach, 2010), this study is the first to focus specifically on food studies and to employ cloudsourcing of individual diet choices from cyberspace.

Social media highlights the foodscape experienced by individuals and its effects on food choices. To date, the majority of food access studies have been based on spatially generalized socioeconomic variables with very few studies investigating food exposure at a household level or around the workplace. Due to the fact that data collection poses a considerable challenge, examining individuals' food procurement and related health implications remains very rare (Kestens, Lebel, Daniel, Thériault, & Pampalon, 2010). Soliciting food-related tweets not only overcomes barriers of traditional data collection methods but more importantly facilitates the exploration of individuals' activity space in a geographically accurate and time-sensitive manner. This possibility is further utilized by geographic approaches to measuring the relationship to individuals' perceived food environment. The conclusion drawn from this study corroborates previous findings that exposure to a healthful food environment in an individual's immediate vicinity facilitates healthful choices while showing that an obesogenic food environment may not necessarily increase the likelihood that individuals will patronize fast food restaurants. A possible underlying

Table 2
The statistics of numbers of green retailers and fast food restaurants in a buffered food access zone of tweet locations. Tweets are categorized as healthful and unhealthful with respect to their content.

	0.5-Mile access zone				1-Mile access zone			
	# of green retailers		# of fast food restaurants		# of green retailers		# of fast food restaurants	
Descriptive statistics								
	Healthful	Unhealthful	Healthful	Unhealthful	Healthful	Unhealthful	Healthful	Unhealthful
N	61	296	61	296	61	296	61	296
Minimum	0	0	0	0	0	0	0	0
Maximum	3	3	63	87	5	6	139	143
Median	1	0	18	12	2	1	39	31
SE	0.13	0.05	1.73	1.05	0.17	0.08	3.77	2.01
Mean	1.28	0.54	19.5	17.71	2.31	1.46	44.49	41.59
ANOVA								
p-value	9.82E-10		0.46		1.08E-05			
							0.54	

Note: *N* is the number of tweets. SE is the standard error for mean. *p*-value is compared with $\alpha = 0.05$.

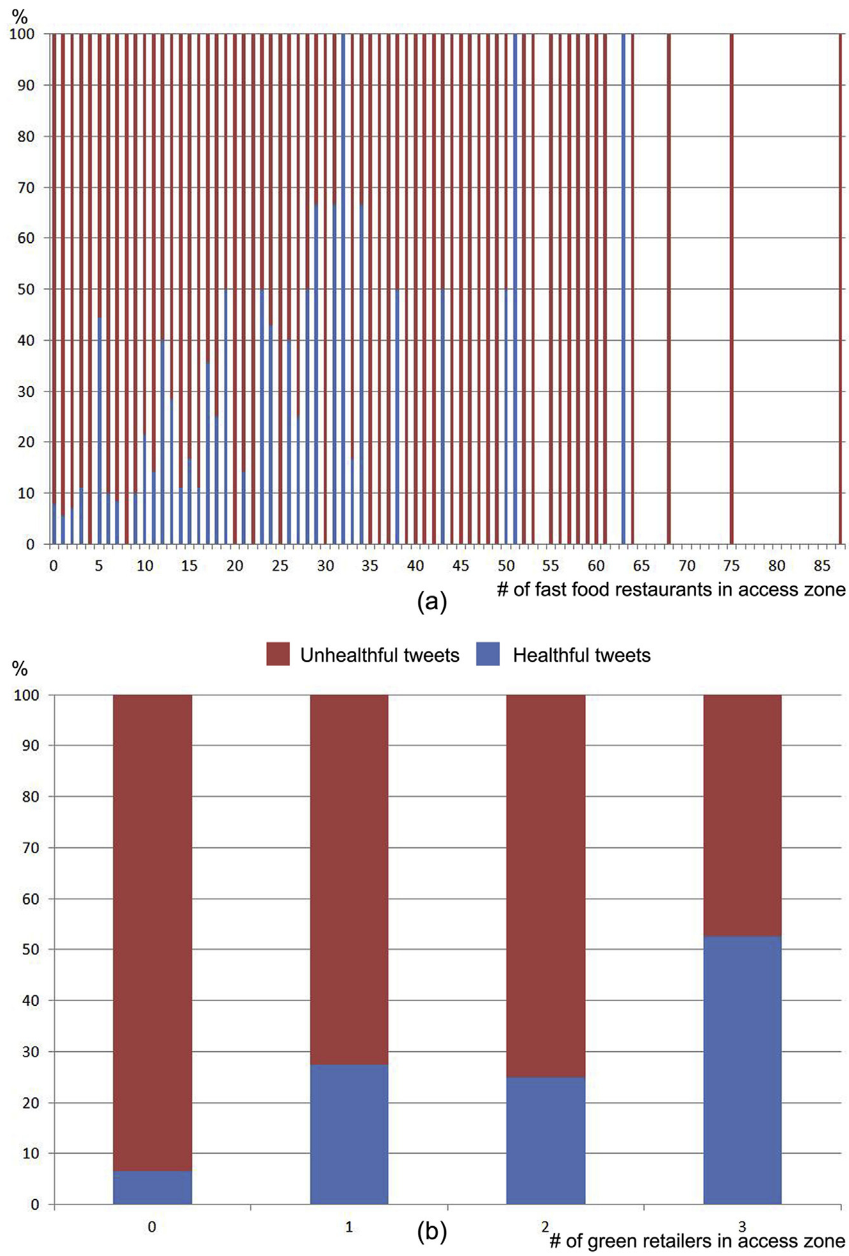

Fig. 2. Percentages of healthful tweets versus unhealthful tweets by (a) number of fast food restaurants and (b) number of green retailers in a 0.5-mile buffered food access zone of each tweet location.

explanation would be that preference for fast food is primarily dictated by non-spatial mediators, such as taste and cost, as reported by a nationwide study (Glanz, Basil, Maibach, Goldberg, & Snyder, 1998).

The real time data collection made possible by social media should be accessed to its full potential. Due to its reliance on user-generated content, VGI on social media allows for analysis at a high spatiotemporal resolution. This availability in space and accuracy in time allows for scrutinizing human activity patterns in unprecedented detail. An example is illustrated in Fig. 3, which shows the variation of healthful tweets and unhealthful tweets by time of day. It can be seen that fast food diet choices are most prevalent at mid day until late evening from 8 PM to 12 AM, while grocery shopping is more frequent from 5 PM to 9:30 PM. This use of VGI not only demonstrates a possible method for identifying daily activity patterns by the frequency of the appearance in a given time period but also provides data for retail analysis, such as the creation of more convenient store hours. As food retailers' hours are limited, how to make complete use of VGI for balancing the tradeoff between the needs of customers and the cost of store operation would be a potential avenue for future studies.

However, our study also incorporates several limitations that should be taken into consideration in future research. First, the demographic composition of Twitter users is not fully representative of the entire population. Only 15% of adults who use the Internet are Twitter users, and young adults and minority groups tend to be overrepresented in tweets (Smith & Brenner, 2012). Therefore, conclusions drawn from Twitter users should also account for the compensation for variations from census data (Mocanu, Baronchelli, Perra, Gonçalves, Zhang, & Vespignani, 2013). Second, the dichotomy of the diet-tweeting population is not warranted in that (1) the preference for tweeting in a grocery store may not be as frequent

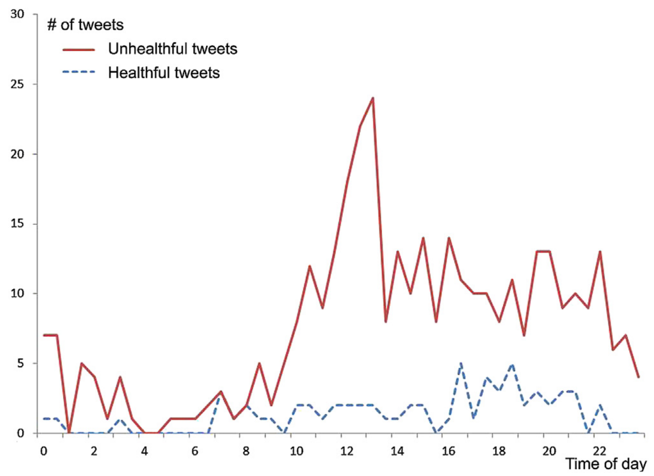

Fig. 3. The variation of healthful tweets and unhealthful tweets consolidated from a five-weekday period aggregated by every half an hour in a day.

as that in a fast food restaurant, as waiting and being seated allow for more time to tweet; (2) patronizing grocery stores does not guarantee purchase of healthful food items; (3) and fast food restaurants may offer healthful food options. Studies on the actual frequency of conducting target activities should be performed to generate more convincing evidence. Third, food choices are not only spatially constrained but also mediated by socioeconomic variables, interpersonal influences, time budgets, and individual preferences that are difficult to infer from tweets (Chen & Kwan, 2012; Glanz, Basil, Maibach, Goldberg, & Snyder, 1998; Glanz, Kristal, Sorensen, Palombo, Heimendinger, & Probart, 1993; Walker, Fryer, Butler, Keane, Kriska, & Burke, 2011; Walker, Block, & Kawachi, 2012). A possible direction to offset these uncertainties is through sentiment analysis of the contextual information in tweets (Mitchell, Frank, Harris, Dodds, & Danforth, 2013). Fourth, as a large portion of tweets contain very succinct messages, compounded with typos, puns, and irrelevant content, filtering the information for the appropriate data demands a significant amount of time and labor (only 0.43% of tweets were considered relevant for our study). More efficient utilization of frontier technologies in semantic analysis to assist in data mining would be beneficial for expanding the sampling period and enlarging the sampling size. Last, in urban food access studies, the maximum values for walkable distance to food retailers were loosely defined as ranging from 0.25 miles to two miles (Block & Kouba, 2006; Jeffery, Baxter, McGuire, & Linde, 2006; Mulangu & Clark, 2012), and some results were interpreted by non-spatial units, such as the percentage of travel cost in food budget (Hallett & McDermott, 2011). Based on the relatively small scale of our study, results showed that a buffer distance of less than 0.5 miles failed to capture a sufficient number of food outlets, while a distance of more than two miles showed little variation. How to standardize this cut-off value as a benchmark of sufficient food access remains to be evaluated.

References

Algert, J., Agrawal, A., & Lewis, S. (2006). Disparities in access to fresh produce in low-income neighborhoods in Los Angeles. *American Journal of Preventive Medicine, 30*(5), 365–370.
Apparicio, P., Cloutier, M. S., & Shearmur, R. (2007). The case of Montréal's missing food deserts: evaluation of accessibility to food supermarkets. *International Journal of Health Geographics, 6*(4).
Arribas-Bel, D. (2014). Accidental, open and everywhere: emerging data sources for the understanding of cities. *Applied Geography, 49*, 45–53.
Ball, K., Crawford, D., & Mishra, G. (2006). Socio-economic inequalities in women's fruit and vegetable intakes: a multilevel study of individual, social and environmental mediators. *Public Health Nutrition, 9*(5), 623–630.
Ball, K., Timperio, A., & Crawford, D. (2009). Neighbourhood socioeconomic inequalities in food access and affordability. *Health & Place, 15*(2), 578–585.
Berg, N., & Murdoch, J. (2008). Access to grocery stores in Dallas. *International Journal of Behavioural and Healthcare Research, 1*(1), 22–37.
Burdette, H., & Whitaker, R. (2004). Neighborhood playgrounds, fast food restaurants, and crime: relationships to overweight in low-income preschool children. *Preventive Medicine Preventive Medicine, 38*(1), 57–63.
Block, D., & Kouba, J. (2006). A comparison of the availability and affordability of a market basket in two communities in the Chicago area. *Public Health Nutrition, 9*(7), 837–845.
Chaix, B., Meline, J., Duncan, S., Merrien, C., Karusisi, N., Perchoux, C., et al. (2013). GPS tracking in neighborhood and health studies: a step forward for environmental exposure assessment, a step backward for causal inference? *Health & Place, 21*, 46–51.
Chen, X., & Clark, J. (2013). Interactive three-dimensional geovisualization of space-time access to food. *Applied Geography, 43*, 81–86.
Chen, X., & Kwan, M. (2012). Choice set formation with multiple flexible activities under space-time constraints. *International Journal of Geographical Information Science, 26*(5), 941–961.
Chew, C., & Eysenbach, G. (2010). Pandemics in the age of Twitter: content analysis of tweets during the 2009 H1N1 outbreak. *Plos One, 5*(11).
Clarke, G., Eyre, H., & Guy, C. (2002). Deriving indicators of access to food retail provision in British cities: studies of Cardiff, Leeds and Bradford. *Urban Studies, 39*(11), 2041–2060.
Drewnowski, A., & Specter, S. (2004). Poverty and obesity: the role of energy density and energy costs. *The American Journal of Clinical Nutrition, 79*(1), 6–16.
Driskell, J., Meckna, B., & Scales, N. (2006). Differences exist in the eating habits of university men and women at fast food restaurants. *Nutrition Research, 26*(10), 524–530.
Ehrenberg, R. (2012). The next epidemics will be tweeted: twitter posts tracked with official data during cholera outbreak. *Science News, 181*(4).
Ferriday, D., & Brunstrom, J. (2008). How does food-cue exposure lead to larger meal sizes? *The British Journal of Nutrition, 100*(6), 1325–1332.
Glanz, K., Basil, M., Maibach, E., Goldberg, J., & Snyder, D. (1998). Why Americans eat what they do: taste, nutrition, cost, convenience, and weight control concerns as influences on food consumption. *Journal of the American Dietetic Association, 98*(10), 1118–1126.
Glanz, K., Sallis, S., Saelens, B., & Frank, L. (2007). Nutrition environment measures survey in stores (NEMS-S): development and evaluation. *American Journal of Preventive Medicine, 32*(4), 282–289.
Glanz, K., Kristal, A. R., Sorensen, G., Palombo, R., Heimendinger, J., & Probart, C. (1993). Development and validation of measures of psychosocial factors influencing fat- and fiber-related dietary behavior. *Preventive Medicine, 22*(3), 373–387.
Guy, C., & David, G. (2004). Measuring physical access to 'healthy foods' in areas of social deprivation: a case study in Cardiff. *International Journal of Consumer Studies, 28*(3), 222–234.
Hallett, L., & McDermott, D. (2011). Quantifying the extent and cost of food deserts in Lawrence, Kansas, USA. *Applied Geography, 31*(4), 1210–1215.
Helling, A., & Sawicki, D. (2003). Race and residential accessibility to shopping and services. *Housing Policy Debate, 14*(1–2), 69–101.
Hunker, H. (2000). *Columbus, Ohio: A personal geography*. Columbus, OH: Ohio State University Press.
Holsten, J. (2008). Obesity and the community food environment: a systematic review. *Public Health Nutrition, 12*(3), 397–405.
Jago, R., Baranowski, T., & Harris, M. (2006). Relationships between GIS environmental features and adolescent male physical activity: GIS coding differences. *Journal of Physical Activity & Health, 3*(2), 230–242.
Jeffery, R., Baxter, J., McGuire, M., & Linde, J. (2006). Are fast food restaurants an environmental risk factor for obesity? *International Journal of Behavioral Nutrition & Physical Activity, 3*(2).
Kestens, Y., Lebel, A., Daniel, M., Thériault, M., & Pampalon, R. (2010). Using experienced activity spaces to measure foodscape exposure. *Health & Place, 16*(6), 1094–1103.
Laraia, B., Siega-Riz, A., Kaufman, J., & Jones, S. (2004). Proximity of supermarkets is positively associated with diet quality index for pregnancy. *Preventive Medicine, 39*(5), 869–875.
Larsen, K., & Gilliland, J. (2008). Mapping the evolution of 'food deserts' in a Canadian city: supermarket accessibility in London, Ontario, 1961–2005. *International Journal of Health Geographics, 7*(16).
McEntee, J., & Agyeman, J. (2010). Towards the development of a GIS method for identifying rural food deserts: geographic access in Vermont, USA. *Applied Geography, 30*(1), 165–176.
McKinnon, R., Reedy, J., Morrissette, M., Lytle, L., & Yaroch, A. (2009). Measures of the food environment: a compilation of the literature, 1990–2007. *American Journal of Preventive Medicine, 36*(4), 124–133.
Miller, H. (2007). Place-based versus people-based geographic information science. *Geography Compass, 1*(3), 503–535.
Mitchell, L., Frank, M., Harris, K., Dodds, P., & Danforth, C. (2013). The geography of happiness: connecting twitter sentiment and expression, demographics, and objective characteristics of place. *PloS One, 8*(5).
Mocanu, D., Baronchelli, A., Perra, N., Gonçalves, B., Zhang, Q., & Vespignani, A. (2013). The Twitter of Babel: mapping world languages through microblogging platforms. *PloS One, 8*(4).
Moore, L., & Diez Roux, A. (2006). Associations of neighborhood characteristics with the location and type of food stores. *American Journal of Public Health, 96*(2), 325–331.

Morland, K., Diez Roux, A., & Wing, S. (2006). Supermarkets, other food stores, and obesity: the atherosclerosis risk in communities study. *American Journal of Preventive Medicine, 30*(4), 333–339.

Mulangu, F., & Clark, J. (2012). Identifying and measuring food deserts in rural Ohio. *Journal of Extension, 50*(3).

Kwan, M. P., Casas, I., & Schmitz, B. (2004). Protection of geoprivacy and accuracy of spatial information: how effective are geographical masks? *Cartographica, 39*(2), 15–27.

Paez, A., Mercado, R., Farber, S., Morency, C., & Roorda, M. (2010). Relative accessibility deprivation indicators for urban settings: definitions and application to food deserts in Montreal. *Urban Studies, 47*(7), 1415–1438.

Pearce, J., Witten, K., & Bartie, P. (2006). Neighbourhoods and health: a GIS approach to measuring community resource accessibility. *Journal of Epidemiology & Community Health, 60*(5), 389–395.

Powell, L., Slater, S., Mirtcheva, D., Bao, Y., & Chaloupka, F. (2007). Food store availability and neighborhood characteristics in the United States. *Preventive Medicine Preventive Medicine, 44*(3), 189–195.

Raja, S., Ma, C., & Yadav, P. (2008). Beyond food deserts: measuring and mapping racial disparities in neighborhood food environments. *Journal of Planning Education and Research, 27*(4), 469–482.

Reisig, V., & Hobbiss, A. (2000). Food deserts and how to tackle them: a study of one city's approach. *Health Education Journal, 59*(2), 137–149.

Robillard, J., Johnson, T., Hennessey, C., Beattie, B., & Illes, J. (2013). Aging 2.0: health information about dementia on Twitter. *Plos One Plos One, 8*(7).

Shaw, H. (2006). Food deserts: towards the development of a classification. *Geografiska Annaler Series B: Human Geography, 88*(2), 231–247.

Simmons, D., McKenzie, A., Eaton, S., Cox, N., Khan, M., Shaw, J., et al. (2005). Choice and availability of takeaway and restaurant food is not related to the prevalence of adult obesity in rural communities in Australia. *International Journal of Obesity, 29*(6), 703–710.

Smith, A., & Brenner, J. (2012). *Twitter use 2012*. Pew Research Institute (technical report) http://pewinternet.org/Reports/2012/Twitter-Use-2012.aspx Accessed 31.01.14.

Smith, C., & Morton, L. (2009). Rural food deserts: low-income perspectives on food access in Minnesota and Iowa. *Journal of Nutrition Education and Behavior Journal, 41*(3), 176–187.

Twitter. (2013). http://www.statisticbrain.com/twitter-statistics/, Accessed 31.01.14.

USDA (United States Department of Agriculture). (2009). *Access to affordable and nutritious food: measuring and understanding food deserts and their consequences*. http://www.ers.usda.gov/Publications/AP/AP036/ Accessed 31.01.14.

Walker, R. E., Fryer, C., Butler, J., Keane, C., Kriska, A., & Burke, J. G. (2011). Factors influencing food buying practices in residents of a low-income food desert and a low-income food oasis. *Journal of Mixed Methods Research, 5*(3), 247–267.

Walker, R., Block, J., & Kawachi, I. (2012). Do residents of food deserts express different food buying preferences compared to residents of food oases? A mixed-methods analysis. *International Journal of Behavioral Nutrition and Physical Activity, 9*(41).

Wrigley, N., Warm, D., Margetts, B., & Whelan, A. (2002). Assessing the impact of improved retail access on diet in a 'food desert': a preliminary report. *Urban Studies, 39*(11), 2061–2082.