The Uncertain Geographic Context Problem

Mei-Po Kwan

Department of Geography, University of California, Berkeley

Any study that examines the effects of area-based attributes on individual behaviors or outcomes faces another fundamental methodological problem besides the modifiable areal unit problem (MAUP). It is the problem that results about these effects can be affected by how contextual units or neighborhoods are geographically delineated and the extent to which these areal units deviate from the true geographic context. The problem arises because of the spatial uncertainty in the actual areas that exert the contextual influences under study and the temporal uncertainty in the timing and duration in which individuals experienced these contextual influences. Using neighborhood effects and environmental health research as a point of departure, this article clarifies the nature and sources of this problem, which is referred to as the *uncertain geographic context problem* (UGCoP). It highlights some of the inferential errors that the UGCoP might cause and discusses some means for mitigating the problem. It reviews recent studies to show that both contextual variables and research findings are sensitive to different delineations of contextual units. The article argues that the UGCoP is a problem as fundamental as the MAUP but is a different kind of problem. Future research needs to pay explicit attention to its potential confounding effects on research results and to methods for mitigating the problem. *Key Words: contextual uncertainty, environmental health, neighborhood effects, uncertain geographic context problem, UGCoP.*

任何探讨以区域为基础的属性对个体的行为或结果之影响的研究，除了可变面元问题（MAUP），还都面临着另一个基本方法的问题。该问题影响了这些研究结果，使它们受到背景单位或社区是如何被地理划定的影响，并受到这些面元偏离真实的地理环境的程度的影响。问题的产生是因为在这些实际区域的空间不确定性的情景，以及个人在经历这些情景影响时，在时间和持续时间上的不确定性的影响。使用邻里效应和环境健康研究作为出发点，本文阐明了这个被称为不确定的地理环境问题（UGCoP）的性质和来源。它突出了一些UGCoP可能导致的推理错误，并讨论了减轻这个问题的一些手段。它回顾了最近的研究，揭示了情景变量和研究成果对情景单位的不同划定是敏感的。文章认为，UGCoP是与MAUP类似的根本性问题，但又是不同类的问题。未来的研究需要明确注意其潜在的对研究结果的复杂影响，并关注减轻这一问题的方法。*关键词：情景的不确定性，环境卫生，邻里效应，地理环境不确定性问题，UGCoP。*

En cualquier estudio que examine los efectos que tienen atributos basados en área sobre conductas o logros individuales se tiene que enfrentar un problema metodológico fundamental aparte del problema de la unidad espacial modificable (PUEM). El inconveniente es que los resultados acerca de estos efectos pueden afectarse por la manera como las unidades contextuales o vecindarios sean delineados geográficamente y el grado en que dichas unidades espaciales se apartan del contexto geográfico verdadero. El problema surge debido a la incertidumbre espacial en las áreas reales que ejercen las influencias contextuales bajo estudio y la incertidumbre temporal presente en el cronograma y duración con que los individuos experimentaron estas influencias contextuales. Utilizando los efectos de vecindad y la investigación sobre salubridad ambiental como punto de partida, este artículo hace claridad sobre la naturaleza y orígenes de este problema, que se denomina *problema de incertidumbre del contexto geográfico* (PICoG). Se destacan algunos de los errores de inferencia que el PICoG podría ocasionar y se discuten algunos medios para mitigar el problema. En el artículo se revisan también estudios recientes para mostrar que tanto las variables contextuales como los hallazgos de la investigación son sensibles a las diferentes demarcaciones de las unidades contextuales. En el artículo se arguye que el PICoG es un problema tan fundamental como el PUEM, aunque se trata de un tipo diferente de problema. La futura investigación sobre el particular deberá poner atención explícita a los potenciales efectos desorientadores de los resultados que se logren y a los métodos para mitigar el problema. *Palabras clave: incertidumbre contextual, salubridad ambiental, efectos de vecindad, problema de incertidumbre del contexto geográfico.*

Any study that examines the effects of area-based attributes on individual behaviors or outcomes faces two fundamental methodological problems. One is that results about these effects can be affected by the zoning scheme or geographic scale of the areal units used (Openshaw 1984; Fotheringham and

Annals of the Association of American Geographers, 102(5) 2012, pp. 958–968 © 2012 by Association of American Geographers
Initial submission, January 2011; revised submission, September 2011; final acceptance, November 2011
Published by Taylor & Francis, LLC.

Wong 1991). This is the well-known modifiable areal unit problem (MAUP), which has received much attention to date. The other problem, however, has received much less attention. This is the problem that findings about the effects of area-based attributes could be affected by how contextual units or neighborhoods are geographically delineated and the extent to which these areal units deviate from the "true causally relevant" geographic context (the precise spatial configuration of which is unknown in most studies to date; Diez-Roux and Mair 2010, 134). This problem is referred to as the *uncertain geographic context problem* (UGCoP) in this article. It arises because of the spatial uncertainty in the actual areas that exert contextual influences on the individuals being studied and the temporal uncertainty in the timing and duration in which individuals experienced these contextual influences. The UGCoP is a significant methodological problem because it means that analytical results can be different for different delineations of contextual units even if everything else is the same. It is perhaps a major reason why research findings concerning the effects of social and physical environments on health behaviors and outcomes are often inconsistent, given that past studies on the same issue (e.g., obesity) often used different contextual units (e.g., Inagami, Cohen, and Finch 2007; Black and Macinko 2008; Wilks et al. 2010).

This article argues that the UGCoP is a problem as fundamental as the MAUP for any study that uses area-based contextual attributes, but it is a different kind of problem because it is not due to the use of different zonal schemes or spatial scales for area-based variables. Because methods for addressing the MAUP do not automatically solve the UGCoP, future research needs to pay explicit attention to its potential confounding effects on research results and to methods for mitigating the problem. The article clarifies the nature and sources of the UGCoP and highlights the inferential errors it might cause. It reviews recent studies to show that both contextual variables and research findings are sensitive to different delineations of contextual units. It suggests some means for addressing the problem, such as delineating more appropriate contextual units based on people's actual or potential activity spaces. The article uses neighborhood effects and environmental health research as a point of departure to highlight the methodological challenges the UGCoP poses to health research in particular and to geographic research in general.[1] It addresses a fundamental methodological problem for any study that uses area-based attributes as explanatory variables.

Contextual Uncertainty: The Spatial Dimension

Studies that examine the effect of contextual or environmental influences (e.g., neighborhood physical and social features) on health using ecological designs often begin by constructing a conceptual model that specifies the causal pathways among the contextual attributes (or variables) and health outcomes (Diez-Roux and Mair 2010). Based on the conceptual model, contextual units or geographic areas for evaluating individual exposure to these contextual influences are then identified. After these units are defined, values of the relevant contextual variables (e.g., neighborhood deprivation) for these contextual units are derived and used as indicators of the exposure individuals in particular contextual units experienced. Effects of the contextual variables on health are finally evaluated using appropriate statistical models (e.g., multilevel models). In this process, two important sources of contextual uncertainty contribute to the UGCoP. One is the uncertainty in the spatial configuration of the appropriate contextual units for assessing the influence of environmental variables on health outcomes, and the other is the uncertainty about the timing and duration to which individuals are actually exposed to these contextual influences (the temporal dimension of contextual uncertainty is discussed in the next section under the rubric of dynamics of geographic context).

Health researchers normally have little or no prior knowledge about the precise spatial configuration and boundary of the geographic area that, through its physical or social characteristics, has significant influence on an individual's health. The "true causally relevant" geographic context is thus unknown in most studies to date (Diez-Roux and Mair 2010, 134). A common practice in the past has been to use residential neighborhoods—operationalized as static administrative areas such as census tracts or postal code areas or buffer areas around individuals' home addresses or centroids of their home census tracts—as contextual units. Leal and Chaix (2011), for instance, observed that 90 percent of studies on environmental influences on cardiometabolic risk factors used the residential neighborhood as the contextual unit. These units not only are convenient but often are the only viable option because available data of censuses and surveys that can be used to derive contextual measures are tied to them.

But residential neighborhoods might not accurately represent the actual areas that exert contextual influences on the health outcome under study (Diez-Roux 1998; Cummins 2007; Matthews 2008; Chaix 2009;

Kwan 2009). For instance, adolescent risk behavior like substance use might be influenced not only by socioeconomic deprivation in the residential neighborhood but also by interactions with friends and peers in various nonresidential contexts (e.g., schools and places for various leisure activities). For working adults, opportunities for physical activities and the quality of food near the workplace could also have important effects on their health. The boundaries of these multiple contexts are often difficult to clearly delineate; even when it is possible, some of them might not be continuous in geographic space (i.e., one contextual unit might consist of several discrete geographic areas) and thus cannot be represented or analyzed in any simple manner even using advanced geographic information systems (GIS) methods (Wiehe et al. 2008).

Further, social contexts such as families, friends, or peers are not in themselves geographically defined and thus cannot be easily delineated as geographic areas with precise boundaries (Diez-Roux 2001; Macintyre, Ellaway, and Cummins 2002). Delineations of contextual units in these cases need to take into account how social networks constituted through people's routine activities and social interactions express themselves in geographic space (Grannis 2009). In other cases, neighborhoods defined on the basis of people's perceptions might be more relevant. The perceived neighborhood for different individuals might not coincide with or might even deviate significantly from the administratively defined home neighborhood or people's activity space, however. For instance, Basta, Richmond, and Wiebe (2010) found that participants' perceived neighborhoods did not correspond to the boundaries of the home census tracts, and time they spent in close proximity to alcohol outlets during their daily activities was not correlated with the prevalence of alcohol outlets in the census tract of their residence. Vallée et al. (2010) found that over 80 percent of the participants have an activity space larger than their perceived neighborhood. This means that even when the boundaries of perceived neighborhoods are identified using appropriate procedures, a considerable portion of people's normal daily activities could still fall outside these boundaries. In these cases, perceived neighborhoods might not correspond well with, and thus are not good proxies for, true geographic contexts.

Besides administrative areas and perceived neighborhoods, other geographic areas have been used as proxies to the true geographic context. For instance, studies on the effect of neighborhood features such as land-use mix and residential density on people's physical activity or body weight have defined neighborhood around each participant's home as a 1-km or 3-km circular zone (Berke et al. 2007), as a 1-km road network buffer (Frank et al. 2005), as a 0.5-mile radius or a ten-minute walk from the respondent's home for some variables, and as a 10-mile radius or a twenty-minute drive from the respondent's home for several other variables (Brownson et al. 2004). It is far from clear, however, which of these areal units appropriately represents the areal extent and spatial configuration of the true geographic context. The mixed results of past studies on neighborhood effects (e.g., neighborhood income inequality and racial composition) on health (e.g., obesity) could thus be partly due to the different neighborhood delineations used (Black and Macinko 2008).

Finally, the relevant contextual unit might vary depending on the population groups under study and according to different factors and processes hypothesized to influence the health outcome in question (Subramanian, Jones, and Duncan 2003). For instance, results in Oliver, Schuurman, and Hall (2007) indicated that larger or smaller contextual units might be more appropriate for different types of built environments and population groups (e.g., smaller for elderly people because of their lower out-of-home mobility). Further, some contextual influences might operate in the block on which a person resides, some might operate in a larger area around the block, some might operate near the person's workplace or school, and still others might exert their influence in the area in which specific types of stores or institutions are located (Diez-Roux 2001; Macintyre and Ellaway 2003). The multilevel and multiscale nature of contextual influences greatly complicates the task of accurately delineating the appropriate contextual units, which could be nested or overlapped in a complex manner. Part of the uncertainty in the spatial configuration and boundaries of contextual units arises from the dynamic characteristics of individuals and contextual influences (Gatrell 2011).

Dynamics of Geographic Context

People move around to undertake their daily activities. They often traverse the boundaries of multiple neighborhoods during the course of a day and come under the influence of many different neighborhood contexts besides their residential neighborhoods (Sampson, Morenoff, and Gannon-Rowley 2002; Chaix 2009). The majority of physical and social resources they use (which affect their health and well-being) might be located outside of or far from their home neighborhoods. Geographers have long observed the spatial and temporal variability of people's daily activities (including those performed in evening hours), and it was also noted that individuals of different social groups tend to have distinctive activity patterns in space–time (e.g., Hanson and Hanson 1981; Kwan 2000; Lee and Kwan 2011).

Recent studies that collected detailed data about people's out-of-home activities and travel routes over the course of one to many days using Global Positioning System (GPS) or other location-aware devices provide further evidence about where and when people spent time in their daily life. Elgethun et al. (2003), for instance, found that participants (children two to eight years old) on average spent most of their time inside schools on weekdays, while spending most of their time in establishments like restaurants and cinemas on weekend days. Basta, Richmond, and Wiebe (2010) observed that half of the sampled participants (fifteen to nineteen years old) spent 92 percent of their time outside of their neighborhood. Wiehe et al. (2008) found that participants (female adolescents) spent one third of their time in locations more than 1 km from home, which is the distance used in many previous studies for defining neighborhood. This means that the participants spent a considerable amount of time in their daily lives outside of what has conventionally been defined as geographic context or neighborhood. The study also found considerable day-to-day variability in participants' activity locations besides their variability by time of day (Wiehe et al. 2008). The daily and day-to-day variability in human activity locations not only raises concerns about using conventional static contextual units in health research but also calls into question the appropriateness of the notion of daytime population, which does not take temporal variability over the course of a day into account.

As Gatrell (2011) cogently argued, exposure to health risks, spread of diseases, and use of health care facilities are inextricably connected to human movements at various spatial and temporal scales. The spatial and temporal variability of human activities thus has significant implications for any study that examines the effect of contextual influences on health. It means that people's activities (and thus exposures) do not take place at one time point and wholly within any conventionally defined neighborhood. Their use of different physical resources and their social interactions with friends, peers, and others might take place at different times of the day and in disparate geographic areas outside of their home neighborhoods (Kwan 2009). The neighborhood of residence is only one of the places people spend their time, and it might not adequately capture people's exposure to relevant contextual influences. Further, besides moving around to undertake their daily activities, people also move around over time. They could change their residence in the same city (residential mobility) or move to another (migration). As a result of moving to different neighborhoods, people's exposure to environmental influences might also change over time. A study on people's exposure to carcinogenic risk factors, for instance, needs to consider their residential history (in addition to individual factors including family predisposition), as knowing where and for how long a person has lived in the past might help more accurately estimate his or her cumulative or lifetime exposure to radioactive substances (Löytönen 1998).

Contextual influences can vary over space and time in a highly complex manner. They might vary with different temporal patterns or time frames. As people move through the changing pollution field over time during the day, for instance, their exposure to traffic-related air pollution also changes (Gulliver and Briggs 2005). Some environmental influences could change over the twenty-four-hour period of a day (e.g., pollutants from truck traffic), and some might change over the seasons. The physical and social characteristics of neighborhoods can also change over time (Entwisle 2007). Population composition and local social ties might change as a result of residential mobility and migration. Government and people's actions could change the physical features and health facilities in a neighborhood over time. When environmental or neighborhood influences have considerable spatial and temporal variability, their health impact often cannot be adequately assessed using data for just one time point (Setton et al. 2010). It might also be difficult to identify which portion of them is causally relevant to a particular individual in relation to the person's daily movement in the study area.

Further, most studies to date assume that the effects of contextual influences on health are most appropriately assessed using data collected at or around the same time point (Entwisle 2007). There might be a variety of response lags that mediate the causal pathways between contextual factors and health outcomes, however. The outcome variable at time point t, for instance, might be determined by the value of a contextual variable at a particular time point or period of time before t (Wheaton and Clarke 2003). For instance, there is some evidence that variation in people's health behaviors and outcomes is related to their exposure to neighborhood characteristics during childhood and adolescence (e.g., Monden, Van Lenthe, and Mackenbach 2006). In addition, the outcome variable might not be determined by the value of the contextual variables at a particular time point before *t* but by their cumulative effect over a period of time before *t*. As a result of these and other cause–effect lags, there could be considerable uncertainty regarding the best time point or time period for deriving the values of the contextual attributes. The lack of significant association between the contextual variables and the outcome variable in any study might be due to a failure to account for this aspect of contextual dynamics (e.g., using the wrong time lag or using a particular time point instead of considering the cumulative effect of a contextual factor).

Inferential Challenges Posed by the UGCoP

The spatial uncertainty and dynamics of geographic context associated with the UGCoP greatly complicate any examination of the effect of contextual influences on health. The error of misspecifying the true geographic context might lead to inconsistent results and inferential errors. Consider a case in which the outcome variable Y (e.g., body mass index) is hypothesized to be determined by n contextual variables X_i (e.g., street network density, land-use mix, and social disadvantage) after taking into account all relevant individual- and household-level factors. Now suppose that significant association (either positive or negative) is found between one or more contextual variables and the outcome variable, assuming that there is no "misspecification of the model at the individual level" (Diez-Roux 1998, 219). This result is normally interpreted in a straightforward manner with few qualifications on the possible confounding effects of the UGCoP. For instance, a study might conclude that neighborhood physical features that encourage physical activity were associated with decreased body mass index (e.g., Berke et al. 2007). The best possibility for this result is that it is true, which not only means that the contextual variables worked as hypothesized in the causal model; it also means that the areal extent and spatial configuration of the true geographic context were correctly identified and used to derive the contextual variables in the study. Because in most cases the true geographic context is not known, the researcher cannot be certain that it had been used in the study. This uncertainty implies that the significant relationships observed might be false positives: There was actually no association between the contextual variables and the outcome variable as hypothesized, but significant association between them was still observed.

Table 1. Inferential errors due to the uncertain geographic context problem

True state of contextual effect	Observed state of contextual effect: Has effect	Observed state of contextual effect: No effect
Has effect	Contextual units correct	Contextual units incorrect
	Correct inference	False negatives (obscured contextual effect)
No effect	Contextual units incorrect	Contextual units correct
	False positives (spurious association)	Correct inference

Table 1 shows how true or incorrect contextual units could lead to different inferential errors. Inferential errors in the form of false positives can occur when the contextual variable actually has no effect on the outcome variable but significant association between them is observed. This situation is called *spurious association* in the parlance of statistical inference and is similar to a Type I error in hypothesis testing (the null hypothesis is rejected although it is true). It can arise for two different reasons. First, it can occur purely by chance even when the causal pathway hypothesized between the contextual variable and the outcome variable is illusionary. Second, the erroneously defined contextual units might have introduced variations in the contextual variable such that positive association between the contextual variable and the outcome variable is observed, even when there is actually no association between them. This type of inferential error could considerably confound research results.

Now suppose that no significant association is observed between the contextual variable and the outcome variable. This result is often interpreted in a straightforward manner with few qualifications, and the best possibility is that it is true: The contextual variable did not work as hypothesized in the causal model. Because the true geographic context is often not known, however, and the researcher cannot be sure that it has been used in the study, failure to observe a significant relationship between the contextual and outcome variables might be due to other reasons. There might be significant association between the contextual variable and the outcome variable as hypothesized, but such association was not observed. In the parlance of statistical inference, this type of error is similar to a Type II error in hypothesis testing (failure to reject the null hypothesis when it is false). One reason why this failure (or false negative) occurs is that the contextual variable might have been misspecified and thus does not correctly capture the true contextual effect. Another reason for the error is that the spatial extent and configuration of the true geographic context were not correctly identified and used in the study. In addition, the contextual influence might be characterized with wrong temporal attributes (e.g., incorrect time point, time lag, or duration). As a result, the effect of the contextual variable on the outcome variable was obscured by the erroneously defined contextual units or inappropriate temporal characterization of the contextual influence.

These two types of error arising from the misspecification of contextual units or inappropriate temporal characterization of the contextual influence might significantly confound research results. For instance, Spielman and Yoo (2009) used simulation experiments to show that linear models tend to underestimate the effects of contextual influences on health outcomes when the size of the true geographic context is underestimated (and vice versa). The study also showed that variation in the characteristics of the population group being studied and the study area can pose a significant problem for inference about neighborhood effects. Kwan et al. (2009) found significant differences in the size and shape of three different delineations of geographic context: two delineations of activity space (the standard deviational ellipse and the kernel density surface) and the home census tract. The study observed that for certain gender and racial groups, neighborhood effects based on people's home census tracts tend to overestimate their actual exposure to social disadvantage (because characteristics of the nonresidential neighborhoods people visit might mitigate the disadvantage they experience in their residential neighborhood). Further, Troped et al. (2010) examined associations among five physical features within 1-km road network buffers of participants' homes and workplaces and the amount of moderate to vigorous physical activity. The study found that three features around the participants' homes were associated with their physical activity near their homes, and two features around their workplaces were associated with their physical activity around their workplaces. None of the five features, however, showed associations with participants' total physical activity. The study not only shows that people's physical activity might vary according to where they are but also suggests that a study that uses only participants' home neighborhoods as the contextual unit might not find any association between its physical features and participants' body mass indexes, because body mass index depends on total physical activity, not just activity around one's home or workplace.

Addressing the UGCoP

As argued in this article, the UGCoP might introduce inferential errors and confound research results in studies that examine the effects of area-based attributes on individual behaviors or outcomes. The problem arises because of our limited knowledge of the precise spatial and temporal characteristics of the true geographic context. The main difficulty it poses is that we cannot tell whether our results are true or confounded and, if confounded, which type of error is involved and to what extent it has obfuscated the results. The UGCoP is a problem as fundamental as the MAUP, but it is a different kind of problem because it is not due to the use of different zonal schemes or spatial scales for deriving area-based variables. These two problems are not necessarily related to each other and can both be present in a particular study. Methods for addressing one of them might not automatically solve the other. For instance, using the best zoning scheme or spatial scale does not help us identify the true geographic context or characterize the temporal attributes of contextual influences, but using delineations of geographic context that capture people's movement in space–time seems to mitigate both the UGCoP and the MAUP (Kwan and Weber 2008).

Given the potential confounding effect of the UGCoP on research results, it is important that future research takes the problem seriously and considers steps to mitigate its impact when using area-based contextual variables. An important initial step is to develop an adequate theoretical model for taking spatial and temporal contextual uncertainties into account (Macintyre, Ellaway, and Cummins 2002). After constructing such a dynamic conceptualization of contextual influences, an explicit statement about what contextual units will be used, how well they approximate the true geographic context, and what temporal attributes of individuals and contextual influences will be taken into account by these contextual units should be given. For example, to evaluate the health impact of traffic-related air pollution, contextual units should be conceived so that they can take the spatiotemporal variations of both air pollution and people's daily movement into account (Hoek et al. 2008).

Further, it is important to recognize that observing significant association between a contextual variable and an outcome variable does not in itself validate the contextual units—as we cannot tell whether our results are produced by erroneous contextual units that have introduced variations in the contextual variable such that association between the contextual variable and the outcome variable is observed, even when there is actually no association between them. It is also important to recognize that, as Spielman and Yoo (2009) have shown using simulation experiments, the model with the best fit is not necessarily the one that uses the true geographic context. This means that appropriateness of contextual units cannot be justified using model fit as a criterion.

In view of these difficulties, it would be particularly helpful (when resources allow) to perform sensitivity analysis to assess how different delineations of contextual units might affect contextual variables and study results (Shi 2010). There is some evidence to date that both are sensitive to the choice of contextual units (e.g., Kwan et al. 2009; Troped et al. 2010; Zenk et al. 2011). Kwan et al (2011), for instance, observed significant difference between the composite deprivation index (as a contextual variable) derived from circular buffers around participants' home addresses and those derived from half-mile road network buffers around participants' GPS tracks. The deprivation index derived using the minimum convex polygon is significantly different from those derived using participants' home census tracts. With respect to research findings, Oliver, Schuurman, and Hall (2007) found that the use of different kinds of buffers around participants' homes (based on centroids of their home postal codes) as contextual units has a considerable influence on the results: Land-use characteristics tend to show greater associations with walking using line-based road network buffers than circular buffers; circular and polygon buffers tend to underestimate the effects of land-use characteristics on walking because they might include large areas that are irrelevant to walking (e.g., industrial land) or inaccessible. These studies indicate that both contextual variables and study results are sensitive to the choice of contextual units. It is thus important to undertake sensitivity analysis to determine their stability and the extent to which they will be affected.

In recent years, geographers and health researchers have explored various methods to address the UGCoP. A promising direction is the use of individual activity space to approximate the true geographic context. An activity space is the area containing all locations that an individual visits as a result of his or her daily activities and travel (Golledge and Stimson 1997). Because humans tend to exhibit a high degree of habitual behavior on most days and circulate on a relatively small island in space–time, their actual or potential activity spaces could provide better proxies to true geographic contexts than conventional administrative areas (cf. Kwan 1999, 2000; González, Hidalgo, and Barabási 2008). Another advantage of this approach is that exposure to contextual influences is evaluated based on personalized contextual units that allow exposure level to vary even for individuals within the same neighborhood or household (Kwan 2009). It also helps transcend the traditional division of health-determining factors into either neighborhood or individual characteristics. Because personalized contextual units are constructed based on people's daily activities and travel as well as their interactions with various places, values of the contextual variables based on these units reflects both individual and place characteristics at the same time.

Some studies construct individual activity spaces using GIS and activity survey data (e.g., Arcury et al. 2005; Sherman et al. 2005; Kwan et al. 2009; Vallée et al. 2010). In these studies the standard deviational ellipse, the kernel density surface, the road network buffer, and the minimum convex polygon are common methods for deriving activity spaces. Although activity surveys provide useful data for delineating people's activity spaces, information about the location and timing of these activities is often very limited. To overcome this limitation, researchers have begun to explore the use of GPS or other location-aware devices in collecting detailed space–time data of people's activities and routes (e.g., Wiehe et al. 2008; Maddison et al. 2010; Troped et al. 2010; Zenk et al. 2011). For example, Kwan et al. (2011) used an integrated GPS–activity diary approach to examine the effect of exposure to protobacco advertisement and socioeconomic deprivation on the use of smokeless tobacco in the Appalachian region of Ohio. The study found that those who traveled in areas with lower socioeconomic status are more likely to use smokeless tobacco heavily.

Because GPS data can record where and how much time people spend as they undertake their daily activities with very high spatial and temporal resolutions, these data allow us to assess people's environmental exposures much more accurately.[2] Detailed GPS data also allow us to perform time-geographic 3D visualizations of people's space–time paths, which will be particularly valuable for studying the health risk of individuals without a stable home or those who live in multiple places such as the homeless (Hägerstrand 1970; Kwan 2004; Kwan and Ding 2008; Lee and Kwan 2011). Using a person's GPS tracks collected over many days (e.g., a week), we can estimate the probability distribution of his or her activities and routes over space and time and more accurately approximate the true geographic context. This will help us move beyond deterministic approaches to the delineation of contextual units and facilitate the development of new stochastic approaches for doing so. Further, using GPS data also helps overcome the conventional dichotomy between daytime and nighttime populations because these data capture people's continuous space–time trajectories, and analysis does not require dividing a day into two distinct segments of time.

Using GPS data to delineate activity spaces and approximate true geographic contexts represents a significant step forward in addressing the UGCoP. Some health behaviors, however, are affected by *situational contingencies* (e.g., interactions with others in real time) that cannot be captured by GPS data alone. For instance, a particular situation might promote or inhibit an adolescent's substance use depending on who is with him or her (e.g., friends or parents) and what they are doing together even for the same context (e.g., movie theater). To take into account the full spectrum of contextual influences on certain health behaviors or outcomes, it might be important to also consider the characteristics of relevant real-time contexts and social interactions. An emerging and promising approach along this line is to integrate GPS methodologies with *ecological momentary assessment* (EMA) and *social network analysis* (SNA). EMA has been used in a wide range of health studies to collect data on people's real-time situations (Shiffman 2009). It involves using wireless devices (e.g., mobile phones) to prompt and collect information from participants about their moods, perceptions, behaviors, and features of the environment as they occur in real time. On the other hand, data about people's social networks—such as attributes of their peers and friends and attributes and structure of the relationships among them—will help shed light on how a person's interactions with others in particular spaces and times could affect their health behaviors (Mennis and Mason 2011). An integrated GPS–EMA–SNA approach seems particularly promising for understanding the transmission of infectious diseases—as information about who has contact with whom and what they are doing together at what times will help shed light on the sociogeographic processes involved.

Another important component in attempts to address the UGCoP involves measuring the spatiotemporal variation of contextual or environmental influences (e.g., airborne pollutants), identifying when individuals are affected by them, and assessing the cumulative exposure of each individual with respect to his or her movement in space–time. This is a highly challenging research area because, for instance, the spatiotemporal dynamics of contextual influences and detailed space–time trajectories of individuals need to be integrated into a suitable analytical framework to accurately assess people's exposures. Some recent studies indicate how this could be accomplished (e.g., Setton et al. 2010). For instance, Gulliver and Briggs (2005) collected twenty-four-hour activity diary data from participants and developed a space–time exposure modeling method to evaluate their journey-time exposure to traffic-related pollution. The method integrated four submodels in a GIS: a traffic model, an air pollution dispersion model, a background pollution model, and a time-activity-based exposure model. Research in this area is sorely needed to fully address the UGCoP.

Promising as these developments might sound, there are still many challenges and limitations. First, collecting GPS, EMA or social network data is costly and time consuming. These methods are thus not suitable for obtaining data for large populations in a short period of time. Second, GPS have their own limitations (e.g., cannot collect reliable indoor data) and are sometimes error prone. Third, collecting space–time data with GPS greatly increases the volume of data, and methods for analyzing these data in health research are still limited to date. This could increase our analytical burden and undermine our ability to identify the true geographic context or to accurately assess people's exposure to contextual or environmental influences (Kwan 2004).

To conclude, the UGCoP is a problem as fundamental as the MAUP, but it is a different kind of problem that calls for new research on its confounding effects and mitigation. Recent studies have shown that both contextual variables and research findings are sensitive to different delineations of contextual units, and model fit by itself is not a reliable criterion for deciding the most appropriate contextual units. It is time to go beyond the static concepts and methods of conventional notions of geographic context and exposure measures. People move around to undertake their daily activities and come under the influence of many different neighborhood contexts besides their residential neighborhood. Their movement, their routes, the places they visit, and the time they spend there are no less important than their residential neighborhood in determining their exposure to contextual influences. Studies on the effects of contextual influences on health thus need to consider where and how much time people spend while engaged in their daily activities in relation to the spatiotemporal patterns of relevant contextual influences. This dynamic conceptualization of geographic context is very much in line with the "new mobilities paradigm" that emerged in social science in the last decade or so (Sheller and Urry 2006). The mobilities turn asserts the ontological significance of people's movement and expands our attention to what people experience while traveling. For health research it helps turn our focus from location to movement, from place to mobility, and from space to space–time. In the final analysis, humans are active agents who construct their own geographic contexts and tie together different spatial scales through their daily activities, movements and social interactions. The interconnections among individuals and places are vastly complex and vibrantly dynamic, and they should be conceptualized and examined as such.

Acknowledgments

I thank Antoinette WinklerPrins for handling the blind review process and making editorial decisions for this article (including the abstract). Her helpful suggestions and the thoughtful comments of three anonymous reviewers have helped improve the article considerably. This article was written while I was supported by the following grants: NSF BCS-0729466, NCI R21 CA129907, and NIDA R01DA025415-01.

Notes

1. Two important qualifications of the article's focus on health studies are in order. First, arguments in this article mainly apply to health studies that are based on ecological designs because other research designs are not primarily concerned with identifying contextual influences from geographic factors and processes (and thus do not need to explicitly delineate contextual areas). Second, discussion in this article is relevant mainly to studies in which area-based contextual variables (e.g., neighborhood poverty) are used to explain or predict individual health behaviors or outcomes. An important goal of many health studies, however, is to identify at-risk populations or areas where the health outcomes are significantly worse than expected. Given their analytical focus on the relationship between area-based contextual variables and area-based outcome variables (e.g., low birthweight rates of census tracts), using conventional administrative units like census tracts in this kind of study is needed and is often the only viable option.
2. A major concern with collecting GPS data in health research is participants' privacy and data confidentiality, because it could be possible to identify a person's identity through reverse geocoding if data are not handled carefully. In countries with strict human subject protection regulations (such as the United States), all persons involved in collecting and analyzing the data are required to go through rigorous human subject protection training and be certified before any involvement in research activities. They are obliged legally and ethically to protect participants' privacy and data confidentiality, and all research procedures (including recruitment, informed consent, data analysis, and dissemination of results) require prior approval from and are closely monitored through continuing review by their institutional or ethical review boards. For instance, in the Appalachian smokeless tobacco usage study, all data were deidentified before being incorporated into the database and no one handling those data or seeing them by chance will be able to identify any participant. Further, no maps or displays of the home or activity sites visited by participants or their daily paths can be printed or disseminated. The GPS data cannot be used in any form other than for deriving activity spaces and related measures or generating aggregate statistical results. In countries without strict regulations and in situations where people provide information without knowing or agreeing to its use for research purposes (e.g., social networking Web sites such as Facebook Places, Twitter, and Foursquare), there might be no informed consent and human subject protection protocol and issues of privacy violation can be a serious concern. It is not clear how the use of location data can be justified with respect to the norms of human subject protection in these contexts.

References

Arcury, T. A., W. M. Gesler, J. S. Preisser, J. E. Sherman, J. Spencer, and J. Perin. 2005. The effects of geography and spatial behavior on health care utilization among the residents of rural region. *Health Services Research* 40 (1): 135–55.

Basta, L. A., T. S. Richmond, and D. J. Wiebe. 2010. Neighborhoods, daily activities, and measuring health risks experienced in urban environments. *Social Science & Medicine* 71 (11): 1943–50.

Berke, E. M., T. D. Koepsell, A. V. Moudon, R. E. Hoskins, and E. B. Larson. 2007. Association of the built environment with physical activity and obesity in older persons. *American Journal of Public Health* 97 (3): 486–92.

Black, L. J., and J. Macinko. 2008. Neighborhoods and obesity. *Nutrition Review* 66 (1): 2–20.

Brownson, R. C., J. J. Chang, A. A. Eyler, B. E. Ainsworth, K. A. Kirtland, B. E. Saelens, and J. F. Sallis. 2004. Measuring the environment for friendliness toward physical activity: A comparison of the reliability of 3 questionnaires. *American Journal of Public Health* 94 (3): 473–83.

Chaix, B. 2009. Geographic life environments and coronary heart disease: A literature review. Theoretical contributions, methodological updates, and a research agenda. *Annual Review of Public Health* 30:81–105.

Cummins, S. 2007. Investigating neighborhood effects on health—Avoiding the "local trap." *International Journal of Epidemiology* 36:355–57.

Diez-Roux, A. V. 1998. Bringing context back into epidemiology: Variables and fallacies in multilevel analysis. *American Journal of Public Health* 88 (2): 216–22.

———. 2001. Investigating neighborhood and area effects on health. *American Journal of Public Health* 91 (11): 1783–89.

Diez-Roux, A. V., and C. Mair. 2010. Neighborhood and health. *Annals of the New York Academy of Sciences* 1186:125–45.

Elgethun, K., R. A. Fenske, M. G. Yost, and G. J. Palcisko. 2003. Time-location analysis for exposure assessment studies of children using a novel global positioning system instrument. *Environmental Health Perspectives* 111 (1): 115–22.

Entwisle, B. 2007. Putting people into place. *Demography* 44 (4): 687–703.

Fotheringham, A. S., and D. W. S. Wong. 1991. The modifiable areal unit problem in multivariate statistical analysis. *Environment and Planning* A 23 (7): 1025–44.

Frank, L. D., T. L. Schmid, J. F. Sallis, J. Chapman, and B. E. Saelens. 2005. Linking objectively measured physical activity with objectively measured urban form—Findings from SMARTRAQ. *American Journal of Preventive Medicine* 28:117–25.

Gatrell, A. C. 2011. *Mobilities and health*. Aldershot, UK: Ashgate.

Golledge, R. G., and R. J. Stimson. 1997. *Spatial behavior: A geographic perspective*. New York: Guilford.

González, M. C., C. A. Hidalgo, and A.-L. Barabási. 2008. Understanding individual human mobility patterns. *Nature* 453:779–82.

Grannis, R. 2009. *From the ground up: Translating geography into community through neighbor networks*. Princeton, NJ: Princeton University Press.

Gulliver, J., and D. J. Briggs. 2005. Time–space modeling of journey-time exposure to traffic-related air pollution using GIS. *Environmental Research* 97:10–25.

Hägerstrand, T. 1970. What about people in regional science? *Papers of Regional Science Association* 24:7–21.

Hanson, S., and P. Hanson. 1981. The travel-activity patterns of urban residents: Dimensions and relationships to sociodemographic characteristics. *Economic Geography* 57:332–47.

Hoek, G., R. Beelen, K. de Hoogh, D. Vienneau, J. Gulliver, P. Fischer, and D. Briggs. 2008. A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmospheric Environment* 42:7561–78.

Inagami, S., D. A. Cohen, and B. K. Finch. 2007. Non-residential neighborhood exposures suppress neighborhood effects on self-rated health. *Social Science & Medicine* 65:1779–91.

Kwan, M.-P. 1999. Gender and individual access to urban opportunities: A study using space–time measures. *The Professional Geographer* 51 (2): 210–27.

———. 2000. Interactive geovisualization of activity-travel patterns using three-dimensional geographical information systems: A methodological exploration with a large data set. *Transportation Research* C 8:185–203.

———. 2004. GIS methods in time-geographic research: Geocomputation and geovisualization of human activity patterns. *Geografiska Annaler B* 86 (4): 267–80.

———. 2009. From place-based to people-based exposure measures. *Social Science & Medicine* 69 (9): 1311–13.

Kwan, M.-P., and G. Ding. 2008. Geo-narrative: Extending geographic information systems for narrative analysis in qualitative and mixed-method research. *The Professional Geographer* 60 (4): 443–65.

Kwan, M.-P., T. Hawthorne, C. Calder, W. Darneider, A. Jackson, and L. J. Krivo. 2009. Activity-space measures for studying spatial crime and social isolation. Paper presented at the annual meeting of the Association of American Geographers, Las Vegas, NV.

Kwan, M.-P., L. L. Kenda, M. E. Wewers, A. K. Ferketich, and E. G. Klein. 2011. Sociogeographic context, protobacco advertising, and smokeless tobacco usage in the Appalachian Region of Ohio (USA). Paper presented at the 2011 International Medical Geography Symposium, Durham, UK.

Kwan, M.-P., and J. Weber. 2008. Scale and accessibility: Implications for the analysis of land use–travel interaction. *Applied Geography* 28:110–23.

Leal, C., and B. Chaix. 2011. The influence of geographic life environments on cardiometabolic risk factors: A systematic review, a methodological assessment and a research agenda. *Obesity Reviews* 12 (3): 217–30.

Lee, J. Y., and M.-P. Kwan. 2011. Visualization of sociospatial isolation based on human activity patterns and social networks in space-time. *Tijdschrift voor Economische en Sociale Geografie* 102 (4): 468–85.

Löytönen, M. 1998. GIS, Time geography and health. In *GIS and health*, ed. A. C. Gatrell and M. Löytönen, 97–110. London and New York: Taylor and Francis.

Macintyre, S., and A. Ellaway. 2003. Neighborhoods and health: An overview. In *Neighborhoods and health*, ed. I. Kawachi and L. F. Berkman, 20–42. Oxford, UK: Oxford University Press.

Macintyre, S., A. Ellaway, and S. Cummins. 2002. Place effects on health: How can we conceptualise, operationalise and measure them? *Social Science & Medicine* 55:125–39.

Maddison, R., Y. Jiang, S. Vander Hoorn, D. Exeter, C. N. Mhurchu, and E. Dorey. 2010. Describing patterns of physical activity in adolescents using global positioning systems and accelerometry. *Pediatric Exercise Science* 22 (3): 392–407.

Matthews, S.A. 2008. The salience of neighborhood: Some lessons from sociology. *American Journal of Preventive Medicine* 34 (3): 257–59.

Mennis, J., and M. J. Mason. 2011. People, places, and adolescent substance use: Integrating activity space and social network data for analyzing health behavior. *Annals of the Association of American Geographers* 102 (2): 279–91.

Monden, C. W. S., F. J. Van Lenthe, and J. P. Mackenbach. 2006. A simultaneous analysis of neighborhood and childhood socio-economic environment with self-assessed health and health-related behaviors. *Health & Place* 12:394–403.

Oliver, L. N., N. Schuurman, and A. W. Hall. 2007. Comparing circular and network buffers to examine the influence of land use on walking for leisure and errands. *International Journal of Health Geographics* 6:41.

Openshaw, S. 1984. *The modifiable areal unit problem*. Norwich, UK: Geo Books.

Sampson, R. J., J. D. Morenoff, and T. Gannon-Rowley. 2002. Assessing "neighborhood effects": Social processes and new directions in research. *The Annual Review of Sociology* 28:443–78.

Setton, E., C. P. Keller, D. Cloutier-Fisher, and P. W. Hystad. 2010. Gender differences in chronic exposure to traffic-related air pollution—A simulation study of working females and males. *The Professional Geographer* 62 (1): 66–83.

Sheller, M., and J. Urry. 2006. The new mobilities paradigm. *Environment and Planning* A 38:207–26.

Sherman, J. E., J. Spencer, J. S. Preisser, W. M. Gesler, and T. A. Arcury. 2005. A suite of methods for representing activity space in a healthcare accessibility study. *International Journal of Health Geographics* 4:24.

Shi, W. 2010. *Principles of modeling uncertainties in spatial data and spatial analyses*. Boca Raton, FL: CRC Press.

Shiffman, S. 2009. Ecological momentary assessment (EMA) in studies of substance use. *Psychological Assessment* 21 (4): 486–97.

Spielman, S. E., and E.-H. Yoo. 2009. The spatial dimensions of neighborhood effects. *Social Science & Medicine* 68 (6): 1098–1105.

Subramanian, S. V., K. Jones, and C. Duncan. 2003. Multilevel methods for public health research. In *Neighborhoods and health*, ed. I. Kawchi and L. F. Berkman, 65–111. Oxford, UK: Oxford University Press.

Troped, P. J., J. S. Wilson, C. E. Matthews, E. K. Cromley, and S. J. Melly. 2010. The built environment and location-based physical activity. *American Journal of Preventive Medicine* 38 (4): 429–38.

Vallée, J., E. Cadot, F. Grillo, I. Parizot, and P. Chauvin. 2010. The combined effects of activity space and neighborhood of residence on participation in preventive health-care activities: The case of cervical screening in the Paris metropolitan area (France). *Health & Place* 16:838–52.

Wheaton, B., and P. Clarke. 2003. Space meets time: Integrating temporal and contextual influences on mental health in early adulthood. *American Sociological Review* 68:680–706.

Wiehe, S. E., S. C. Hoch, G. C. Liu, A. E. Carroll, J. S. Wilson, and J. D. Fortenberry. 2008. Adolescent travel patterns: Pilot data indicating distance from home varies by time of day and day of week. *Journal of Adolescent Health* 42:418–20.

Wilks, D., H. Besson, A. Lindroos, and U. Ekelund. 2010. Objectively measured physical activity and obesity prevention in children, adolescents and adults: A systematic review of prospective studies. *Obesity Reviews* 12 (5): 119–29.

Zenk, S. N., A. J. Schulz, S. A. Matthews, A. Odoms-Young, J. Wilbur, L. Wegrzyn, K. Gibbs, C. Braunschweig, and C. Stokes. 2011. Activity space environment and dietary and physical activity behaviors: A pilot study. *Health & Place* 17 (5): 1150–61.

Correspondence: Department of Geography, University of California, Berkeley, 507 McCone Hall, Berkeley, CA 94720-4740, e-mail: mpk654@gmail.com.