LAND USE PATTERN AND HOSPITALIZATION FOR CHRONIC CONDITIONS: A SPATIAL STUDY ON MODIFIABLE HEALTH RISKS IN SAN FRANCISCO

Background

Healthcare costs for diabetes, hypertension, and congestive heart failure (CHF) have shown an alarming trend in the past few decades in the United States. According to the Centers for Disease Control and Prevention, diabetes prevalence in the last decade was up 40 percent among the end of the baby-boomer cycle (aged 40 to 49), and up 70 percent in those aged 30 to 39. Similar statistics from CDC/NCHS on hospital discharge show that the number of cases for congestive heart failure rose from 377,000 in 1979 to 995,000 in 2001. A recent study by Anderson and Horvath (2004), based on the 1998 Medical Expenditure Panel Survey (MEPS, 1998) shows that 78 percent of the total health care dollars are spent on people with one or more of these chronic conditions.

Active lifestyle plays an important role in the management and control of diabetes, hypertension, and congestive heart failure (Yen and Kaplan, 1999; Ellaway and Macintyre, 1996). Researchers in the medical fields are increasingly of the view that increased obesity and sedentary lifestyles in the population, in general, and urban population, in particular, are creating an explosion in the prevalence of these diseases (Abrahamson and Hu, 2004). However, opportunities to adopt and maintain an active lifestyle depend on the environmental and the infrastructural amenities available in the community where a person lives (Bell *et al*. 2002). Much of the variation in health status across individuals may be attributed to unequal access to those amenities required for a healthy lifestyle (Anderson and Horvath, 2004).

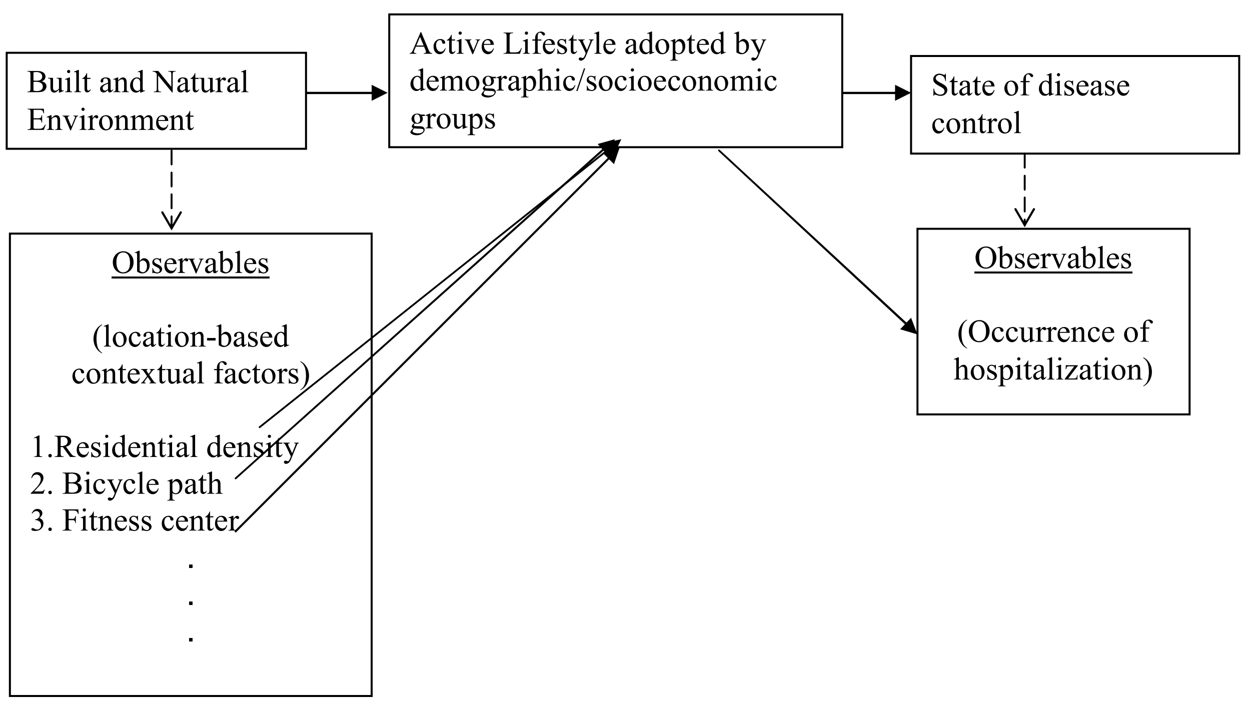
Public policies on land use that encourages active lifestyle can play an important role in this direction (Sallis *et al*., 2002). The Public health literature emphasizes the need to evaluate the role of urban sprawl and residential environment on physical activity (Handy et al., 2002), cardiovascular risk (Roux, 2003) and obesity (Frumkin, 2002). Similarly, a growing body of work in urban planning and transportation has investigated how features of the built environment enrich urban environment (Moudon and Untermann, 1991) and influence travel behaviors of individuals including walking and bi cycling (Bento *et al*., 2004; Saelens *et al*. 2003; Ross, 2000). Unfortunately, public policies of land use to develop and maintain community-based infrastructure and the environment have, so far, been framed in isolation, without recognizing their tremendous potential in impacting population health and health care costs. Yet population-based strategies that shift the community-wide distribution of health risk factors are likely to have much greater public health impact than individual-based approaches (Bell *et al*., 2002; Denton *et. al*., 2002; Schmid et al., 1995; Rose, 1985).

One potentially useful but as yet unexplored measure of morbidity associated with reduced physical activity is hospitalization for chronic conditions such as diabetes, hypertension and CHF. These diseases are amenable to control through active lifestyle and timely and effective treatment in an outpatient setting (Anderson and Horvath, 2004; McGinnis and Foege 1993; Bindman et al. 1995). Hospitalizations for these conditions therefore, reflect a significant decline in disease management and thereby may be used to provide a measure of the health consequence of access barriers to adopting an active lifestyle. Yet to our knowledge, the relationship between opportunities for active lifestyle and the risk of hospitalization for these diseases has not been studied previously.

This research will measure access to recreational facilities, transportation options and community safety and will link these data to hospitalizations for diabetes, hypertension and CHF in San Francisco to explore the relationship between the built environment and the occurrences of hospitalization for the three aforementioned chronic conditions.

Conceptual Basis

The literature on the association between active lifestyle and chronic conditions recognizes the importance of the built environment in two distinct cause-and-effect sequences. The first sequence links the built and natural environment to opportunities for active lifestyle (Sallis *et al.*, 2002). The second sequence follows the first and links active lifestyle to the control of chronic diseases (Anderson and Horvath, 2004; McGinnies and Foege 1993). The following diagram illustrates the two sequences, with the solid arrows showing the direction of cause-and-effect relationships:



In the proposed project, we do not analyze the two cause-and-effect sequences, separately. Instead, we combine the two sequences to evaluate the extent to which differential access to public amenities (location-based contextual factors) that promote active living translates to variation in hospitalization and cost between different racial, ethnic, and demographic groups. There are three advantages in the proposed approach. First, analysis of the two separate sequences requires measurement of physical activity at the individual level. Unfortunately, objective measures (as opposed to self-reported) are extremely rare in an individual-level approach. A location-based approach has the ability to bring in more objectivity in the measurements of factors that determine the occurrence of hospitalization. Secondly, location-based approaches can often eliminate individual-level idiosyncrasies through the process of aggregation. Thirdly, designing and evaluating policy instruments that affect community behavior *through* changes in the contextual factors are comparatively more straightforward than designing and evaluating those that *directly* affect individual behavior. WHY ZIP CODE.

Data and Measurement Issues

The study is based in San Francisco. Our primary unit of analysis is the zip code. San Francisco Bay Area has been widely studied by researchers in urban planning and transportation (e.g. Southworth and Owens, 1993; Dunphy and Fisher, 1994; Cervero and Radisch, 1995). There are two reasons that make San Francisco Bay Area a good candidate for this research. First, health risks associated with chronic diseases, namely diabetes, hypertension, and congestive heart failure and the associated costs of hospitalization are particularly high in the area. Secondly, its land use design varies significantly across locations within the metropolitan center (Southworth and Owens, 1993). The data for the research can be broadly classified in to three distinct categories. They are described below.

(i) Healthcare and Hospitalization data

Information in this category includes patient-level hospital discharge data with zip code as the identifier of the patient residence and are obtained from the Patient Discharge Data files made available to the public by California's Office of the Statewide Health Planning and Development (OSHPD). The data file provides a record for every hospitalization in California for the year 2000. The study population consists of all adult residents of San Francisco aged 40 and above during the year 2000. This dataset also includes patient-level information on number of hospitalizations, demographics (age, sex, race/ethnicity), location of residence (zip code, county), diagnosis, and insurance status. We will use ICD9 codes of the principal diagnosis to identify hospitalizations for diabetes, hypertension, and congestive heart failure of all patients aged 40 and above. In addition, information on number of primary care clinics and number of physicians serving each zip code were obtained through the OSHPD and American Medical Association master file respectively.

(ii) Socioeconomic and demographic data

Demographic factors, such as age, gender, and race/ethnicity are important determinants of hospitalization for the three study conditions. We have obtained census data on demographic compositions for all the zip codes in California from the Department of Finance, Sacramento, CA. These zip code-level exposure data from 2000-census figures along with the incidence data from OSHPD is used to calculate the hospitalization rates for the three disease conditions for each zip code in San Francisco. A list of other location-based socioeconomic factors that may be associated with hospitalization as well as physical activity, such as poverty rate, educational attainment, and crime rate have been obtained from the TIGER files of the published 2000-census records.

(iii) Infrastructural and built environmental data

One major challenge in the study of environmental determinants of physical activity and health pertains to the conceptualization and measurement of relevant environmental features. Currently, few operational measures of environmental quality that can be used in health studies exist. GIS-based approaches show a promising trend in the development of objective measures of resource accessibility. Such measures include housing density, land use mix, street connectivity, aesthetics, recreational facilities, walking and cycling facilities, to name a few. The proposed research creates an inventory of objective measures of the structural features of land use associated with physical activity pattern for the city of San Francisco at the zip code level incorporation into our analysis.

Physical activity includes both recreational activity and utilitarian activity. Recreational activity is primarily for exercise, whereas utilitarian activity is secondary to other purposes such as commute or shopping. Both types account for physical activity and hence promote health. In the proposed project, environmental characteristics that affect both types of physical activity are considered. Furthermore, we focus on the three modes of activity, which are closely associated with the design of the built environment: exercise with designated facilitates (e.g. fitness club, golf courses, etc.), walking, and bicycling.

Environmental factors that influence physical activity can be grouped into three dimensions: land use patterns, design characteristics, and transportation systems (Frank et al., 2003). Specifically, the following factors have been found to be relevant (Frunkmin et al., 2004):

- Overall neighborhood design such as street network pattern, land use mix, etc. Traditional neighborhoods with grid street pattern are found to be associated with more walking than hierarchical, curvilinear street network (Moudon et al., 1997). High mixing and balance between residential and commercial-retail uses are also found to encourage walking (Kockelman, 1997).

- Neighborhood density measured by the density of housing units. An analysis of 1995 National Personal Transportation Survey revealed that the difference between densest areas and sparse areas in terms of walking and bicycling can be up to five folds (Ross and Dunning, 1997).

- Nearby sidewalks, footpaths, and bicycle facilities whose quality and access affect their usage by pedestrians and bikers, both for utilitarian and recreational purposes. Accessibility can be measured by the length of sidewalk and bike path. Quality is affected by many factors, such as cross walks, number of intersections, etc.

- Enjoyable scenery promotes people to get out and be active (Ball, 2001). Architecture, distant view and greenery can make a place attractive and aesthetically appealing. Enjoyable scenery is associated with people's perception and thus is best determined through survey on street level. Because of the scale of the proposed project, this street-level factor is modified into overall attractiveness of the zip code and is measured by tree coverage, presence of water body and open space (including parks and recreation areas).

- Availability of designated exercise facilities, e.g. fitness centers, golf courses, public parks and open space. These facilities and land use can influence recreational activities other than walking and bicycling, and therefore are included in the study.

- Safety and topography. All of the above factors are facilitators or promoters of physical activity. However, barriers and inhibitors also need to be addressed. Safety, which includes both traffic safety and crime safety, is perhaps the single most significant factor that influences people's decision to walk or bicycle. In fact, its influence can exceed that of the facilities and neighborhood design (Cervero and Duncan, 2003). Topography is another inhibitor. If the sidewalk or bike path is too steep, which is often true in the San Francisco, it is unlikely that the pedestrian or bicycling facilities will be widely used.

The above environmental characteristics may coexist in a zip code. For example, zip codes with grid street pattern are often associated with high population density and greater mixing of land use. This covariance may pose some challenges for the analysis and we will take full consideration of it.

TABLE 1: DATA SOURCES FOR ENVIRONMENT FACTORS ASSOCIATED WITH PHYSICAL ACTIVITY PATTERN

Environmental Factor	Measurements	Data Required	Source
Neighborhood design	Block size and density	Street network	TIGER file from Census Bureau
	Vegetation coverage	Distribution of vegetation	SPOT 10-meter satellite imagery
Pedestrian/Bike friendliness	Sidewalk and bike path continuity and length	Sidewalk network and bicycle path network	Metropolitan Transportation Committee (MTC)
	Proportion of three-, four-, and five-way intersections and dead ends	Street network	TIGER file
Topography	slope	Digital Elevation Model (DEM)	USGS (http://www.usgs.gov)
Crime safety	Police-reported crime data on zip code level	Crime incidence	Police department of each county
Recreational facilities	1. Distance to beach	Coastal line, beaches	NOAA (http://www.noaa.gov)
	2. area of open space & parks	Open space and parks distribution	GreenINFO Network
	3. number of fitness centers, golf courses, and water sports	Count of Fitness centers, golf courses, water sport facilities, etc.	Zipcode business pattern from Census Bureau.

Analysis
Our preliminary analysis reveals significant variation in health care costs among zip codes in California. For example, the average cost of hospitalization in a zip code for the three study conditions in California, derived from the statewide hospital discharge data for the year 2000 is $18,989, with a significantly high standard deviation of $28,205. The proposed project aims at modeling this high variation systematically in a multivariate framework, for the city of San Francisco.

We propose to adopt the poisson count model that has been recently used by authors in the health economics literature (e. g. Bindman et al., 2004; Deb and Trivadi, 2002). The poisson count model involves maximum likelihood estimation of the model characterized by an incidence rate parameter (hospitalization rate in the current context), commonly called λ (lambda). In particular, the number of hospitalization in an area will be estimated using the following specification:

$$P(Y = y) = \frac{e^{-\lambda}\lambda^{y}}{y!} \tag{1}$$

Here the parameter λ is influenced by a set of explanatory variables X_is, as follows:

$$\lambda = \exp(b_0 + \sum b_i X_i) \tag{2}$$

where b_0 is the constant term, and b_i's are the effect coefficients. Our approach will be unique in that we will use an exhaustive set of location-based demographic, environmental, and infrastructural factors as explanatory variables that are believed to affect the incidence rate in (2) and consequently, counts or events of hospitalization for the three chosen conditions. In the demand-side analysis, each discharge will be regarded as a single count for the zip code location where the patient resides. Total counts will be created for multiple levels of categorizations based on multiple contextual factors for each zip code residential location in San Francisco. An appropriate scale factor will be introduced to account for any over-dispersion.

In order to estimate the poisson rate, observations on the number of hospitalizations for the three study conditions will be ascertained from the hospital discharge files and will be aggregated into analytic cells defined by different combinations of values for the explanatory variables that affect the rate. For example, based on the data that we have already integrated for the San Francisco County for the year 2000, one cell contains 13 cases of hospitalizations, among African-American female, aged 40-55, residing in zip-code 94102, which has 68,633 sq meters of open space, is 2,090 meters from the beach, has 3 fitness centers, 25.41 percent population below high school degree, and a median income of $22,351. Such an approach can accommodate changes in environmental characteristics across zip codes, such as residential density, street connectivity, accessibility of parks, etc. to reflect a change in the rate and, consequently, the counts of hospitalization. For econometric estimation, the corresponding denominator population for calculating the hospital admission rate for each cell has been obtained from census tabulations that provide age, sex, and racial breakdown of the number of residents of each zip-code. The population “at risk” for a hospitalization will vary across cells and will therefore be included as an offset variable in the model. The coefficient estimates from the poisson regression model will be used to obtain standardized predicted rates adjusted for differences in demographic composition. In this way we will eliminate the role of demographic compositions in generating area wise variation in hospitalization rates. These rates will then be used to derive the predicted counts of hospitalization for residents exposed to different levels of environmental resources and infrastructure for physical activity. The model will also account for other contextual factors such as physician availability, hospital bed availability neighborhood socioeconomic environment (poverty rate, and educational attainment), and neighborhood safety (crime rate) which impact either or both hospitalization rate and physical activity rate in a neighborhood. The effect of environmental factors of the current location on active living and, consequently, hospitalization will depend to a large extent on the patient’s length of stay in the current zip code location. Although we cannot obtain data on the exact length of stay in the current zip code for each individual, we expect that the census data on percentage living in the same zip code five years ago will control for the residential mobility of the population in general. Moreover, despite inclusion of a large number of control variables, the possibility of omitting some confounding variables remains. We will test for such omitted variable bias and correct it using appropriate econometric techniques.

Bibliography

Abrahamson M. J. and F. Hu. 2004. "Diabetes and Boomers: An Epidemic Waiting to Happen." Times, March 29.

Anderson G. and J. Horvath. 2004. "The Growing Burden of Chronic Disease in America." *Public Health Reports*, 119 (3): 263-700.

Ball, K., Bauman,A., Leslie E., Owen, N., 2001. "Perceived Environmental and Social Influences On Walking for Exercixse in Austrialian Adults." *Preventive Medicine*, 33, 434-440.

Bell, J. D., J. Bell, R. Colmenar, R. Flournoy, M. McGehee, V. Rubin, M. Thompson, J. Thompson, and V. B. Vasquez. 2002. "Reducing Health Disparities Through a Focus on Communities." "Policy Link" Report.

Bento, A. M., M. L. Cropper, A. M. Mobarak, and K. Vinha. 2004. "The Impact of Urban Spatial Structure on Travel Demand in the United States." Forthcoming, *Quarterly Journal of Economics*.

Bindman, A., A. Chattopadhyay, D. Osmond, W. Huen, and P. Baccheti. 2004. "The Impact of Medicaid managed Care on Preventable Hospitalizations for Ambulatory Care Sensitive Conditions." Forthcoming, *Health Services Research*.

Cervero R. and C. Radisch. 1995. "Travel Choice in Pedestrian Versus Automobile Oriented Neighborhoods." Working Paper 644. University of California at Berkeley, Institute of Urban and Regional Development.

Cervero, R., and Duncan, M., 2003. "Walking, Bicycling, and Urban Landscapes: Evidence From the San Francisco Bay Area." *American Journal of Public Health*, 93, 1478-1483.

Denton, F. T., A. Gafni, and B. G. Spencer. 2002. "Exploring the Effects of Population Change on the Cost of Physician Services." *Journal of Health Economics*, 21 (3): 781-803.

Deb P. and P. K. Trivedi. 2002. "The Structure of Demand for Healthcare: Latent Class Versus Two-part Models." *Journal of Health Economics*, 21 (3): 601-25.

Dunphy R. and K. Fisher. 1994. "Transportation, Congestion, and Density: New Insights." *Transportation Research Record*, 1552: 89-92.

Ellaway A. and S. Macintyre. 1996. "Does Where You Live Predict Health Related Behavior? A Case Study in Glasgow." *Health Bulletin (Edinburgh)*. 54 (2): 443-46.

Frank, L.D., Engelke, P.E., Schmid, T.L., 2003. "Health and Community Design: The Impacts of the Built Environment On Physical Activity. Washington, DC: Island Press.

Frumkin H. 2002. "Urban Sprawl and Public Health." *Public Health Report*, 117 (3): 201-17.

Handy, S. L., M. G. Boarnet, R. Ewing, R. E. Killingsworth. 2002. "How the Built Environment Affects Physical Activity: View From Urban Planning." *American Journal of Preventive Medicine*, 23 (1): 64-73.

Kockelman, K.M., 1997. "Travel Behavior As Function of Accessibility, Land Use Mixing, and Land Use Balance: Evidence From San Francisco Bay Area." *Transportation Research Record*, 1607, 116-125.

McGinnis J. M. and W. H. Foege. 1993. "Actual Causes of Death in the United States." *Journal of American Medical Association*. 270: 2207-12.

Moudon, A.V., Hess, P., Snyder, M.C., Stanilov, K., 1997. Effects of Site Design on Pedestrian Travel in Mixed-use, Medium-density Environments. *Transportation Research Record*, 1578, TRB, National Research Council, Washington D.C., pp. 48-55.

Moudon A. V. and R. Untermann. 1991. “Grids Revisited.” In A. V. Moudon (Ed.), Public Streets for Public Use. New York: Van Nostrand Reinhold Co., Inc.

Rose G. 1985. “Sick Individuals and Sick Populations.” *International Journal of Epidemiology*, 14 (1): 32-38.

Ross C. E. 2000. “Walking, Exercising, and Smoking: Does Neighborhood Matter?” *Social Science and Medicine*, 51 (1): 618-20.

Ross, C.L., and Dunning, A.E., 1997. “Land Use Transportation Interaction: An Examination of the 1995 NPTS Data.” Atlanta, USDOT, Federal Highway Administration, 1997. Available http://npts.ornl.gov/npts/1995/Doc/landuse3.pdf.

Roux A. V. D. 2003. “Residential Environments and Cardiovascular Risk.” *Journal of Urban Health*, 80 (4): 569-89.

Saelens, B. E., J. F. Sallis, and L. D. Frank. 2003. “Environmental Correlates of Walking and Cycling: Findings from the Transportation, Urban Design, and Planning Literatures. *Annals of Behavioral Medicine*, 25 (1): 80-91.

Sallis, J. F., K. Kraft, and L. S. Linton. 2002. “How the Environment Shapes Physical Activity: A Transdisciplinary Research Agenda.” *American Journal of Preventive Medicine*, 22 (1) 202 (commentary).

Schmid, T., M. Pratt, and E. Howze. 1995. “Policy as Intervention: Environmental and Policy Approaches to the Prevention of Cardiovascular Disease.” *American Journal of Public Health*, 85 (9): 1207-11.

Southworth M. and P. Owens. 1993. “The Evolving Metropolis: Studies of Community, Neighborhood, and Street Form at the Urban Edge.” *Journal of American Planning Association*, 59 (3): 271-87.

Yen I. H. and G. A. Kaplan, 1999. “Neighborhood Social Environment and risk of death: Multilevel Evidence Form the Alameda County Study.” *American Journal of Epidemiology*, 149 (4): 898-907.