Journal of Urban Health: Bulletin of the New York Academy of Medicine, Vol. 86, No. 1
doi:10.1007/s11524-008-9306-7

Neighborhood Effects on an Individual's Health Using Neighborhood Measurements Developed by Factor Analysis and Cluster Analysis

Yu-Sheng Li and Ying-Chih Chuang

ABSTRACT *This study suggests a multivariate-structural approach combining factor analysis and cluster analysis that could be used to examine neighborhood effects on an individual's health. Data were from the Taiwan Social Change Survey conducted in 1990, 1995, and 2000. In total, 5,784 women and men aged over 20 years living in 428 neighborhoods were interviewed. Participants' addresses were geocoded with census data for measuring neighborhood-level characteristics. The factor analysis was applied to identify neighborhood dimensions, which were used as entities in the cluster analysis to generate a neighborhood typology. The factor analysis generated three neighborhood dimensions: neighborhood education, age structure, and neighborhood family structure and employment. The cluster analysis generated six types of neighborhoods with combinations of the three neighborhood dimensions. Multilevel binomial regression models were used to assess the effects of neighborhoods on an individual's health. The results showed that the biggest health differences were between two neighborhood types: (1) the highest concentration of inhabitants younger than 15 years, a moderate education level, and a moderate level of single-parent families and (2) the highest educational level, a median level of single-parent families, and a median level of elderly concentrations. Individuals living in the first type had significantly higher chances of having functional limitations and poor self-rated health than the individuals in the second neighborhood type. Our study suggests that the multivariate-structural approach improves neighborhood measurements by addressing neighborhood diversity and examining how an individual's health varies in different neighborhood contexts.*

KEYWORDS *Cluster analysis, Factor analysis, Neighborhood, Multilevel analysis, Taiwan*

INTRODUCTION

The influences of neighborhoods on an individual's health have been a popular area of research in the past decade. Neighborhood influences on various health outcomes, such as self-rated health, mental health, chronic disease, and various health behaviors, have been examined. Reviews of neighborhood research on an individual's health have pointed out that previous studies suffered many methodological limitations with one of the major limitations being the underdevelopment of neighborhood measurements.[1–4]

Li is with the Center for Health Policy Research and Development, National Health Research Institutes, Miaoli County, Taiwan; Chuang is with the Graduate Institute of Public Health, Taipei Medical University, Taipei, Taiwan.

Correspondence: Ying-Chih Chuang, Ph.D., Graduate Institute of Public Health, Taipei Medical University, 250 Wu-Hsing St., Taipei 110, Taiwan. (E-mail: yingchih@tmu.edu.tw)

The most typical approach for measuring the neighborhood context has been to use some key variables (i.e., unemployment rates) or established indices (i.e., the Townsend score) derived from census data.[5,6] Some researchers questioned whether census data can adequately measure neighborhoods because census data only comprise neighborhood population characteristics and do not necessarily reflect social process variables (i.e., neighborhood norms or informal social control).[7–10] In addition, neighborhoods defined by census tracts or block groups might not match the perception of its residents and, therefore, might not be closely related to residents' physical and mental health.[11–13] Nevertheless, using census data in neighborhood research has its own advantages because of full coverage of every neighborhood, low cost, and convenience, and in some occasions, it is the only data available to characterize neighborhood influences. It would be helpful to develop a systematic method for researchers to use census data to measure neighborhood contexts.

According to the disciplines of urban geography and sociology, two methodological approaches are often used to analyze census data to understand neighborhood effects.[14] One of the methods is factor analysis, which seeks to discover if the observed variables can be explained largely in terms of a much smaller number of variables and can further uncover the underlying theoretical dimensions of the interrelationships among the observed variables.[15,16] For example, prior studies have applied factor analysis to US census data and found a low socioeconomic status, low residential mobility, and a high concentration of minorities to underlie neighborhood dimensions in the city of Chicago.[9]

Although factor analysis can help in developing neighborhood dimensions, it does not help in understanding the spatial patterning of the dimensions. It assumes that all neighborhoods are homogeneous in possessing these neighborhood dimensions and ignores the possibility that a diversity of neighborhoods with different combinations of these dimensions might exist.[17] For example, disadvantaged neighborhoods can be characterized by high poverty and high residential mobility but can also be characterized by high poverty and low residential mobility.[18]

Cluster analysis, on the other hand, was primarily developed for classification purposes, which helps to understand the nature of neighborhood diversity. Cluster analysis attempts to classify neighborhoods into relatively homogeneous groups and creates a typology based on selected variables.[19] Using this kind of typological approach can help discern underlying forces and processes in each type of neighborhood. For example, Dupere and Perkins used cluster analysis to identify six types of neighborhoods based on residents' perception of physical disorder, fear of crime, informal ties with neighbors, and formal participation. They found that communities with relatively few stressors and high levels of formal participation were associated with better mental health; however, the protective effects of formal participation were not found in communities with higher levels of stressors, suggesting that interpersonal ties are not always helpful in different social contexts.[20]

One major shortcoming associated with cluster analysis is that, although it is a useful tool for classification purposes, it provides no information about the sources of variation among the observed variables.[14,21] In other words, cluster analysis does not address how variables ought to be weighted and combined and does not directly contribute to theoretical construct development. In addition, if a large number of variables are used in a cluster analysis, interpretation of profiles is far more difficult.[17]

In order to address the limitations of factor analysis and cluster analysis, some researchers suggested integrating the two methods in neighborhood research. This combined approach first examines neighborhood dimensions by looking at interrelationships among variables through factor analysis procedures. The second step is to apply a cluster analysis using neighborhood dimensions identified by the factor analysis to generate a neighborhood typology. Finally, individual health differences are compared across different neighborhood types. Urban geographers have called this method the multivariate-structural approach method.[21] Because this approach is based on the real distribution of neighborhoods in a given region or country, it is especially appropriate for contexts that were barely examined in previous studies. In this study, we used a nationwide sample in Taiwan as an example to illustrate this methodological approach. We also compared this approach with the factor analysis approach, which was regarded as the more conventional approach adapted by prior studies.

METHODS

Data

The individual-level data were from the 1990, 1995, and 2000 Taiwan Social Change Surveys.[22–24] A multistage cluster sampling method was used to select adults 20 years and older for the survey. Taiwan administrative structure has the following geographical hierarchy: counties/cities (200,000–3,500,000 people), townships/districts (60,000 people), lis/villages (2,000 people), and lins (500 people). This study first stratified all 359 townships and districts into ten strata according to geographic location and degree of urbanization. Townships or districts in each stratum were then selected by probability proportional to their size (PPS). In each selected township/district, lis and villages were selected by PPS, and individuals were randomly selected in lis and villages using a systematic sampling method, which selects every *K*th cases(sampling interval) from a list of household registration, starting with a randomly chosen case from the first *K* cases on the list. The sampling interval is the ratio of the number of cases in the population to the desired sample size.

Data were collected by interpersonal interviews using a structured questionnaire. Interviewers were required to attend a standardized 2-day training workshop before conducting interviews. The overall response rate was 67% after excluding ineligible cases. The major reasons for not completing the interview included an inability to find the person (18.3%) and a refusal to participate (11.2%). Ten percent of the cases were rechecked for quality control. Participants' residential addresses were geocoded with the 1990 and 2000 Taiwan census data; linear interpolation was used for the 1995 data. Six percent of the respondents were not accurately geocoded to their neighborhoods based on their home addresses, resulting in a final sample size of 428 neighborhoods and 5,784 people. This study defined neighborhoods by the geographical level of lis and villages. They were created by visible boundaries such as streets and rivers and to be as homogeneous as possible with population characteristics. The size of a li/village is smaller than a US census tract (4,000 people) but larger than a census block group (1,000 people). Of the residents in the neighborhoods in this study, 37.6% had less than a middle school education (SD=14.1), 7.5% had a college degree (SD=6.8), 6.3% were single-parent families (SD=2.4), 7.7% were older than 65 years (SD=3.3), and 28.8% were younger than 18 years (SD=5.2). About 57%, 29%, and 14% of neighborhoods were located in urban, suburban, and rural areas, respectively.

Informed consent was obtained from each participant. The ethics committee of the Taiwan National Science Council approved this study.

Measurements

Neighborhood-Level Characteristics Neighborhood-level characteristics were derived from the 1990 and 2000 Taiwan census data; linear interpolation was used for the 1995 data. Neighborhood-level characteristics included: the percentages of residents with less than a middle school education and those with a college degree, the percentage of employed, the percentage of divorced/separated, the percentage of single-parent families, the percentage of residents younger than 15, the percentage of residents older than 65, the percentage of residents who lived in the same house 5 years ago, and the percentage of households that were occupied by the owners. The last two items address the concept of residential mobility. Since the two items neither form a single factor nor load well on other factors, we disregarded them in further analyses. We selected these variables according to the traditions of urban sociology in which a neighborhood's socioeconomic status, family structure, and age distribution are regarded as the fundamental neighborhood characteristics describing neighborhood contexts.[25] We also selected variables based on neighborhood theories, which suggest that high rates of single-parent families and high rates of unemployment are responsible for the conditions of disadvantaged neighborhoods where local basic organizations collapse and social problems are prevalent.[26] We did not include average family income and the percentages of minorities because family income data were not collected through the Taiwan census and the percentages of minorities, such as aborigines, were very small in a neighborhood-level study in Taiwan.

Individual-Level Characteristics Three outcome variables, chronic diseases, functional limitations, and self-rated health, were derived from the Taiwan Social Change Survey, which represent different perspectives of individual health conditions. Chronic disease was measured by the question: "Do you have any chronic disease?" The responses were recorded as yes=1 and no=0. Functional limitations were measured by the question: "In the past 2 weeks, were you limited in any way in your ability to work at a job, do housework, or go to school because of impairment or a physical health problem?" The responses were rated on a four-point scale of "not at all serious," "not very serious," "fairly serious," and "very serious." We recorded "fairly serious" and "very serious" as 1 to represent functional limitations and "not at all serious" and "not very serious" as 0. Self-rated health was measured by the question: "How would you rate your general state of health?" The responses were rated on a four-point scale: "very poor," "poor," "good," and "very good." We recoded "very poor" and "poor" as 1 to represent poor health and "very good" and "good" as 0 to represent good health.

Control variables included age (continuous), gender, race/ethnicity (Taiwanese, Hakka, mainlanders, indigenous populations, and others), marital status (single, married, divorced or separated, and others), monthly household income (seven-point scale), and educational attainment (seven-point scale). Because more than 70% of the participants' ethnicity was Taiwanese, we created a dummy-coded variable and used non-Taiwanese as the reference group. Marital status was recoded as 1 for married and 0 for the others. Income and education were stratified into tertiles (low, middle, high) based on the distribution.

Analysis

We first conducted a factor analysis to identify dimensions of neighborhood indicators. We adopted the principal component method because it is the most common factor extraction approach in exploratory studies.[16] We then used an oblique rotation method, Promax, to rotate the factors because previous studies showed that neighborhood dimensions are correlated.[27,28] The number of factors was decided by a significant jump in the slope of a scree plot and an eigenvalue of >1 and the factors' theoretical coherence. Then, we used a cluster analysis to identify a typology of neighborhoods using dimensions generated by the factor analysis. We selected *K*-means as the clustering algorithm with Euclidean distance as the distance measure to generate clusters. The process of *K*-means starts with specifying the expected number of clusters and the expected centroid (mean) of each cluster. Previous empirical studies indicated that the *K*-means has the best ability to recover the true groupings of data if a nonrandomized starting point is assigned.[29] Because there was insufficient information about the expected number of neighborhood types, we first used Ward's algorithm that requires no initial assigned point to identify the number of groupings. *K*-means was then conducted with the starting point obtained from the results of Ward's method.[14,30]

To determine the number of clusters, we used an inverse scree plot in conjunction with a cross-validation method. An inverse scree plot graphs the number of clusters against the fusion coefficient, which is the numerical value at which various cases merge to form a cluster.[19] A significant jump in the fusion coefficient was used to inform the number of clusters extracted from the data. We used cross-validation to assess the stability of clusters.[14,30] We randomly divided our sample into two samples: a test sample and a validation sample. We first conducted a cluster analysis on the test sample and obtained the centroids of the clusters from the results. Then, we conducted the cluster analysis in the validation samples both with and without specifying the centroids obtained from the test sample. The two results (validation samples with and without specifying centroids) were compared to determine which solution (number of cluster) had higher stability.

After the neighborhood typology was established, multivariate regression models were applied to assess the influences of neighborhood factors and neighborhood types on an individual's health. We used random-intercept models in which the mean of the outcome varied by neighborhood. Multilevel analyses were performed using the SAS GLIMMIX procedure.

RESULTS

Sample Characteristics

Table 1 shows the descriptive statistics for individual- and neighborhood-level characteristics. For chronic diseases, participants who were aged over 60 years, were indigenous, were widowed/divorced, who had not completed elementary school, and who had a household income of ≤NT$29,999 (US$1≈NT$30) reported a higher probability of having chronic diseases. Furthermore, individuals living in neighborhoods with a median rate of elderly concentration were more likely to have chronic diseases. Similar patterns were found for functional limitations and self-rated health except that women were more likely to report functional limitations and rated themselves as having worse health than men. In addition, those who lived in neighborhoods with a high rate of inhabitants who had not finished middle school

TABLE 1 Percentages of individual and neighborhood characteristics by chronic diseases, functional limitations, and self-rated poor health, Taiwan Social Change Survey, 1990, 1995, and 2000, *N*=5,784

	Total	Chronic diseases	Functional limitations	Self-rated poor health
Individual characteristics				
Age				
20–29	18.06	10.01*	20.02*	11.66*
30–39	31.16	13.15	17.66	11.55
40–49	23.72	20.70	17.68	12.35
50–59	12.48	33.15	24.47	18.85
≥60	14.58	51.09	31.40	21.66
Gender				
Male	49.92	22.96	18.29*	10.69*
Female	50.08	22.13	23.94	17.93
Race/ethnicity				
Taiwanese	70.99	22.08*	20.51*	14.58
Hakka	13.34	19.23	21.13	13.79
Mainlander	13.20	26.17	19.95	12.76
Indigenous and others	2.47	34.27	44.44	18.75
Marital status				
Single	18.54	13.51*	20.93*	13.30*
Married	71.93	22.80	19.62	13.34
Widowed/divorced and others	9.53	37.57	32.55	23.01
Education				
<Elementary	8.50	40.73*	36.36*	28.48*
Elementary	26.00	30.13	25.03	16.99
Middle school	15.56	19.01	16.13	12.36
High school	26.98	15.66	19.10	11.64
≥College	22.96	17.69	16.84	10.48
Household income				
≤NT29,999	29.14	26.70*	26.24*	19.81*
NT30,000–NT49,999	24.99	22.03	20.38	13.49
NT50,000–NT69,999	16.69	20.44	18.11	11.20
NT70,000–NT99,999	13.92	19.89	19.50	10.92
≥NT100,000	15.26	23.35	18.17	11.66
Neighborhood characteristics				
Percentage of less than middle school				
Low	33.52	20.49	18.86*	13.63*
Median	33.19	21.90	20.31	13.39
High	33.29	25.91	24.15	16.23
Percentage of greater than age 65				
Low	33.37	22.55*	18.81*	13.35*
Median	33.25	23.66	19.41	13.47
High	33.38	21.52	25.14	16.54
Percentage of single-parent families				
Low	33.34	22.61	20.76	14.60
Median	33.22	24.04	21.51	14.69
High	33.44	21.16	21.08	13.71

*$P<0.05$

and a high rate of elderly concentration were more likely to report functional limitations and poor health than their counterparts.

Factor Analysis and Cluster Analysis

The results of the factor analysis are shown in Table 2. Neighborhood educational level is composed of the percentage of inhabitants with less than a middle school education and the percentage with a college degree (mean=–0.005; range=–3.415 to 1.981). This factor captures the levels of local education and indirectly represents the social class of a neighborhood. The neighborhood age structure is composed of the percentages of inhabitants younger than 15 years of age and those older than 65 years of age (mean=0.016; range=–2.789 to 2.137). This factor represents proportions of dependency attributable to youths and elderly and implies the strain placed upon the productive population. Neighborhood family structure and employment is composed of the percentage of employed, the percentage of divorced/separated, and the percentage of single-parent families (mean=0.003; range=–3.090 to 4.024). Hence, the predominant interpretation of this factor suggests how the concentration of single-parent households and divorced persons affect job searching networks and job opportunities. These three factors explained 76% of the total variance. The correlations among the factors ranged from 0.11 to 0.22, showing small to moderate correlations. Higher scores represent a lower neighborhood educational level, more inhabitants younger than 15 years of age, and more single-parent families and unemployed population.

We used the neighborhood dimensions generated from the factor analysis to conduct the cluster analysis. The cross-validation method showed four, six, and ten potential cluster solutions, which had the best stability. The ten-cluster solution was represented by only a few neighborhoods in some neighborhood types, which thus made interpretation of profiles difficult. The six-cluster solution appeared to be a subtype of the four-cluster solution. For example, the four-cluster solution identified the following types: (1) median education, median youth concentration, and high percentages of single-parent families; (2) high education, median youth concentration, and median percentages of single-parent families; (3) median education, median youth concentration, and median percentages of single-parent families; (4) low education, high elderly concentration, and low percentages of single-parent families. The six-cluster solution further divided type 1 of the four-cluster solution into two types and type 3 of the four-cluster solution into two

TABLE 2 Factor analysis of neighborhood characteristics, Taiwan census data, 1990, 1995, and 2000, *N*=428

	Neighborhood education	Neighborhood age structure	Neighborhood family structure and employment
Percent less than middle school	0.96	−0.08	−0.005
Percent greater than college	−0.94	−0.03	−0.05
Percent less than age 15	0.19	0.87	0.006
Percent greater than age 65	0.21	−0.89	−0.02
Percent divorced and separated	−0.17	−0.33	0.68
Percent single-parent families	0.06	0.32	0.67
Percent employment	−0.17	−0.06	−0.76

types. Because the six-cluster solution was more likely to represent the theoretical diversity of neighborhoods, we chose it as the final typology. Table 3 showed that neighborhood type 1 was characterized by the highest scores on single-parent families, divorced/separated families, and unemployed residents and a moderate score on neighborhood educational level as well as a large proportion of elderly people. Neighborhood type 2 had the highest score on the concentration of inhabitants younger than 15 years and relatively high scores on single-parent families and unemployment. Neighborhood type 3 was characterized by the lowest scores for the concentration of single-parent families and unemployment and relatively moderate scores on neighborhood educational level and concentration of youth. Neighborhood type 4 was a middle-class neighborhood type with no highest or lowest score for any neighborhood dimension. Neighborhood type 5 was characterized by the lowest educational level and the highest elderly concentration. Neighborhood type 6 was characterized by the highest neighborhood educational level among all types of neighborhoods and moderate levels of single-parent families and elderly concentration.

Multilevel Analysis

Table 4 presents the results of the multilevel analysis using neighborhood dimensions generated by the factor analysis. For chronic diseases, the unadjusted model showed significant effects of neighborhood age structure (OR=0.87), but the effects disappeared after adjusting individual characteristics. For functional limitations, the unadjusted model showed that residents in neighborhoods with lower educational levels and higher proportion of elderly were more likely to have functional limitations (OR=1.18 and 0.90, respectively). After including individual characteristics, the effects of neighborhood education remained (OR=1.17), but the effects of elderly concentration disappeared. For self-rated health, neighborhood education level and neighborhood age structure were associated with self-rated health in unadjusted models (OR=1.10 and 0.92, respectively); however, the effects disappeared after controlling for individual characteristics.

Table 5 presents the results of the multilevel analysis using neighborhood types generated by the cluster analysis. We used type 6 as the reference group because it had the highest education level and thus served as a typical advantaged neighborhood type. In addition, it was represented by a large number of neighborhoods, which can lead to more stable reference comparisons. For chronic diseases, both unadjusted and adjusted models suggested no effects of neighborhood

TABLE 3 Cluster analysis of neighborhood dimensions, Taiwan census data, 1990, 1995, and 2000, N=428

	Type 1 (N=11)	Type 2 (N=56)	Type 3 (N=29)	Type 4 (N=143)	Type 5 (N=58)	Type 6 (N=131)
Neighborhood education[a]	0.28	−0.58	−0.37	0.59	1.06	−0.82
Neighborhood age structure[b]	−0.85	0.98	−0.36	0.79	−0.92	−0.63
Neighborhood family structure and employment[c]	2.45	1.08	−1.38	−0.13	−1.13	0.30

[a]Higher scores represent a lower neighborhood educational level
[b]Higher scores represent more inhabitants younger than 15 years of age
[c]Higher scores represent more single-parent families and unemployed population

TABLE 4 Odds ratios and 95% confidence intervals of chronic diseases, functional limitations, self-rated poor health based on the factor analysis approach, Taiwan Social Change Survey, 1990, 1995, and 2000, N=5,784

	Chronic diseases		Functional limitations		Self-rated poor health	
	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
Neighborhood education[a]	1.00[b] (0.92–1.08)[c]	1.00 (0.91–1.09)	1.18** (1.09–1.28)	1.17** (1.08–1.27)	1.10* (1.02–1.20)	1.05 (0.96–1.14)
Neighborhood age structure[d]	0.87** (0.80–0.94)	1.09 (0.99–1.19)	0.90** (0.83–0.97)	0.96 (0.89–1.04)	0.92* (0.85–0.99)	1.02 (0.94–1.11)
Neighborhood family structure and employment[e]	1.04 (0.96–1.13)	1.07 (0.98–1.17)	1.04 (0.97–1.12)	1.03 (0.95–1.11)	0.96 (0.89–1.04)	0.96 (0.89–1.05)
Age		1.06** (1.05–1.07)		1.02** (1.01–1.03)		1.02** (1.0.1–1.03)
Gender		0.92 (0.80–1.05)		0.65** (0.57–0.75)		0.52** (0.44–0.61)
Education (middle/low)		0.86 (0.72–1.03)		1.04 (0.88–1.23)		0.93 (0.77–1.14)
Education (high/low)		1.02 (0.52–2.00)		1.62 (0.87–3.02)		0.97 (0.41–2.30)
Income (middle/low)		1.13 (0.96–1.34)		0.81* (0.69–0.96)		0.71** (0.59–0.87)
Income (high/low)		1.11 (0.94–1.32)		0.78** (0.66–0.92)		0.70** (0.58–0.85)
Taiwanese		0.93 (0.77–1.13)		0.87 (0.72–1.04)		1.10 (0.88–1.37)
Marriage		0.84* (0.72–0.99)		0.69** (0.59–0.79)		0.72** (0.61–0.85)
Employment		0.75* (0.56–1.00)		0.73* (0.55–0.96)		0.60** (0.44–0.81)

*P<0.05; **P<0.01

[a]Higher scores represent a lower neighborhood educational level

[b]Odds ratios

[c]95% confidence intervals

[d]Higher scores represent more inhabitants younger than 15 years of age

[e]Higher scores represent more single-parent families and unemployed population

TABLE 5 Odds ratios and 95% confidence intervals of chronic diseases, functional limitations, self-rated poor health based on the multivariate-structural approach, Taiwan Social Change Survey, 1990, 1995, and 2000, $N=5{,}784$

	Chronic diseases		Functional limitations		Self-rated poor health	
	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
Type 1[a]	0.92[b] (0.53–1.60)[c]	0.88 (0.49–1.59)	0.87 (0.50–1.50)	0.87 (0.50–1.51)	1.47 (0.84–2.59)	1.42 (0.81–2.50)
Type 2	1.17 (0.85–1.62)	1.23 (0.86–1.75)	1.49* (1.08–2.04)	1.48* (1.07–2.04)	1.75** (1.22–2.51)	1.55* (1.07–2.25)
Type 3	1.38 (0.91–2.09)	1.42 (0.90–2.22)	1.38 (0.92–2.07)	1.41 (0.94–2.12)	1.62* (1.04–2.53)	1.58* (1.01–2.48)
Type 4	1.05 (0.79–1.41)	1.16 (0.84–1.59)	1.32 (0.99–1.76)	1.34* (1.00–1.80)	1.43* (1.03–1.99)	1.33 (0.95–1.87)
Type 5	1.14 (0.86–1.52)	1.33 (0.98–1.82)	1.18 (0.88–1.57)	1.22 (0.91–1.63)	1.32 (0.95–1.85)	1.32 (0.94–1.86)
Age		1.06** (1.05–1.06)		1.02** (1.01–1.02)		1.02** (1.01–1.03)
Gender		0.91 (0.80–1.04)		0.66** (0.58–0.75)		0.52** (0.45–0.61)
Education (middle/low)		0.90 (0.75–1.07)		1.04 (0.89–1.22)		0.95 (0.78–1.15)
Education (high/low)		1.09 (0.56–2.15)		1.55 (0.83–2.88)		0.99 (0.42–2.34)
Income (middle/low)		1.12 (0.95–1.33)		0.79** (0.67–0.93)		0.70** (0.58–0.85)
Income (high/low)		1.10 (0.93–1.30)		0.77** (0.65–0.91)		0.71** (0.58–0.85)
Taiwanese		0.92 (0.76–1.10)		0.88 (0.73–1.05)		1.08 (0.87–1.34)
Marriage		0.86 (0.74–1.00)		0.68** (0.59–0.79)		0.72** (0.61–0.85)
Employment		0.73* (0.55–0.97)		0.72* (0.55–0.95)		0.59** (0.44–0.80)

*$P<0.05$; **$P<0.01$
[a]Reference group: type 6
[b]Odds ratios
[c]95% confidence intervals

types on the opportunity to have a chronic disease. For functional limitations, the unadjusted model suggested that compared to neighborhood type 6, participants living in type 2 were more likely to have functional limitations (OR=1.49). The relationship still existed in the adjusted model (OR=1.48). In addition, participants living in type 4 were also more likely to have functional limitations compared to type 6 (OR=1.34). For self-rated health, individuals living in type 2 (OR=1.75), type 3 (OR=1.62), and type 4 (OR=1.63) neighborhoods were more likely to rate themselves as having poor health than individuals living in type 6. After including individual characteristics, the effects of type 2 and type 3 remained (OR=1.55 and 1.58); however, the effects of type 4 compared to type 6 disappeared, suggesting that individuals living in type 2 and type 3 were more likely to report poor health than individuals in type 6, above and beyond individual characteristics.

DISCUSSION

By using a factor analysis, we derived three neighborhood dimensions: the neighborhood educational level, the age structure, and the neighborhood family structure and employment. This phenomenon reflects that the neighborhood educational level and the neighborhood family structure are two distinct neighborhood dimensions in Taiwan. In contrast, employment loaded with proportions of single-parent families and divorced persons on the same factor. The multivariate analysis further showed that inhabitants living in less educated neighborhoods were more likely to report functional limitations, above and beyond individual characteristics (Table 4). This finding was consistent with prior studies suggesting that neighborhood education is an important neighborhood dimension, which contributes to one's human capital and further influences the outcome of one's health.[31]

In contrast, the cluster analysis showed six neighborhood types with different combinations of the identified neighborhood dimensions. The multivariate analysis further showed that compared to neighborhood type 6, inhabitants living in type 2 and type 4 were more likely to have functional limitations and inhabitants living in type 2 and type 3 were more likely to report worse health (Table 5). Our results suggest that the most adverse neighborhood type is not simply characterized by a higher rate of residents with less than junior high school education, a higher rate of single-parent families, or a higher concentration of elderly. Instead, our results suggest that the characteristics can combine in different clusters to exert a range of effects. Compared with the factor analysis approach, the multivariate-structural approach is thus more likely to explain the diversity and the natural groupings of neighborhood environments. This method is more sophisticated in assessing how individual health is shaped by different contexts with combinations of neighborhood elements.

In order to provide explanations about the health differences between the reference type 6 and types 2, 3, and 4, we mapped out neighborhoods according to their locations. We found that more than half of the type 2 neighborhoods located in remote and mountainous areas were characterized by large percentages of youths and aborigines. Type 3 neighborhoods located in rural midsize townships were characterized by the lowest rates of single-parent families, and type 4 neighborhoods, located in midsize cities and surrounding suburban areas, had an average level of education, age structure, and single-parent families nationally. On the contrary, a majority of type 6 neighborhoods located in the capital city of Taipei were characterized by the highest educational level and a better access to medical care and social welfare programs. Therefore, the inequality within the socioeconomic structure and the accessibility to health and social services may partially explain why types 2, 3, and 4 have worse health profiles than the neighborhoods of Taipei cities in type 6. This explanation is supported by prior studies suggesting that people in remote areas had higher disease-specific standardized mortality rates and lower health care access than other regions in Taiwan.[32,33]

The different findings by different health outcomes suggest that an individual's health status needs to be measured by multiple indicators because each one has its own theoretical meaning.[34,35] While the number of chronic diseases represent one's physical status, status of functional limitations represent behavioral barriers which impact daily functions and self-rated health represents one's own assessment of his/her health. We are not aware of any study investigating the differential effects of neighborhood types using different health outcomes. Therefore, it is difficult to provide possible explanations about different results based on the health measures. However, our results showed that neighborhood types are associated with functional limitation and self-rated health. This may be due to the fact that functional limitation is a direct indicator of how daily functions and social activities can be shaped in an adverse circumstance. Self-rated health reflects a subjective evaluation of whether the adverse environment can be influential. In contrast, chronic disease is a more distal indicator assessing neighborhood impact since individuals may have several chronic diseases but no functional limitations if they manage their illness well. Therefore, its effect is less likely to be shown.

Our findings should be considered in light of the following limitations. The first is the small sample size for some neighborhood types (types 1 and 3). This may have limited our ability to examine how these rarer neighborhood types influence individual health, although the size of our neighborhood sample was relatively large (N=428) compared to previous studies. Second, we did not conduct longitudinal neighborhood measurements, which may have generated a selection bias. The relationship between neighborhood characteristics and health outcomes may have been due to the nonrandom selection of individuals in neighborhoods and not because of neighborhood influences.[36] Specifically, the relationships between neighborhood characteristics and individual health outcomes could be explained by some unmeasured individual characteristics, which would lead to biased estimates of the neighborhood effects. Thus, the relationships found between neighborhood characteristics and individual health should perhaps be more cautiously interpreted as associations rather than as evidence of neighborhood influence. Third, we did not measure the length of time that participants had spent in their neighborhoods or the extent of their exposure to the neighborhood environment. We were thus unable to determine whether effects of neighborhood characteristics on health outcomes were due to cumulative exposure. Fourth, while our survey measures of chronic diseases, functional limitations, and self-rated health are far from perfect, the Taiwan Social Change Survey is one of a few long-term, nationwide surveys from which participants' residential addresses can be released for geocoding. These data have been extensively used to investigate a wide range of topics. Fifth, although this study attempted to introduce a methodological approach using census data, more indicators of the social environment such as the extent of informal social control provided by adults in the neighborhood and the level of residents' participation in local organizations should be included in future studies.[37,38]

CONCLUSIONS

Despite these limitations, the study suggests that a multivariate-structural approach combining factor analysis and cluster analysis could be used to examine neighborhood effects on an individual's health. This approach provides guidelines for researchers to understand neighborhood influences based on the reality of neighborhood distributions. Investigation of neighborhood diversity defined by a combination of well-delineated neighborhood constructs is warranted in future research. A sufficiently large sample that captures different aspects of diversity and indicators of social processes that describe the neighborhood environment would be highly interesting and useful.

REFERENCES

1. Diez Roux AV. Invited commentary: places, people, and health. *Am J Epidemiol.* 2002;155(8):516–519.
2. Macintyre S, Ellaway A, Cummins S. Place effects on health: how can we conceptualise, operationalise and measure? *Soc Sci Med*. 2002;55:125–139.
3. Picket KE, Pearl M. Multivariate analyses of neighbourhood socioeconomic context and health outcomes: a critical review. *J Epidemiol Community Health*. 2001;55:111–122.
4. Sampson RJ, Morenoff JD, Gannon-Rowley T. Assessing "neighborhood effects": social processes and new directions in research. *Annu Rev Sociology*. 2002;28:443–478.
5. Cubbin C, Marchi K, Lin M, et al. Is neighborhood deprivation independently associated with maternal and infant health? Evidence from Florida and Washington. *Matern Child Health J*. 2008;12(1):61–74.
6. Sundquist J, Malmstrom M, Johansson S-E. Cardiovascular risk factors and the neighbourhood environment: a multivariate analysis. *Int J Epidemiol.* 1999;28(5):841–845.
7. Kawachi I, Kennedy BP, Lochner K, Prothrow-Stith D. Social capital, income inequality, and mortality. *Am J Public Health*. 1997;87(9):1491–1498.
8. Lochner K, Kawachi I, Brennan R, Buka S. Social capital and neighborhood mortality rates in Chicago. *Soc Sci Med*. 2003;56(8):1797–1805.
9. Sampson RJ, Raudenbush SW, Earls F. Neighborhoods and violent crime: a multivariate study of collective efficacy. *Science*. 1997;277(15):918–924.
10. Subramanian SV, Kim DJ, Kawachi I. Social trust and self-rated health in US communities: a multivariate analysis. *J Urban Health*. 2002;79(4):s21–s34.
11. Coulton CJ, Cook T, Irwin M. *Aggregation issues in neighborhood research: a comparison of several levels of census geography and resident defined neighborhoods*. Proceedings of the Association for Public Policy and Management Conference, Atlanta, GA; 2004.
12. Coulton CJ, Korbin J, Chan T, Su M. Mapping residents' perceptions of neighborhood boundaries: a methodological note. *Am J Community Psychol*. 2001;29(2):371–383.
13. Haynes R, Daras K, Reading R, Jones A. Modifiable neighbourhood units, zone design and residents' perceptions. *Health Place*. 2007;13(4):812–825.
14. Chow J. Differentiating urban neighborhoods: a multivariate structural model analysis. *Soc Work Res*. 1998;22(3):131–142.
15. Gorsuch RL. *Factor Analysis*. 2nd edn. Hillsdale, NJ: Erlbaum; 1983.
16. Pett MA, Lackey NR, Sullivan JJ. *Making Sense of Factor Analysis: The Use of Factor Analysis for Instrument Development in Health Care Research*. Thousand Oaks, CA: Sage; 2003.
17. Rapkin BD, Luke DA. Cluster analysis in community research: epistemology and practice. *Am J Community Psychol*. 1993;21(2):247–277.
18. Drukker M, Kaplan C, Os JV. Residential instability in socioeconomically deprived neighbourhoods, good or bad? *Health Place*. 2005;11(2):121–129.

19. Aldenderfer MS, Blashfield RK. *Cluster Analysis*. Newbury Park, CA: Sage; 1984.
20. Dupere V, Perkins DD. Community types and mental health: a multivariate study of local environmental stress and coping. *Am J Community Psychol*. 2007;39(1–2):107–119.
21. Davies W. *Urban Social Structure: A Multivariate-Structural Analysis of Cardiff and Its Region*. Cardiff, Wales: University of Wales Press; 1983.
22. Chang Y-H. *Taiwan Social Change Survey Report*. vol. 4-1, Taipei: Institute of Sociology, Academia Sinica, Taiwan; 2000.
23. Chiu H-Y. *Taiwan Social Change Survey Report*. vol. 2-1, Taipei: Institute of Sociology, Academia Sinica, Taiwan; 1990.
24. Chiu H-Y. *Taiwan Social Change Survey Report*. vol. 3-1, Taipei: Institute of Sociology, Academia Sinica, Taiwan; 1995.
25. Johnston RJ. Residential area characteristics: research method for identifying urban subareas-social area analysis and factorial ecology. In: Theodorson GA, ed. *Urban Patterns: Studies in Human Ecology*. University Park, PA: Pennsylvania State University Press; 1982.
26. Wilson WJ. *The Truly Disadvantaged: The Inner-City, the Underclass and Public Policy*. Chicago: The University of Chicago Press; 1987.
27. Browning CR, Cagney KA. Moving beyond poverty: neighborhood structure, social processes, and health. *J Health Soc Behav*. 2003;44(4):552–557.
28. Cagney KA. Neighborhood age structure and its implications for health. *J Urban Health*. 2006;83(5):827–834.
29. Milligan GW, Cooper MC. Methodology review: clustering methods. *Appl Psychol Meas*. 1987;11(4):329–354.
30. Punj G, Stewart DW. Cluster analysis and marketing research: review and suggestions for application. *J Mark Res*. 1983;20:134–148.
31. Sundquist K, Winkleby M, Ahlen H, Sven-Erik J. Neighborhood socioeconomic environment and incidence of coronary heart disease: a follow-up study of 25,319 women and men in Sweden. *Am J Epidemiol*. 2004;159(7):655–662.
32. Chang H, Chou Y. Amending health disparities in Taiwan's indigenous population. *Harvard Health Policy Rev*. 2002;3(2):40–46.
33. Lee MC, Jones AM. Understanding differences in income-related health inequality between geographic regions in Taiwan using the SF-36. *Health Policy*. 2007;83(2–3):186–195.
34. Bailis DS, Segall A, Chipperfield JG. Two views of self-rated general health status. *Soc Sci Med*. 2003;56(2):203–217.
35. Molarius A, Janson S. Self-rated health, chronic diseases, and symptoms among middle-aged and elderly men and women. *J Clin Epidemiol*. 2002;55(4):364–370.
36. Tienda M. Poor people and poor places: deciphering neighborhood effects on poverty outcomes. In: Huber J, ed. *Macro–Micro Linkages in Sociology*. Newbury Park, CA: Sage; 1991:204–212.
37. Ellaway A, Macintyre S. Is social participation associated with cardiovascular disease risk factors? *Soc Sci Med*. 2007;64(7):1384–1391.
38. Hyyppa M, Maki J. Social participation and health in a community rich in stock of social capital. *Health Educ Res*. 2003;18(6):770–779.