Spatial Analysis of Body Mass Index and Smoking Behavior among WISEWOMAN Participants

LEE R. MOBLEY, Ph.D.,[1] ERIC A. FINKELSTEIN, Ph.D., M.H.A.,[1]
OLGA A. KHAVJOU, M.A.,[1] and JULIE C. WILL, Ph.D., M.P.H.[2]

ABSTRACT

Background: The WISEWOMAN program focuses on reducing cardiovascular disease (CVD) risk factors by providing screening and lifestyle interventions for many low-income and uninsured women. To provide the most effective interventions possible, it is important to understand the characteristics of WISEWOMAN participants and their communities.

Methods: We used baseline data collected for WISEWOMAN participants from five states (Connecticut, Michigan, Nebraska, North Carolina, and South Dakota) who had enrolled in WISEWOMAN between January 2001 and December 2002 in order to examine body mass index (BMI) and smoking behavior for evidence of spatial clustering. We then examined whether neighborhood characteristics in clusters of high-risk factors differed from neighborhood characteristics in other locations.

Results: Six percent of the WISEWOMAN participants lived in ZIP codes with high-BMI clusters, and 4% lived in ZIP codes with high-smoking clusters. High-BMI and high-smoking clusters occurred, however, in different locations from each other. The high-BMI-clustered ZIP codes were, on average, located in more disadvantaged areas. Most of the differences between the high-smoking-clustered ZIP codes and the remaining ZIP codes were not statistically significant.

Conclusions: Our analysis revealed spatial clustering in CVD risk factors among WISEWOMAN participants. We also found evidence of a correlation between high-BMI clusters and low socioeconomic status of the surrounding community. A more in-depth analysis of the relationship between risk factors (e.g., BMI) and community characteristics in clustered locations will provide further information concerning the role of the community in affecting individual behavior and should allow for tailoring interventions to reduce these risk factors more effectively.

INTRODUCTION

CARDIOVASCULAR DISEASE (CVD), which includes heart disease, infarctions, and stroke, is the leading cause of death for women in the United States.[1,2] It is also a primary contributor to morbidity and decreased quality of life, especially among older women. Women in lower-income brackets with lower levels of education and without health insurance have an increased risk

of CVD morbidity and mortality,[3] as do women from some racial and ethnic minority groups (i.e., African Americans and Hispanics).[4] Low-income, less educated, uninsured, and minority women have limited access to health services and are more likely to smoke cigarettes, engage in limited physical activity, and have poor nutrition.[5]

In 1995, the Centers for Disease Control and Prevention (CDC) began funding the Well-Integrated Screening and Evaluation for Women Across the Nation (WISEWOMAN) demonstration projects.[6] WISEWOMAN provides CVD screening and intervention services for low-income women aged 40–64 who participate in the National Breast and Cervical Cancer Early Detection Program (NBCCEDP), a cancer-screening program for underinsured and uninsured women. We used baseline data, which were collected when a participant enrolled in WISEWOMAN, to test for spatial clustering of CVD risk factors at the ZIP code level.

Spatial clustering occurs when adjacent ZIP codes have systematically high (or low) median values for a risk factor. Because our analysis is at the ZIP code level rather than the individual level, we tested for ZIP code-level correlations in risk factors. If the correlation in risk factors was caused by variables unique to specific neighborhoods, we would have expected to see clustering in those neighborhood variables. We tested for spatial clustering for two CVD risk factors: elevated body mass index (BMI) and smoking. When the clustering of risk factors was present, we assessed whether unique neighborhood characteristics existed that might be associated with the higher prevalence of BMI and smoking among WISEWOMAN participants in the high-risk-factor clusters.

Neighborhood characteristics are aspects of a person's geographic location that can indirectly influence health by affecting knowledge, attitudes, beliefs, and behavior.[7–11] Wallerstein[12] cites a number of risk factors for poor health that often characterize disadvantaged neighborhoods, including high levels of poverty, poor working conditions, high unemployment, discrimination, and limited social capital (defined as community infrastructure that supports education, health, and welfare). A recent Institute of Medicine report[13] and other studies[14–16] support Wallerstein's claim that neighborhood characteristics, independent of an individual's status in the community, have additional effects on health. For example, the prevalence of race-based residential segregation among counties in urban Michigan was found to be a fundamental cause of observed disparities in health, independent of an individual's race, education, and income.[14]

Studies by Diez-Roux et al.[15,16] show that neighborhood effects are correlated with CVD risk factors. One of their studies found that living in a disadvantaged neighborhood (as measured by the percentage of adults who never completed high school, median household income, area occupational characteristics, and median home value) was associated with increased coronary heart disease (CHD) prevalence and increased levels of risk factors.[15] The increased risk generally persisted even after controlling for individual-level variables. Another study found significant associations between community-level income inequality and three CVD risk factors (BMI, history of hypertension, and sedentariness) among women, particularly at low-income levels (annual household income < $25,000).[16] The associations persisted after adjusting for individual-level income. Community-level income inequality was also positively associated with smoking, but generally only for women at higher income levels.

MATERIALS AND METHODS

Data

WISEWOMAN data have been described previously.[17] We focused on baseline BMI and smoking rates of participants from five states (Connecticut, Michigan, Nebraska, North Carolina, and South Dakota) who had enrolled in WISEWOMAN between January 2001 and December 2002. BMI was defined as weight in kilograms/height in square meters. Smoking was represented by a dichotomous variable that equaled 1 if the woman answered "yes" to a question asking whether she smoked "every day or some days" and equaled 0 otherwise. We obtained neighborhood characteristics by merging the baseline WISEWOMAN data by ZIP code or county with data from the 2000 Census, the 2002 Area Resource File, and the 2000 National Archive of Criminal Justice Data.

Based on our review of the literature and data availability, we included seven county-level variables and three ZIP code-level variables in our

analysis (see Appendix A for variable definitions). The seven county-level variables were (1) index of dissimilarity, (2) proportion of work force in manufacturing jobs, (3) proportion of families in poverty, (4) robbery arrests per 100,000 county residents, (5) proportion of population that was urban, (6) unemployment rate, and (7) median home value. The three ZIP code-level variables were (1) median household income, (2) median earnings of females, and (3) proportion of the adult (≥25 years old) female population with a high school diploma as highest educational attainment.

After deleting all observations with missing geographic or risk factor data, the final number of records was 3364 WISEWOMAN participants representing 717 ZIP codes for the BMI analysis and 4048 participants representing 719 ZIP codes for the smoking analysis.

Analysis

To determine spatial clustering, we used a univariate approach to assess whether ZIP codes with higher and lower median levels for each risk factor among participants tended to cluster geographically. Higher and lower level clusters were defined as those adjacent ZIP codes that systematically had above and below average median risk factor values, respectively. The clustering algorithm is systematic in that clusters are identified based on standardized values of variables and their neighborhood averages.

We employed standard tools for exploratory spatial data analysis (ESDA).[18] In brief, the univariate cluster approach uses spatial statistics that are appropriate for small area aggregate measures (not individual data points) to test for statistically significant clustering in the ZIP code-
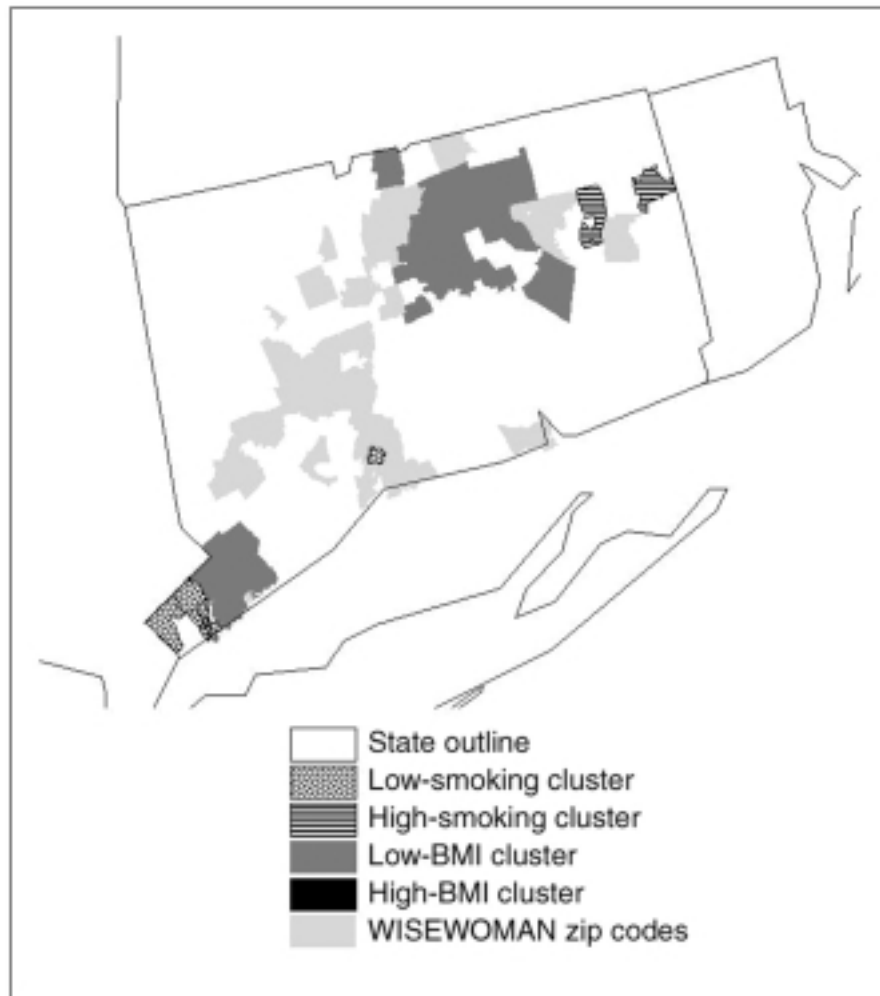

FIG. 1.   Connecticut: clusters of high and low BMI and high and low smoking rates.

level risk measures. A finding of significant clustering suggests that values for the observed risk factor variables are too similar across neighboring ZIP codes to have occurred by chance. After completion of the analysis, we mapped the high and low clusters for BMI and smoking for each state. Analyses were conducted using SpaceStat software (TerraSeer, Ann Arbor, MI) and ArcView GIS (geographic information systems) software (ESRI, Redlands, CA). More details on the spatial-clustering methodology are provided in Appendix B.

After assessing spatial clustering in median BMI and median smoking rates across WISE-WOMAN ZIP codes, we assessed whether the neighborhood characteristics in the high-risk-factor-clustered locations differed from the neighborhood characteristics in the other locations by pooling low-risk-factor-clustered and nonclus-

tered ZIP codes. If neighborhood effects contributed to the observed geographic clustering in BMI or smoking, we would have expected significant differences in the means of the neighborhood variables in the high-risk-factor-clustered locations compared with those in other locations. We assessed the significance of these differences using a t test.

RESULTS

Four percent ($n = 30$) of the included ZIP codes (accounting for 6% [$n = 181$] of participants) were associated with statistically significant clustering of high BMI among participants, and 86% ($n = 26$) of these ZIP codes (accounting for 96% [$n = 178$] of participants in high-BMI-clustered ZIP codes) were in North Carolina. Four percent ($n =$
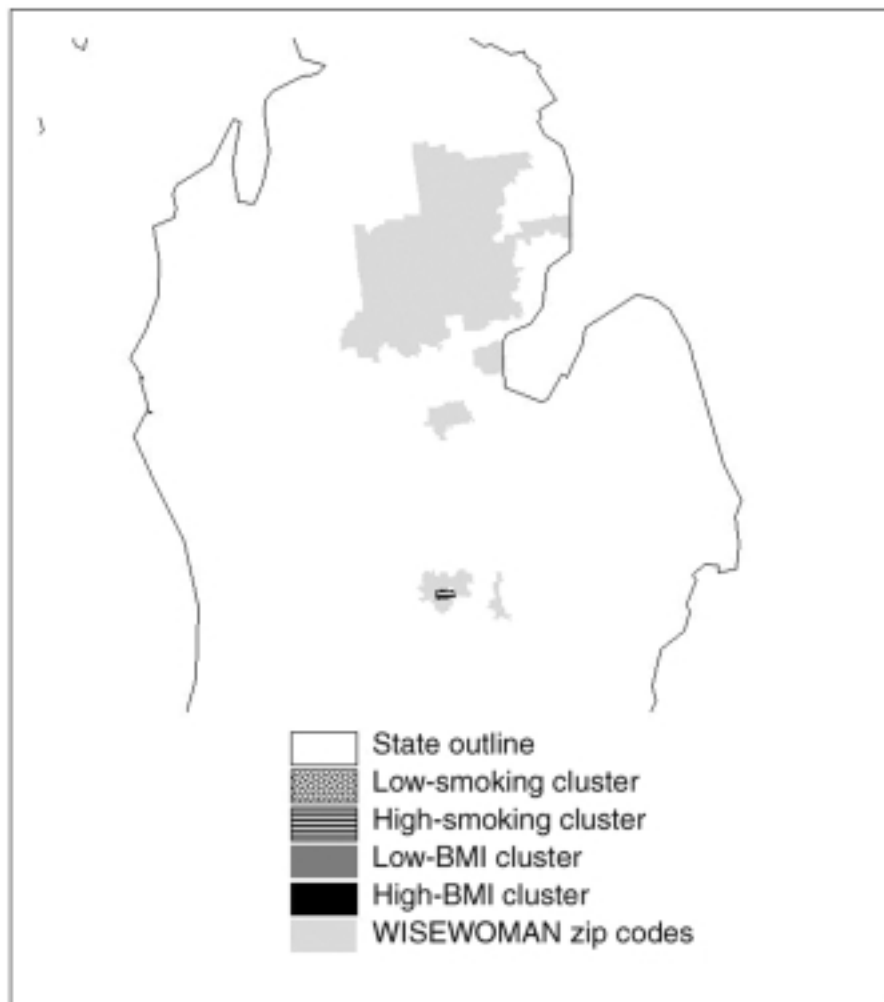

FIG. 2. Michigan: clusters of high and low BMI and high and low smoking rates.

29) of all ZIP codes (accounting for 2% [$n = 65$] of participants) were associated with statistically significant clustering of low BMI among participants, and 82% ($n = 24$) of these ZIP codes (accounting for 78% [$n = 51$]) of participants in low-BMI-clustered ZIP codes) were in Connecticut.

The smoking analysis revealed that 3% ($n = 23$) of included ZIP codes (accounting for 4% [$n = 135$] of participants) showed statistically significant clusters of high smoking rates among WISEWOMAN participants. Fifty-seven percent ($n = 13$) of the ZIP codes with high-smoking clusters were in South Dakota (accounting for 67% [$n = 92$] of participants in high-smoking-clustered ZIP codes), and 4% ($n = 1$) (accounting for 15% [$n = 21$] of participants in high-smoking-clustered ZIP codes) were in Michigan. Six percent ($n = 43$) of all ZIP codes (accounting for 3% [$n = 97$] of participants) were associated with

statistically significant clustering of low smoking among participants; 57% ($n = 25$) of these ZIP codes (accounting for 36% [$n = 35$] of participants in low-smoking-clustered ZIP codes) were in Connecticut, and 39% ($n = 18$) (accounting for 63% [$n = 61$] of participants in low-smoking-clustered ZIP codes) were in North Carolina.

Figures 1, 2, 3, and 4 show the high-risk-factor and low-risk-factor clusters for both BMI and smoking. Figure 1, which shows clustering among participants in WISEWOMAN ZIP codes in Connecticut, indicates that there are no high-BMI clusters and that the low-BMI clusters are in the southwest, northcentral, and central regions of the state. There are low-smoking clusters in the southwest and southcentral regions of Connecticut and high-smoking clusters in the northeast region. Figure 2 shows one high-BMI cluster in the southcentral part of Michigan and no low-BMI or
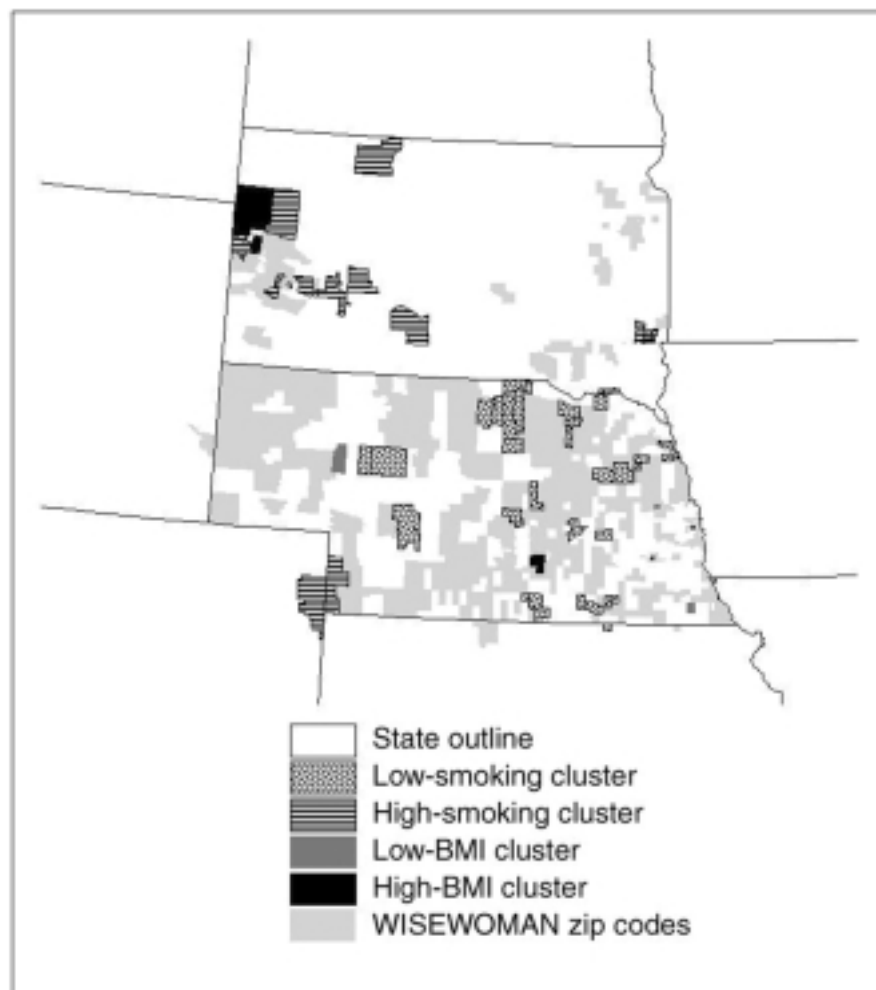

FIG. 3. Nebraska and South Dakota: clusters of high and low BMI and high and low smoking rates.
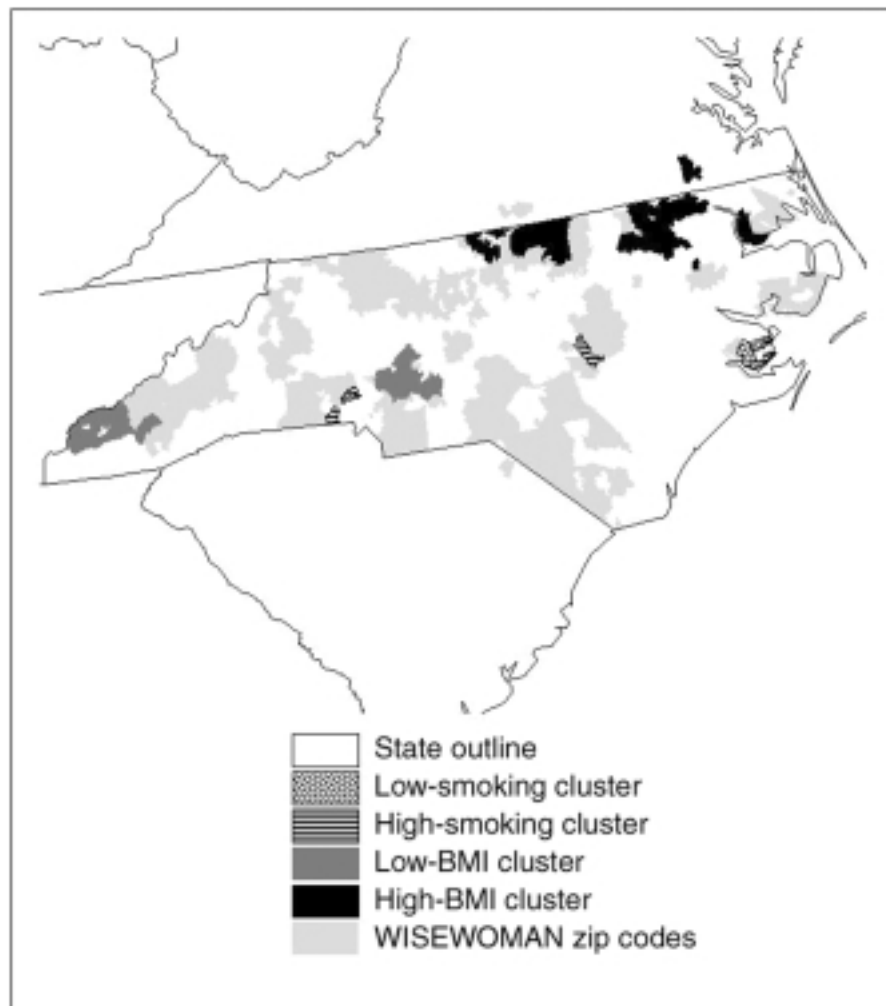
FIG. 4. North Carolina: clusters of high and low BMI and high and low smoking rates.

smoking (neither high nor low) clusters. Figure 3, which depicts clustering in Nebraska and South Dakota, reveals that whereas high-BMI clusters occur in both states, only Nebraska has low-BMI clusters. High-smoking clusters occur in the southwestern corner of Nebraska and the western part of South Dakota, with one high-smoking cluster in the southeast. Low-smoking clusters are geographically dispersed in Nebraska, and there are no low-smoking clusters in South Dakota. Figure 4 shows the results for North Carolina, which has high-BMI clusters in the northcentral part of the state and low-BMI clusters in the western and central parts of the state. There are three high-smoking clusters in the central part of North Carolina and low-smoking clusters on the coast.

Table 1 shows neighborhood variables for high-risk-factor clustered and other locations (including ZIP codes with no clustering and low-risk clusters). High-BMI-clustered ZIP codes (column 1), compared with other BMI ZIP codes (column 2), were, on average, located in more disadvantaged counties, with a higher proportion of the work force in manufacturing jobs (21% vs. 15%), fewer urban residents (30% vs. 46%), and a higher average number of robbery arrests (average of 41 vs. 24 arrests per 100,000 residents). There was also a higher percentage of families in poverty in high-BMI-clustered areas (14% vs. 9%) and a slightly higher unemployment rate (5% vs. 4%). In addition, median county home value, median household income, and median earnings of females were lower for high-BMI-clustered ZIP codes than for other ZIP codes. There were no statistically significant differences, however, in the

TABLE 1. MEANS (SE) OF HIGH-RISK-FACTOR CLUSTERS AND
OTHER ZIP CODES FOR BODY MASS INDEX (BMI) AND SMOKING

Variable	High-BMI-clustered ZIP codes n = 30	Other BMI ZIP codes n = 689	High-smoking clustered ZIP codes n = 22	Other smoking ZIP codes n = 697
County-level variables				
Index of dissimilarity (white vs. other)	0.24 (0.02)	0.29 (0.01)	0.28 (0.02)	0.29 (0.01)
Proportion of work force in manufacturing jobs	21%* (0.01)	15% (0.00)	11%* (0.02)	15% (0.00)
Proportion of families in poverty	14%* (0.01)	9% (0.00)	10% (0.02)	9% (0.00)
Robbery arrests per 100,000 county residents	41* (7)	24 (1)	17 (5)	25 (1)
Proportion of population that is urban	30%* (0.04)	46% (0.01)	56% (0.07)	45% (0.01)
Unemployment rate in 2000	5%* (0.00)	4% (0.00)	3%* (0.00)	4% (0.00)
Median home value (in $1000)	72.54* (3.0)	91.4 (1.7)	82.8 (6.2)	90.9 (1.7)
ZIP code-level variables				
Median household income (in $1000)	29.1* (1.1)	38.4 (0.6)	34.8 (1.9)	38.2 (0.6)
Median earnings of females (in $1000)	15.8* (0.5)	16.6 (0.2)	15.7 (1.0)	16.6 (0.2)
Proportion of the adult (25+) female population with a high school diploma as highest educational attainment	33% (0.01)	34% (0.00)	32% (0.01)	34% (0.00)

*The mean difference between high-risk-factor clusters and other areas is significantly different from 0 at the 5% level.

index of racial dissimilarity or in educational attainment of women for high-BMI-clustered ZIP codes vs. other ZIP codes.

The only significant differences between the high-smoking-clustered ZIP codes (Table 1, column 3) and the remaining ZIP codes (column 4) were the proportion of the work force in manufacturing jobs and the unemployment rate. Areas with high-smoking clusters had a lower proportion of the work force in manufacturing jobs (11% vs. 15%) and a slightly lower unemployment rate (3% vs. 4%).

DISCUSSION

This analysis confirms the presence of spatial clustering for two CVD risk factors among WISE-WOMAN participants. Six percent ($n = 181$) of the participants lived in high-BMI-clustered ZIP codes, and 4% ($n = 135$) lived in high-smoking-clustered ZIP codes. Moreover, the high-BMI and high-smoking clusters occurred in different locations, with many of the high-BMI clusters located in North Carolina.

Unlike the high-smoking clusters, the finding of significant differences in neighborhood characteristics in the high-BMI clusters suggests that these variables may have played a causal role in increasing BMI among the WISEWOMAN participants. The high-BMI-clustered areas had a higher average unemployment rate, a greater proportion of families in poverty, lower earnings, and higher crime rates. Each of these factors may have adversely influenced participants' BMI by affecting the participants' ability to eat healthy foods and engage in regular physical activity. Further analysis of the relationship between CVD risk factors (e.g., high BMI) and neighborhood characteristics in clustered locations represents an important step in developing intervention strategies. For example, an in-depth analysis of high-BMI clusters might reveal that concerns about community safety limit women's ability to exercise, indicating a need for strategies to provide participants with access to safe exercise facilities.

Several factors may influence our BMI and smoking results. First, the cluster analysis was carried out at the ZIP code level and focused on the median value of each risk factor among WISE-WOMAN participants living in the different ZIP

code areas. Because some ZIP codes had few participants during the period in question, there is likely to be substantial uncertainty associated with the median values.

Because of data limitations, the means comparison of many of the socioeconomic variables among clusters of high-BMI and non-high BMI and smoking was at the county level, whereas the clustering analysis was at the ZIP code level. Because some counties contained both clustered and nonclustered ZIP codes and because of variance in socioeconomic variables across ZIP codes within counties, the county comparison introduced additional measurement error. Moreover, both the cluster and means analyses were univariate. Multivariate analysis might provide more insight into the role that individual and neighborhood characteristics play in influencing CVD risk factors. For example, it may be that participants in the high-BMI ZIP codes were more likely to be African American. African American women have, on average, higher BMIs than do white women and live in communities with lower socioeconomic status.[13,19] Although confirmation of this hypothesis through multivariate analysis would not change the conclusions noted, it might provide further information concerning how best to tailor interventions to specific communities.

In summary, we found evidence of spatial clustering in BMI and smoking behavior among WISEWOMAN participants. Although we cannot reject the hypothesis that neighborhood characteristics were similar in high-smoking and low-smoking neighborhoods, our BMI analysis is consistent with other studies that find higher BMI among African Americans and people of lower socioeconomic status.[13,19] A more in-depth analysis of the relationship between risk factors and neighborhood characteristics in clustered locations will provide further information concerning the role of communities in affecting individual behavior and should allow for tailoring interventions to reduce these risk factors more effectively.

Spatial analysis is a relatively young field and has only recently been applied to health services research. Our analysis shows that relatively simple spatial methods can provide valuable information for program planning and evaluation purposes. Continued use of spatial analysis in health services research will further contribute to our understanding of the interrelationships between neighborhood variables and health.

REFERENCES

1. Casper MI, Barnett E, Halverson JA, et al. Women and heart disease: An atlas of racial and ethnic disparities in mortality. Morgantown, WV: Office for Social Environment and Health Research, West Virginia University, 1999.

2. Mosca L, Manson J, Sutherland S, Langer R, Manolio T, Barrett-Conner E. Cardiovascular disease in women: A statement for healthcare professionals from the American Heart Association. Circulation 1997;96: 2468.

3. Shumaker SA, Czajkowski SM. A review of health-related quality of life and psychosocial factors in women with cardiovascular disease. Ann Behav Med 1993;15:149.

4. Casper ML, Barnett E, Williams GI Jr, Halverson JA, Braham VE, Greenlund KJ. Atlas of stroke mortality: Racial, ethnic, and geographic disparities in the United States. Atlanta, GA: Department of Health and Human Services, Centers for Disease Control and Prevention, 2003.

5. Ford ES, Will JC, De Proost Ford MA, Mokdad AH. Health insurance status and cardiovascular disease risk factors among 50–64-year-old U.S. women: Findings from the third National Health and Nutrition Examination Survey. J Wom Health 1998;7:997.

6. Centers for Disease Control and Prevention. WISE-WOMAN. Available at *www.cdc.gov/wisewoman/* Accessed June 1, 2003.

7. Porell W, Miltiades HB. Regional differences in functional status among the elderly. Soc Sci Med 2002;54: 1181.

8. Stineman M. A model of health environmental integration. Top Stroke Rehabil 2001;8:34.

9. Von Korff M, Koepsell T, Curry S, et al. Multilevel analysis in epidemiological research on health behaviors and outcomes. Am J Epidemiol 1992;132:1077.

10. Kington R, Carlisle D, McCaffrey D, et al. Racial differences in functional status among elderly U.S. migrants from the south. Soc Sci Med 1998;47:831.

11. Duncan C, Jones K, Moon G. Health-related behavior in context: A multilevel modeling approach. Soc Sci Med 1996;42:817.

12. Wallerstein N. Empowerment to reduce health disparities. Scand J Public Health 2002;30(Suppl 59):72.

13. Smedley BD, Stith AY, Nelson AR, eds. Institute of Medicine. Unequal treatment: Confronting racial and ethnic disparities in healthcare. Washington, DC: The National Academies Press, 2002.

14. Schultz A, Williams D, Israel B, Lempert L. Racial and spatial relations as fundamental determinants of health in Detroit. Milbank Q 2002;80:677.

15. Diez-Roux AV, Nieto FJ, Szklo M. Neighborhood en-

vironments and coronary heart disease: A multilevel analysis. Am J Epidemiol 1997;146:48.

16. Diez-Roux A, Link B, Northridge M. A multilevel analysis of income inequality and cardiovascular disease risk factors. Soc Sci Med 2000;50:673.

17. Finkelstein EA, Khavjou OA, Mobley LR, Haney DM, Will JC. Racial/ethnic disparities in coronary heart disease risk factors among WISEWOMAN enrollees. J Wom Health 2004;13:xxx.

18. Anselin L. Interactive techniques and exploratory spatial data analysis. In: Longley P, Goodchild M, Maguire D, Rhind D, eds. Geographical information systems, 2nd ed. New York: John Wiley & Sons, 1999;1:253.

19. The Surgeon General's call to action to prevent and decrease overweight and obesity. Available at *www.surgeongeneral.gov/topics/obesity* Accessed August 25, 2003.

20. Massey DS, Denton NA. The dimensions of residential segregation. Soc Forces 1988;67:281.

21. Williams DR, Collins C. Racial residential segregation: A fundamental cause of racial disparities in health. Public Health Rep 16:116:404.

22. Getis A. Spatial statistics. In: Longley P, Goodchild M, Maguire D, Rhind D, eds. Geographical information systems, 2nd ed. New York: John Wiley & Sons, 1999;1:239.

23. Anselin L. Local indicators of spatial association—LISA. Geographical Anal 1995;27:93.

Address reprint requests to:
Lee R. Mobley, Ph.D.
RTI International
Health, Social and Economics Research
3040 Cornwallis Road
P.O. Box 12194
Research Triangle Park, NC 27709

E-mail: lmobley@rti.org

APPENDIX A. DETAILS ON VARIABLE CONSTRUCTION

In this appendix, we provide a definition for each community-level variable used in the analysis. We describe how the variable was created and from what dataset it was obtained.

County-level variables

Index of dissimilarity. Residential segregation is defined as the degree to which two or more racial/ethnic groups live separately from one another.[20] One of the measures of residential segregation that has been used extensively in previous research is the index of dissimilarity.[14,20,21] In our analysis, the index measures how whites and nonwhites are distributed across the ZIP codes that make up each county. We used the 2000 Census data to create this variable. The index ranges between 0 and 1 and is calculated as follows:

$$D = 0.50 \sum_{z=1}^{z} \left| \frac{Pzw}{Pw} - \frac{Pznw}{Pnw} \right|$$

where *Pzw* is the number of white people in the ZIP code, *Pw* is the number of white people in the county, *Pznw* is the number of nonwhite people in the ZIP code, *Pnw* is the number of nonwhite people in the county, and Z is the number of ZIP codes in the county.

For example, a value of 0.80 for the dissimilarity index can be interpreted as follows: 80% of whites in the county would have to move from some ZIP codes to others to produce a completely even distribution of whites and nonwhites across all ZIP codes in the county. An index of 0 means each ZIP code has the same ratio of whites to nonwhites as the county ratio; an index of 1 means that no whites share their ZIP code with nonwhites, and *vice versa*.

Manufacturing jobs, families in poverty, unemployment rate, and median home value. We used the 2002 Area Resource File to create variables for the proportion of the county's work force in manufacturing jobs and the proportion of families in poverty and to obtain county unemployment rates and median home values.

Robbery arrests. We obtained data on the number of robbery arrests per 100,000 county residents from the 2000 National Archive of Criminal Justice Data.

Urban population. We used the 2000 Census data to create the county-code-level variable for the proportion of the population that was urban.

ZIP code-level variables

Median household income, median earnings of females, and adult female population with high school diploma only. We obtained median ZIP code household income and median ZIP code earnings of females from the 2000 Census. We also used the 2000 Census data to create the ZIP code-level

variable for the proportion of the adult (≥25 years old) female population with a high school diploma as their highest level of educational attainment.

APPENDIX B. ANALYSIS OF SPATIAL CLUSTERING

Spatial clustering (or spatial autocorrelation) in a variable manifests when similar values of the variable are found in geographic proximity. Two different statistics have been used extensively to assess spatial autocorrelation: Geary's c and Moran's I.[22] Moran's statistic is based on the covariance among values found at associated locations, much like the more familiar Pearson correlation coefficient, which compares associated variables. The association among locations for Moran's I is defined using a spatial weights matrix. The weights matrix describes for each location all of the closest neighboring locations and is derived from a distance matrix that contains distances between all possible pairs of locations. The researcher must define the neighborhood set for each observation by determining how the spatial weights matrix is to be constructed. A simple approach is to define the k closest observations as the neighborhood set. For this analysis, we let $k = 6$.

Moran's I statistic is a global statistic. We calculated Moran's I using all data points from all five states with WISEWOMAN participants. Because we expected differences in the strengths of association from region to region, we were most interested in employing a local Moran test, which detects local spatial autocorrelation in group-level data.[23] We used the median value for body mass index (BMI) among all WISEWOMAN enrollees in each ZIP code and the proportion of enrollees in the ZIP code who were smokers as our ZIP code-level measure of BMI and smoking behavior, respectively. We then looked for evidence of spatial clustering in the ZIP code-level BMI and smoking behavior measures.

The local Moran decomposes the global Moran's I statistic into contributions for each location, termed local indicators of spatial association (LISAs). Anselin[23] defines the local Moran statistic for observation I as follows:

$$I_i = p_i \sum_j w_{ij} p_j$$

The local Moran statistic is based on the gamma index, a general index of matrix association. In this equation, p_i is the difference between the median risk value in ZIP code i and the mean overall ZIP code-specific risk values. The w_{ij} represents the strength of connection between ZIP codes i and j, developed from the neighbor information in the spatial weights matrix. This weights matrix ensures that only neighboring p_js are included in the statistic. The local Moran LISA statistic will be positive when median risk factor values at neighboring ZIP codes are similar and will be negative when they are dissimilar. The statistical significance of these local LISA values is evaluated in SpaceStat using Monte Carlo randomization.