Transactions of the Illinois Academy of Science
(1988), Volume 81, 1 and 2, pp. 147-152

SOME POSSIBLE FACTORS IN THE GEOGRAPHIC VARIATION OF ISCHEMIC HEART DISEASE MORTALITY RATES

Anthony J. Dzik, Ph.D.
Division of Social Sciences
Shawnee State University
Portsmouth, Ohio 45662

ABSTRACT

Regional differences exist in ischemic heart disease mortality rates in the United States. These differences appear to be partially related to geographic variations in environmental risk factors. Age-adjusted county mortality rates (sample of 101 counties) were examined as to their associations with seven environmental variables. Inverse relationships with altitude and drinking water hardness are noted while positive associations are found for atmospheric sulfur dioxide concentration, snowfall, and weather instability. A multivariate regression model was formulated and suggests that 44 percent of the variance in ischemic heart disease mortality rates can be explained by a combination of altitude, air pollution, income level, and snowfall frequency. The remaining variance may be accounted for by other environmental factors and perhaps by personal factors that may be difficult to quantify.

INTRODUCTION

Regional variations in cardiovascular disease mortality rates exist in the United States. Rates are higher in the southwestern part of the Atlantic Coastal Plain and in the Manufacturing Belt of the Northeast than they are in the Mountain States or Great Plains. Previous studies [Enterline and Stewart, 1956; Sauer, 1962; Pyle, 1971; Sauer, 1978] seem to indicate that this general pattern has remained relatively constant over the past thirty years. It is virtually impossible to attribute such geographical differences solely to personal factors, such as genetic composition, occupation, dietary habits, tobacco, etc. [Kleinman et al., 1981; Meade, 1983]. Environmental risk factors or characteristics of place may have better explanatory value. Some of these extrinsic factors that may play roles in the etiology of ischemic heart disease are air quality, drinking water hardness, weather and climate, and socio-economic stress [Bean, 1938; Daubert, 1952; Stallones, 1965; Carroll, 1966; Schroeder and Kraemer, 1974; Rogot and Padgett, 1976; Singh et al., 1977]. Such factors are not randomly distributed in space and correlation of the spatial variation of any such variable with variation in ischemic heart disease mortality may suggest some causal relationship.

A study was undertaken to try to determine which environmental factors might explain the spatial variation of ischemic heart disease mortality rates in the United States. Environmental factors were evaluated through the construction and testing of a linear regression model that may explain the degree to which these extrinsic factors explain mortality variations.

METHODOLOGY

One hundred and one counties in the coterminous United States were selected as the sample for analysis. Because spatial contiguity presents problems of social, economic, and environmenal interaction between adjacent enumeration units, a random areal sample was chosen in a manner that would preserve geographical separation so as to avoid spatial autocorrelation and to reduce measurement errors. Ischemic heart disease (ISCN 410-413) mortality rates (mean annual rate 1976-1978) were computed for each county and age-adjusted to the 1980 U.S. population structure. These rates serve as the dependent variable.

Based upon literature review, seven environmental variables were selected. These variables are (1) Drinking Water Hardness (as ppm $CaCO_3$), (2) Mean Annual Number of Days With Snowfall Exceeding One Inch (adjusted for latitude), (3) Index of Weather Instability (annual number of days with unsettled weather), (4) Mean Elevation Above Sea Level, (5) Air Pollution (indexed by the average annual arithmetic mean concentration of sulfur dioxide [see Dzik and Cember, 1985]), (6) Median Family Income, and (7) Percentage of the Labor Force That is Unemployed.

In order to develop a predictive model, Pearson product-moment correlation coefficients were computed for each of the environmental variables and the mortality rates to determine the degree to which individual factors might be associated with the death rates. Stepwise regression was performed to obtain a regression model which might be useful in explaining geographic variation in ischemic heart disease mortality rates.

RESULTS AND DISCUSSION

Correlation coefficients are presented in Table 1. IHD mortality rates were significantly ($p < .01$) associated with five of the seven variables. Elevation exhibited the highest coefficient ($r = -.380$). Sulfur dioxide concentrations were also moderately correlated with the mortality rates (r = + .365). An inverse relationship was found for water hardness (r = –.318). There was a positive association between mortality rates and the annual number of days with snowfall (r = + .310) and a weaker association with weather instability (r = + .296). These bivariate associations appear to be in accord with the findings of earlier studies [Daubert, 1952; Schroeder, 1960; Carroll, 1966; Dudley et al., 1969; Rogot, 1974; Singh et al., 1977; Baker-Blocker, 1982].

Ischemic heart disease death rates exhibited no significant relationships with the two socio-economic variables (unemployment and median family income), but the signs of the coefficients are in agreement with the hypotheses concerning these variables. That is not to say that the spatial variations of IHD mortality rates cannot be partly explained by these variables for there may be interactions which cannot be accounted for by simple bivariate analysis. Because none of the independent variables exists alone in space, the desired explanatory model must be multivariate.

A forward addition stepwise regression procedure was applied to the data using the Statistical Package for the Social Sciences (SPSS) computer program. Stepwise regression is one method for selecting the "best" regression equation ("Best" implies that the model is the one that produces the highest r^2 statistic, when only statistically significant coefficients are included, and also that possesses the least amount of deviation about the regression line).

The multivariate model appears in Table 2. This explanatory model uses four environmental variables, elevation, annual number of days with snowfall, median family income, and sulfur dioxide concentration, to explain 44 percent of the variance in the IHD death rates. The model suggests that higher rates are likelier to be in areas of cold climate at low elevations where there is considerable air pollution and incomes are low. The predictive power, however, is somewhat limited, as indicated by a rather wide spread about the regression line.

The relative importance of any two predictor variables in a regression model can be obtained by taking the ratio of the squares of their respective standardized regression coefficients [Kachigan, 1982]. When this method is applied to the model, elevation accounts for about 1.4 times as much of the variance as does the snowfall variable and for 3 times as much as does sulfur dioxide. The standardized regression effect [St. Leger and Sweetnam, 1979] or the percentage change in the death rate that is induced by increasing an environmental variable by one standard deviation while all other independent variables are held constant shows that elevation induces the largest change (13.1%) and sulfur dioxide the least (7.2%).

The model confirms several of the hypothesized relationships, namely that elevation, air pollution, income, and snowfall characteristics may play some part in explaining the ischemic heart disease mortality rate of a place. The relationships concerning water hardness, unemployment, and weather instability were not substantiated by the model. The statistical problem of multicollinearity could be suspected.

Multicollinearity is a condition in which several independent variables are strongly related to one another in a linear relationship. With multicollinearity it is difficult to interpret the strengths of each predictor variable. It may be that a particular variable's relevance is clouded by the presence in the model of another variable to which it is related. Tests, suggested by Kmenta [1971], indicated that multicollinearity in the variable set was not a problem.

Another question concerning the model is the unexplained variance. Factors of the physical and socio-economic environment, such as air temperature, cardiotoxins in the soil, and occupational structures (which were not evaluated in this research), are perhaps likely candidates for explaining some of the remaining variance. Part of the unexplained variance might also be related to personal risk factors that are difficult to quantify.

CONCLUSIONS

The regression model suggests that 44 percent of the variance in ischemic heart disease mortality rates in the U.S. can be explained by a combination of four environmental features. The model indicates that, on the average, IHD mortality rates will be greater in counties at low elevations, with greater frequency of heavy snowfall, with lower income levels, and with considerable air pollution than in places at higher altitude that have infrequent snowfall, higher income levels, and little air pollution.

There are, of course, some problems with using aggregate data like that utilized here, but the results of this examination lend further credence to several hypotheses concerning the epidemiology of ischemic heart disease, namely:

1. High altitude dwellers may be in some way protected from certain forms of cardiovascular disease because of their physiologic attributes that developed in response to barometric pressure conditions at higher altitudes.

2. Atmospheric contamination has an aggravating role in cardiovascular disorders as certain pollutants disrupt normal pulmonary functions, which in turn stresses the cardiovascular system.

3. Stress from poor living conditions may be a contributory factor in ischemic heart disease. Poverty itself is probably a factor as it can be a barrier to accessing preventative and curative treatment.

4. Snowfall may bring about physical strain, mental stress, and access to service problems, all of which may enhance the risk of mortality from ischemic heart disease.

With the apparent multifactorial etiology of ischemic heart disease, further research into the interactions of environmental and personal risk factors may be the key to control of this disease. Further examination of suspected high risk physical and social environments is necessary.

LITERATURE CITED

Baker-Blocker, A. 1982. Winter Weather and Cardiovascular Mortality in Minneapolis-St. Paul. Am J Public Health 72:261-265.

Bean, W.B. 1938. Coronary Occlusion, Heart Failure, and Environmental Temperature. American Heart Journal 16:701.

Carroll, R.E. 1966. The Relationship Between Cadmium in the Air to Cardiovascular Disease Death Rates. J Am Medical Association 198:177-179.

Daubert, K. 1952. Tropospharische Einflusse Bei Per Entstehung der Thrombose und Embolie. Berichte Des Deutschen Wetterdienstes 42:418-552.

Dudley, E.F., R.A. Beldin, and B.C. Johnson 1969. Climate, Water Hardness, and Coronary Heart Disease. J Chronic Diseases 22:25-48.

Dzik, A.J. and H. Cember 1985. Power Plant Emissions and Cancer Mortality Rates. Trans. Ill. Acad. of Sciences 78:179-182.

Enterline, P.E. and W.H. Stewart 1956. Geographic Patterns in Deaths from Coronary Heart Disease. Public Health Reports 71:849-855.

Kachigan, S.K. 1982. *Multivariate Statistical Analysis: A Conceptual Introduction* (Radius Press, New York).

Kleinman, J.C., V.G. DeGruttola, B.B. Cohen, and J. Madans 1981. Regional and Urban-suburban Differentials in Coronary Heart Disease Mortality and Risk Factor Prevalence. J Chronic Diseases 34: 11-19.

Kmenta, J. 1971. *Elements of Econometrics* (MacMillan, New York).

Meade, M.S. 1983. Cardiovascular Disease in Savannah, Georgia. in McGlashan and Blunden eds. *Geographical Aspects of Health* (Academic Press, London).

Pyle, G.F. 1971. *Heart Disease, Cancer, and Stroke in Chicago*. Univ. of Chicago Dept. of geography Research Paper # 134.

Rogot, E. 1974. Associations Between Coronary Mortality and the Weather in Chicago. Public Health Reports 89:330-338.

Rogot, E. and S. Padgett 1976. Associations of Coronary and Stroke Mortality with Temperature and Snowfall in Selected Areas of the United States 1962-1966. Am J Epidemiology 103:565-575.

Sauer, H.I. 1962. Epidemiology of Cardiovascular Mortality- Geographic and Ethnic. Am J Public Health 52:94-105.

Sauer, H.I. 1978. Two-hundred Extreme-rate U.S. Counties from Cardiovascular Diseases. in *Geochemistry and the Environment vol. III* (National Academy of Sciences, Washington).

Schroeder, H.A. 1960. Relations Between Hardness of Water and Death Rates from Certain Chronic and Degenerative Diseases in the United States. J Chronic Diseases 12:586-591.

Schroeder, H.A. and L. Kraemer 1974. Cardiovascular Mortality, Municipal Water, and Corrosion. Arch Environmental Health 28:303-311.

Singh, I., I.S. Chohan, M. Lal et al. 1977. Effects of High Altitude Stay on the Incidence of Common Diseases in Man. Int J of Biometeorology 21:93-122.

St. Leger, A. and P. Sweetnam 1979. Statistical Problems in Studying Relative Specificities of Association Between Environmental Agents and Disease, a Solution Suggested. Int J of Epidemiology 8:73-77.

Stallones, R.A. 1965. Epidemiology of Cardiovascular Disease, a Review. J Chronic Diseases 18:859-872.

Table 1. Associations between ischemic heart disease mortality rates and environmental factors.

Environmental Variable	HR*	r
Mean Elevation Above Sea Level	–	–.380**
Annual Arithmetic Mean Concentration of Sulfur Dioxide	+	.365**
Drinking Water Hardness	–	–.318**
Mean Annual Number of Days With Snowfall Exceeding One Inch (Adjusted for Latitude)	+	.310**
Weather Instability	+	.296**
Per Cent Unemployed	+	.176
Median Family Income	–	–.092

* HR = Hypothesized direction of relationship between environmental variable and IHD death rates.

** Coefficient significant at $p < .01$, N = 101.

Table 2. Regression model for ischemic heart disease death rates. Equation: $Y = 451 - 0.031X_1 + 336X_2 - 0.013X_3 + 3.43X_4$. Standard deviation of Y about the regression line: 95. $r^2 = .44$. $F = 17.7$ (significant at $p < .01$, N = 101).

Variable	Standardized Coefficient	t Statistic
X_1 : Mean Elevation Above Sea Level	–.470	–5.6
X_2 : Mean Annual Number of Days With Snowfall Exceeding One Inch (Adjusted for Latitude)	.394	4.6
X_3 : Median Family Income	–.370	–4.4
X_4 : Annual Arithmetic Mean Concentration of Sulfur Dioxide	.277	3.2

All coefficients significant at $p < .01$, N = 101.