

COMPUTER MAPPING OF CANCER MORTALITY BY CENSUS TRACT:

COLUMBUS, OHIO 1956-1974*

Eugenia E. Calle
Health and Safety Research Division
Oak Ridge National Laboratory**
Oak Ridge, Tennessee 37830

Introduction

A phenomenon of considerable epidemiologic interest in cancer research is the pronounced effect of geographic location on cancer mortality. Most cancer rates vary markedly by country and within regions of the same country. These differences have given rise to national and international research in an attempt to identify the environmental and demographic factors associated with high risk areas.

The technique of mapping provides a most effective means of describing the geographic differences in cancer rates, identifying patterns of elevated mortality, and generating etiologic hypotheses.

*Research supported by National Library of Medicine Grant No. 5T01 LM00159 and The Ohio State University, Department of Preventive Medicine.

**Operated by Union Carbide Corporation under contract W-7405-eng-26 with the U.S. Department of Energy.

Mortality maps have been widely used to display cancer frequency at varying geographic scales of analysis (Dunham and Bailar, 1968; Burbank, 1971; Mason et al., 1975). While the majority of studies have employed fairly large areal units of comparison (e.g. countries, states, counties), some have focused on the intracommunity variations in cancer rates within an urban area (Cohart, 1954; Greenburg et al., 1967; Pyle, 1971; Henderson et al., 1975).

The present study examines the geographic variation of cancer mortality in Columbus, Ohio using the Poisson distribution to judge the significance of standard mortality ratios adjusted for age, sex, and race. Total Columbus age, sex, and race-specific rates for a variety of cancer sites are used to generate expected numbers of deaths. The census tract serves as the basic unit of analysis and comparison of intracommunity risk.

Methodology

Mortality tapes from the Ohio Department of Health (ODH) provided the main data source for the study. The tapes document all deaths which occurred in Ohio during the 19-year period 1956 to 1974; each mortality record (for Franklin county) also includes the census tract of residence at the time of death. The causes of death considered were limited to 16 of all the most common individual cancers and the combined category of 'all malignant neoplasms'. Cancers less common than these would yield (over the 19-year period) too few cases in each census tract to allow for any meaningful description or analysis.

Because three revisions of the International Classification of Disease (ICD) were used to code cause of death on the ODH tapes, the more specific seventh and eighth revisions were converted to the less specific sixth revision. Following conversion of ICD codes, the mortality tapes yielded number of deaths from each cancer in each census tract stratified by race (white, nonwhite), by sex, and by 10-year age groups.

Population data by census tract are available on summary tapes for the 1960 and 1970 censuses and in printed form for the 1950 census. Intercensal values were estimated by linear interpolation and estimates for the years after 1970 were derived by extrapolation based on

the growth rate of the 1960's. Total population at risk values for each tract in each age and race-sex group were then computed by summing across all 19 years.

Three types of situations required some modification to correct for changes in census tract boundaries over the three censuses. The first, and the easiest to remedy, results from simple population growth and consequent subdivision of already existing tracts. To achieve comparability, any tracts that were subdivided in 1960 or 1970 were recombined and assigned to the original 1950 tract number. All mortality and population data from these subdivided tracts were also combined and assigned to the larger 1950 tract.

The second situation arises due to changes in individual tracts lying inside (i.e., not on the periphery of) the study area. Changes in census tract boundaries occur between each census for a variety of reasons and to varying degrees. Block maps and block statistics were used to make rough estimates of the degree of geographic and population changes that resulted from boundary changes. In most cases, tract boundary changes inside the study area involved only one or two blocks and less than 10 percent of the population. However, where changing boundaries between tracts shifted at least 10 percent of the tract's population, these contiguous tracts were combined into larger tract groups, thus maintaining stable geographic units across the three census periods.

The third, and most troublesome, situation results from changes in the outer boundaries of tracts that lie on the periphery of the study area. Because these tracts delineate the external borders of the study area, they could not be combined with other tracts to compensate for boundary changes. Major boundary changes occurred for three border tracts between the 1960 and 1970 censuses. In each case, these tracts became much larger (geographically) with the 1970 census, encompassing land lying outside of the study area. The problem lies in estimating the 1970 population values of the original tract area, i.e., the change that would have occurred in the tract's population between 1960 and 1970 if the tract had not experienced such a large geographic addition. The 1970 Census of Population values cannot be used directly; presumably these values are too large. In order to estimate the

1970 population values, growth rates for each race, sex, and age group that can then be applied to the 1960 population values are needed. For each tract, these growth rates were derived from an average of the growth rates in the five closest neighboring tracts.

Many previous studies (Cohart, 1954; Greenburg, et al., 1967; Henderson et al., 1975) aggregated census tracts according to some predefined criteria of homogeneity (socioeconomic status, air pollution, demographic factors). While grouping on a certain factor facilitates correlation of cancer rates with that factor, it also tends to obscure other relationships that might exist. Grouping was not attempted in this study because hypotheses relating cancer to specific geographic or demographic influences were not being tested. This study is, rather, a descriptive study, and as such, attempts to disclose any geographic patterns in cancer mortality that are present in the study area. Consequently, no grouping of census tracts has been done other than the aggregation necessary to insure equivalent units of study between the three census periods.

From the ODH mortality data and the population data, age-specific mortality rates were computed for the study area as a whole (all tracts combined) for each of the 17 cancer sites and each of the four race-sex groups. These age-specific mortality rates were then applied to the population values in each individual tract to arrive at an expected number of deaths for each of the race-sex groups in each tract and for each cancer site. The entire study area thus served as the standard for each individual census tract. These unexpected deaths were totalled across the race-sex groups so that, ultimately, one age, sex, and race-adjusted standard mortality ratio (SMR) was computed for each tract and each cancer site.

In an attempt to reveal geographic variation in cancer mortality between census tracts, the tracts were grouped into five classes according to their SMR's: higher and lower than the entire study area at $p \leq 0.05$; higher and lower than the entire study area at $0.05 < p \leq 0.2$; and no different from the study area. The choice of a significant level $p \leq 0.05$ is dictated by convention while the less conservative level of $p \leq 0.2$ is used to highlight geographic clusters or patterns that might exist. Attention is thus focused on a tract significant

at the $p \leq 0.2$ level if it borders or is surrounded by tracts significant at the $p \leq 0.05$ level. These five classes are represented visually by five different shading schemes on maps produced of the study area. All maps were computer generated and produced by a light pen plotting on film, and then printed following photographic processing.

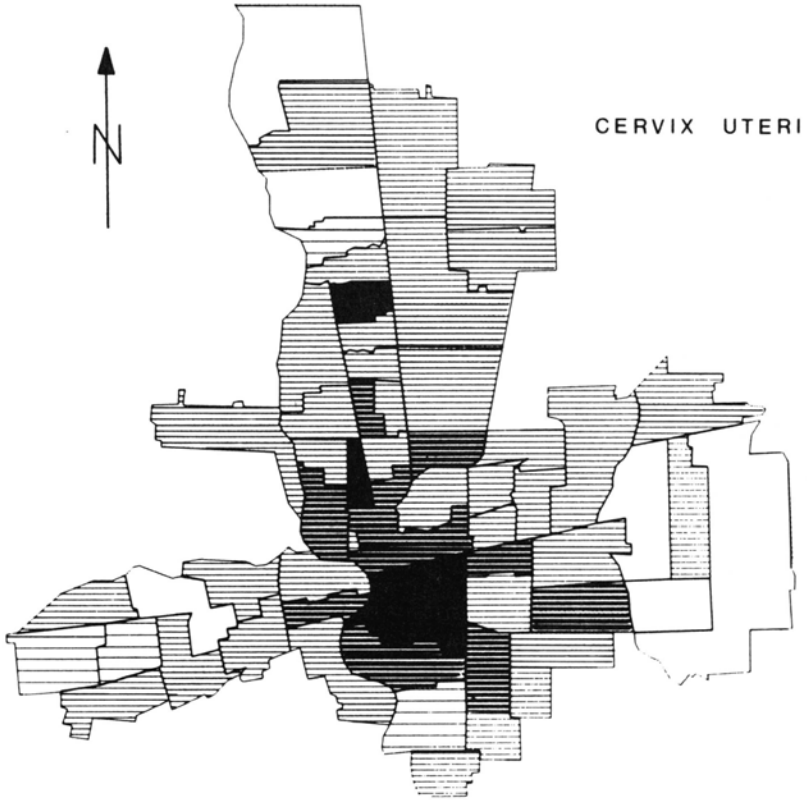
Results and Conclusions

The finished maps provide a visual indication of variation between census tracts for each individual cancer site and for 'all malignant neoplasms'. A number of sites form what might be called a central city pattern, displaying high SMR's in the centrally located tracts and low SMR's in the outlying areas. Cervical cancer (Figure 1) is the best example of this type of pattern; others are stomach, liver, lung, and 'all malignant neoplasms'. This particular type of geographic variation is quite probably the result of the negative relationship between socioeconomic status and certain cancers since central city residents are of lower socioeconomic status (as indicated by median incomes). A marked association exists between low socioeconomic status and cancer of the cervix, stomach, lung (Levin, et al., 1974), and liver (Hoover et al., 1975). Two cancers show an opposite type of pattern, that is low SMR's in the center of the city with higher SMR's in the outlying areas. Both of these cancers (large intestine, lymphosarcoma and reticulosarcoma) have shown a positive association between socioeconomic status and mortality (Hoover et al., 1975), which may be reflected in the Columbus distribution. No obvious city-wide patterns appear for cancers of the rectum, pancreas, breast, uterus, ovary, prostate, kidney, bladder, brain, or leukemia.






A second method of examining variation among the census tracts is to note those tracts that have SMR's significantly high or low for a number of cancer sites. Most Columbus tracts having high SMR's are similar in terms of socioeconomic status, percent foreign stock, and types of cancer for which mortality is elevated. Four of these tracts form a geographic block around the Scioto River in south central Columbus. Residents of these four low income inner city tracts are at high risk of mortality for cancers of the lung, cervix, liver, and 'all malignant neoplasms'.

CANCER MORTALITY BY CENSUS TRACT:
1956 - 1974

COLUMBUS, OHIO



Age, Sex and Race Adjusted SMR*

-  High $p \leq 0.05$
-  High $0.05 < p \leq 0.2$
-  No Difference
-  Low $0.05 < p \leq 0.2$
-  Low $p \leq 0.05$



*Standard Population is Columbus, Ohio

Figure 1

The most remarkable census tract in Columbus, in terms of cancer mortality, is tract 91. This tract is significantly high ($p < 0.05$) for seven individual sites and for "all malignant neoplasms". It is quite dissimilar from the other high mortality tracts, displaying a greater percentage of foreign stock, a much higher median income level, and a different group of cancer sites for which mortality is elevated (large intestine, breast, ovary, kidney, brain, leukemia). The tract has a sizable Jewish population, many of them foreign born or of foreign stock from the U.S.S.R., Germany, and Poland. People of foreign stock comprised 25.2 percent of the population in 1960 and 24.4 percent in 1970. These people may experience genetic and/or cultural influences that predispose them to increased risk of mortality from certain cancers.

This study offers one application of mapping techniques to the investigation of disease and illustrates large differences in risk between census tracts within an urban area. These differences may be due to differences in local exposure to carcinogenic influences or to differences in the social, cultural, and demographic risk profiles of the local populations. Describing the variation in cancer mortality within the population is the first step toward interpreting patterns, increasing current knowledge of the disease, and suggesting theories of etiology.

References

- Burbank, F., Epidemiology Branch, National Cancer Institute. "Patterns in Cancer Mortality in the United States: 1950-1967". National Cancer Institute Monograph 33: May 1971.
- Cohart, E. M. "Socioeconomic Distribution of Stomach Cancer in New Haven". Cancer 7:455-461, May 1954.
- Dunham, L. J. and Bailar, J. C. "World Maps of Cancer Mortality Rates and Frequency Ratios". Journal of the National Cancer Institute 41:155-203, July 1968.
- Greenburg, L., Field, F., Reed, J.I., and Glasser, M. "Air Pollution and Cancer Mortality". Archives of Environmental Health 15:356-361, September 1967.
- Henderson, B. E., Gordon, R. J., Menck, H., Soohoo, J., Martin, S. P., and Pike, M.C. "Lung Cancer and Air Pollution in Southcentral Los Angeles County". American Journal of Epidemiology 101:477-488, 1975.
- Hoover, R., Mason, T.J., McKay, F.W., and Fraumeni, J.F. "Geographic Patterns of Cancer Mortality in the United States". In Persons at High Risk of Cancer: An Approach to Cancer Etiology and Control. pp. 343-360. Edited by J. F. Fraumeni. New York: Academic Press, Inc., 1975.
- Levin, D. L., Devesa, S.S., Godwin, J.D., and Silverman, D.T. Cancer Rates and Risks. DHEW Publ. No. (NIH) 75-691. Washington, D.C.: U. S. Government Printing Office, 1974.
- Mason, T. J., McKay, F.W., Hoover, R., Blot, W. J., and Fraumeni, J.F. Atlas of Cancer Mortality for U. S. Counties: 1950-1969. DHEW Publ. No. (NIH) 75-780. Washington, D.C.: U. S. Government Printing Office, 1975.
- Pyle, G. F. Heart Disease, Cancer and Stroke in Chicago: A Geographical Analysis with Facilities, Plans for 1980. Chicago: The Department of Geography, The University of Chicago, 1971.