THE GEOGRAPHY OF HEALTH AND DEATH AND THE QUALITY OF THE ENVIRONMENT IN MANHATTAN, KANSAS, 1965-1969

by 6791

KEITH PIERCY

B.A., Sheffield University, 1969

A MASTER'S THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER OF ARTS

Department of Geography

KANSAS STATE UNIVERSITY
Manhattan, Kansas

May 1971

Approved by:

David E Kromm

Major Professor

LD
2668
T4
1971
P54
C.2

ACKNOWLEDGEMENTS

The author wishes to thank Dr. David E. Kromm for all his help and guidance, which made the successful completion of this thesis possible. He would also like to thank Dr. H. L. Seyler for his assistance, especially with the analysis.

TABLE OF CONTENTS

CHAPTER

LIST OF TABLES

LIST OF MAPS

CHAPTER I

INTRODUCTION

As awareness of city problems has grown, geographers have become increasingly active in researching urban questions. The quality of urban life is an especially important question of interest to geographers. This thesis looks at one specific urban social problem, that of the nature and adequacy of the provision of health services as related to variations in the need for medical care. It will specifically investigate the variation in the need for health care in one urban area, the city of Manhattan, Kansas.[1]

I. BACKGROUND TO THE PROBLEM

Urban problems in geography. There has been much fruitful research by geographers into the problems of the urban environment, especially since 1945. This research has concentrated on urban problems in two ways. The first, more economically oriented approach, has been to study the pattern of the use of urban space, emphasizing the intraurban economic and

---

[1] This thesis was originally motivated by the personal observation of the writer, in the summer of 1970, that throughout the city of Detroit there were considerable spatial variations in the death rate from each of a number of different diseases. In Detroit this spatial variation in death was personally and qualitatively judged to be related to other social, economic, and demographic variables, which were in turn all spatially related to the quality of the urban environment.

social locational forces and the problems consequent upon the development of sub-optimal land use patterns.[2] Considerable emphasis has also been placed on the closely related topic of the forces and patterns of urban growth.[3] The second approach has been to consider the spatial aspects of specific urban issues, such as race,[4] air pollution,[5] and poverty.[6] Health has been given little attention as part of the geographic discussion of these individual urban problems. But undoubtedly there are spatial variations in the incidence of disease, health, and medical care facilities within urban areas. In this thesis health problems are investigated as a part of the geographical field of study of urban problems; specifically the spatial variation in the need for health services will be studied as to how it is related to variations in the quality of the urban environment.

---

[2] A relatively recent review of intraurban location studies in geography is given in B. J. Garner, "Models of Urban Geography and Settlement Location," in R. J. Chorley and P. Haggett (eds), Models in Geography (London: Methuen, 1967), pp. 335-355.

[3] Geographic research into this problem is summarized in H. M. Mayer, The Spatial Expression of Urban Growth, Association of American Geographers, Commission on College Geography Resource Paper No. 7, (Washington, 1969).

[4] For a review of geographic studies of urban racial issues see H. M. Rose, Social Processes in the City: Race and Urban Residential Choice, Association of American Geographers, Commission on College Geography Resource Paper No. 6, (Washington, 1969).

[5] A summary of geographic research into this problem is given in R. A. Bryson and J. E. Kutzbach, Air Pollution, Association of American Geographers, Commission on College Geography Resource Paper No. 2, (Washington, 1968).

[6] For a discussion of geographic research into poverty see "The Geography of American Poverty," Antipode, vol. 2 (2), December 1970. This is an entire issue devoted to the subject of American poverty, and of perhaps the greatest use is the comprehensive bibliography, pp. 84-106.

The most important previous contribution to this question on the part of geographers is to be found in the Chicago Regional Hospital Study.

The Chicago Regional Hospital Study. The most comprehensive research by American geographers into urban health problems has been on behalf of the Chicago Regional Hospital Study. This survey was supported by a grant from the Office of Economic Opportunity to the Chicago Board of Health to investigate the problem of providing health care to indigent poverty populations. The investigation concentrated on the present state of provision of health care throughout the city, with an emphasis on how the pattern of health services could be improved to give better care to the less privileged social groups of Chicago.[7] The geographic contribution to the project has been largely based on the premise that "hospitals, like schools and indeed most retail businesses, are central services."[8] Because the geographic approach to the Chicago Regional Hospital Study is derived from central place theory, the work has consisted partially of building models of optimum hospital size and location,[9] and identification of the hierarchy of hospital services and classification of hospitals.[10] Study of

---

[7] M. H. Lepper et at., "Approaches to Meeting the Needs of Large Poverty Populations," American Journal of Public Health, 57: 1153-1157, 1967.

[8] R. L. Morrill and P. Kelley, "Optimum Allocation of Services: An Hospital Example," The Annals of Regional Science, 3: 55, 1969.

[9] Ibid., pp. 55-66.

[10] R. L. Morrill and R. J. Earickson, "Variation in the Character and use of Chicago Area Hospitals," Health Services Research, 3: 225-235, 1968.

patient flows and the distance traveled by patients to hospitals has also been a major emphasis of the geographical contribution; simulation and gravity models have been used to investigate the effects of such variables as the level of services offered by the hospital, the type of care needed, income, race, and religion on the decision of the patient as to which medical facility to use.[11] The results of the studies of patient distance traveled are used to suggest locational shifts of hospitals and/or changes in capacity to meet the needs of both hospitals and patients, with an emphasis on the need to overcome barriers of race and income.[12]

The work of the geographers on the Chicago Regional Hospital Study is a valuable contribution to the spatial study of urban health problems, but it contains certain weakensses. The papers dealing with minimum distance traveled implicitly assume that use of health services will be related to the need for medical care. But Muller suggests that, although they need more care, the propensity of low income people to use health services is

---

[11] R. L. Morrill and R. J. Earickson, "Hospital Variation and Patient Travel Distances," Inquiry, 5: 1-9, December 1968; Morrill and Earickson, "Problems in Modeling Interaction: The Case of Hospital Care," in K. R. Cox and R. G. Golledge (eds), Behavioral Problems in Geography: A Symposium (Northwestern University, 1969), pp. 254-276; Morrill, Earickson and P. Rees, "Factors Influencing Distances traveled to Hospitals," Economic Geography, 46: 161-171, April 1970; and Morrill and Earickson, "Effects of Race, Income, and Religion on Hospital Travel," in Pierre de Vise et al., Slum Medicine: Chicago's Apartheid Health System Report No. 6 of the Interuniversity Scoial Research Committee -- Metropolitan Chicago Area, (Chicago: University of Chicago, 1969)

[12] Morrill and Earickson, "Locational Efficiency of Chicago Hospitals: An Experimental Model," Health Services Research, 4: 128-142, Summer, 1969; and Morrill, Earickson and Rees, op. cit.

less than that of higher income persons.[13] Hochstim and others also found an inverse relationship between the need for and use of health services in the low quality environment areas of Oakland.[14]

There is consideration in depth of only one level of health care in the Chicago study -- the hospital. Secondary importance is given to the analysis of the role of the physician,[15] and no attention at all is paid to the provision of care in the home or to the problem of community health as a whole. Only in the excellent report of Pierre de Vise is there an attempt to define the areas where need of medical services is greatest and to relate the distribution of health services to this need. This study define the spatial variations in the incidence of disease and death as they are related to the poverty areas of the city, assesses the demand of the poor population for health care, analyzes the medical services available to meet that demand, and finally suggest how adequate health services could be provided. Because of its consideration of the linkage between health care and poverty, it would seem that future research on the part of geographers into the issues of health services could well benefit from following a format similar to the de Vise report.[16]

Outside of the Chicago survey, geographers seem to have contributed

---

[13] C. Muller, "Income and the Receipt of Medical Care," American Journal of Public Health, 55: 510-521, April 1965.

[14] J. R. Hochstim, D.A. Athanasopoulos and J. H. Larkins, "Poverty Area under the Microscope," American Journal of Public Health, 58: 1815-1827, 1968.

[15] The physician is considered only in as far as he affects the patients decision as to which hospital to use.

[16] Pierre de Vise et al., op. cit.

little to the study of urban health problems. But a number of other investigations have been carried out which concerned themselves with aspects of urban health problems and techniques that are of considerable interest to geographers. Included in these are the previously mentioned investigation of the Oakland poverty and non-poverty area,[17] the study by Abrams and others which includes discussion of the spatial aspects of drug addiction in Chicago,[18] and the application of computer mapping to health research in Boston.[19] The geographic research into health, disease, and death in the city has probably been restricted by the lack of interest in urban health problems thus far shown by medical geographers.

The contribution of medical geography. The most rigorous statements of the scope and purpose of medical geography are by Dr. J. M. May of the American Geographical Society. As an example, he suggests "Medical geography professes to make its principle objective the study of the areal distribution of disease and its relationship to the environment."[20] This definition incorporates the two main themes apparent in contemporary medical geography. First, it has essentially adopted a descriptive approach, in that much of the work has concentrated on the mapping of variations in

---

[17] J. R. Hochstim et al., op. cit.

[18] A. Abrams, J. H. Gagnon and J. J. Levin, "Psychosocial Aspects of Addiction," American Journal of Public Health, 58: 2142-2155, 1968.

[19] R. A. Greenes and V. W. Sidel, "The Uses of Computer Mapping in Health Research," Health Services Research, 2: 243-258, 1968.

[20] J. M. May, "Medical Geography," in P. E. James and C. Jones, American Geography: Inventory and Prospect (Syracuse University Press, 1954) p. 453.

the incidence of disease and death at a national or even a world scale.[21] In the present state of advance of medical geography this is an unavoidable circumstance, because knowledge of the geographical aspects of health, disease, and death is at present very limited, and a sound data base must still be completed. Nevertheless, McGlashan already looks forward to the time when, through development of appropriate quantitative techniques, geographic theories of the causation of disease can be constructed.[22]

The second major theme is that of the relationship of disease to the environment, and this would appear to be a much more controversial aspect of medical geography. May stresses that "the social as well as the physical environment is taken into account,"[23] and he obviously sees the man-environment relationship in the context of disease as an interrelated physical-cultural-biotic complex. But he tends to stress the influence of the physical environment at the expense of other elements, occasionally coming near to defining a simple cause and effect relationship between the physical, especially the climatic, environment and man as the explanation of health and disease in a given area. May's overall concept of a

---

[21] Examples include: G. M. Howe, National Atlas of Disease Mortality in the United Kingdom (Royal Geographical Society, London: Nelson, 1963); M. A. Murry, "The Geography of Death in the United States and the United Kingdom," Annals, Association of American Geographers, 57: 301-314, 1967; A. T. A. Learmonth, "Medical Geography in India and Pakistan," Geographical Journal, 127: 10-20, 1961; and J. M. May, The Ecology of Human Disease (New York: MD Publications, 1958).

[22] N. D. McGlashan, "The Nature of Medical Geography," Pacific Viewpoint, 10: 60-65, September 1969.

[23] J. M. May, The Ecology of Human Disease, p. 1.

symbiosis between man and his environment is sound, however. In this thesis the environment will be regarded in this symbiotic context, in that it will be investigated as a major factor in explaining the spatial variation of disease and death.[24]

The tendency of medical geography not to wholly disassociate itself with determinism can only be seen as harmful to its cause in a discipline still aware of its struggle to refute this philosophy. But the influence of climatic determinists such as Ellsworth Huntington, S. F. Markham, and Clarence Mills is still seen in recent medical geography.[25] In the investigation of urban health problems a tendency to stress the role of the physical environment may on occasion be appropriate, in that the effects of such elements as air pollution on health may vary spatially throughout a given city.[26] But many of the physical environmental influences are relatively uniform in their impression on the small areal extent of the urban area, and the effects of many of the others may be controlled by

---

[24] Ibid., pp. 1-34.

[25] As late as 1964 Stamp was prepared to discuss the ideas of Huntington, Markham, and Mills at some length. This generally uncritical use of determinist ideas should not be allowed to detract from the impact of the rest of the book, or to suggest that all his comments about the influence of environment on health are invalid, despite the tendency in part to imply a cause and effect relationship between the two. L. D. Stamp, The Geography of Life and Death (Cornell University Press, 1964), pp. 70-79 and 129-43.

[26] For a further discussion of the effects of air pollution on health see "Air Pollution and Health: A Statement by the Committee on Air Pollution, American Thoracic Society," American Review of Respiratory Disease, 93: 2-12, 1966.

human device, as with floods.

In the city it is the social environment that dominates human activity. In studying the role of the environment in explaining the spatial variation of health and death in an urban area it is usually more appropriate to concentrate on the social aspects of the environment. Terris comments on the importance to urban health research of the social environment, even going so far as to suggest that the main influences of the physical environment are socially caused:

> The individual, however, lives in a social as well as a physical environment. It is society which has created the hazards with which we contend in the physical environment... And it is society, the social environment, which largely determines individual attitudes and behavior.[27]

If Terris' social environment and the cultural environment of May are equated, then undoubtedly the attempt to explain spatial patterns of health, disease, and death in an environmental context for urban areas is justified. Morris states the case:

> And today the study of urban problems and poverty has focused attention on the total urban environment as the medical milieu. This approach involves more than the familiar demographic descriptions of a population, such as age, sex, and ethnic and economic class. To these must be added at least one further dimension: the symbiosis of the individual and his environment.[28]

It is hoped, therefore, to place the study of urban health problems in an environmental context, with a view towards synthesizing the diverse

---

[27] M. Terris, "A Social Policy for Health," American Journal of Public Health, 58: 5, January 1968.

[28] R. Morris, "The City of the Future and Planning for Health," American Journal of Public Health, 58: 13, January 1968.

approaches used by the geographers of the Chicago Regional Hospital Study and by many medical geographers. The spatial distribution of deaths within the city of Manhattan, Kansas, is related to selected variations in the quality of the environment of the city. This is seen as a basis for the recognition of variation within Manhattan of the need for health care. From this position it is possible to consider further study, beyond the aims of this thesis, on the adequate provision spatially of medical services for all of the population of the city.

II. STATEMENT OF THE PROBLEM

This thesis attempts to describe and explain the spatial expression of death in Manhattan, Kansas. The approach used is based on the hypothesis that the variation in the incidence of death will be inversely related to the quality of the urban environment. The course of investigation is as follows:

1. the collection and mapping of the data showing the distribution of death for residents of the city for the period 1965-1969;

2. the identification of selected variables measuring spatial variations in the quality of the environment throughout the city;

3. the relating of variations in the quality of the environment to the incidence of death.

III. THE DISTRIBUTION OF DEATH

The need to use mortality data. Mortality data was considered in preference to that of morbidity because of the problems of obtaining

morbidity data for Manhattan.[29] The Riley County Health Department does not make records available to the public to indicate the location or the extent of the occurrences of those major diseases that are notifiable.[30] The substitution of mortality data for that of morbidity in the analysis of health problems is an accepted technique that has been commonly used because of the lack of good statistics pertaining to health and diseases. Mortality statistics are obviously biased towards diseases which cause death, and will not be a good indicator of those that rarely, if ever, have fatal consequences. But they have been regarded as reliable indicators of the overall pattern of health in an area. In the past they have formed the major source for American health data.[31] The U. S. National Health Survey was started in 1957 to try to improve the quality of health statistics, but the small size of the sample used in gathering the data, the basing of statistics on illnesses known by the respondent to the questionnaire, and the lack of comparability with earlier statistics have all combined to reduce the utility of this source.[32]

---

[29] Mortality data is that which refers to the incidence of death and is usually derived from death certificates. Morbidity data is that which refers to the incidence of disease, which may or may not result in death, and which may be compiled from a number of sources.

[30] Oblinger-Smith, Neighborhood Analysis of Manhattan, Kansas (Wichita, 1968), p. 30; and personal interview at Riley County Health Department, September 17, 1970.

[31] M. Lerner and O. W. Anderson, Health Progress in the United States (Chicago: University of Chicago Press, 1963), pp. 1-96.

[32] T. R. Ford, Health and Demography in Kentucky (University of Kentucky Press, 1964), pp. 102-104; and Lerner and Anderson, op. cit., pp. 1-2.

The source of mortality data for Manhattan. The Registration and Health Statistics Services Division of the Kansas State Department of Health provided the certificates for all deaths occurring within the state for the period 1965-1969. From the death certificate it was possible to identify the place of residence of the deceased, but unfortunately the deaths were recorded as to where they occurred rather than on the basis of residence. This meant it was not possible to find certificates of the deaths of Manhattan residents who died outside Kansas without reviewing the records of other states. But it was felt these were not such a large part of the whole that their omission would prejudice the results of the study. The deaths of non-residents within the city were disregarded; this included most of the students and Fort Riley personnel who died during this period, although a few students and soldiers were recorded as residents and thus included in the population studied.

IV. DEATH AND THE QUALITY OF THE URBAN ENVIRONMENT

The relationship between death and the environment. The environment has been defined as "each person's room for living,"[33] suggesting that an individual's health is related to the physical and social characteristics of the space he or she lives in. This definition is suited to the ideas expressed previously, where it was suggested the environment of the city dweller was dominated by social and cultural elements.[34] This is

---

[33] Health is a Community Affair Report of the National Commission on Community Health Services, (Cambridge, Massachusetts: Harvard University Press, 1966), p. 38.

[34] See above, pp. 7-10.

particularly applicable to Manhattan, where there is little or no evidence of any significant variation about the city in the quality of the physical environment. Thus this thesis, because it concerns itself with an urban area, is especially interested in the "social room for living" of the individual.

The relationship between the quality of a person's environment and state of health is frequently stressed. According to the National Commission on Community Health Services, "Whatever a man's genetic inheritance, much of his health is dependent on the quality of his relation with his environment."[35] Benjamin suggests that "the risk of dying varies with a number of factors -- sex and age and those factors which either influence the physicial condition or the environment of the people..."[36] May argues that disease is synonymous with maladjustment of the human being in the man-environment symbiosis:

> Disease is very simply that alteration of living tissues that jeopardizes their survival in their environment.... It expresses a temporary state of the living cell in conflict with environmental challenges and trying to cope with them and survive.[37]

Thus it would be expected that the incidence of death varies as the environmental factors change spatially throughout the city of Manhattan.

---

[35] Report of the National Commission on Community Health Services, op. cit., p. 7.

[36] B. Benjamin, Health and Vital Statistics (London: George Allen and Unwin, 1968), p. 81.

[37] J. M. May, The Ecology of Human Disease p. 1.

Analysis of the death-environment relationship. The hypothesis of an inverse relationship between environmental quality and the incidence of death will be tested through the use of a multiple regression analysis relating death to nine variables based on population overcrowding, value of housing, non-white population, the marital status of the population, and the availability of plumbing facilities. Data on these variables was obtained from the first count tapes of the 1970 Census of Population as provided by the Institute of Social and Environmental Studies at the University of Kansas. But before proceeding with the analysis it is necessary to investigate the distribution of death around the city of Manhattan, and it is to this that the thesis turns in Chapter II.

CHAPTER II

THE DISTRIBUTION OF DEATH IN MANHATTAN, 1965-1969

It is the purpose of this chapter to describe the occurrence and spatial variation of death in Manhattan for the period 1965-1969. To place this consideration in perspective it will be necessary to discuss the advantages and disadvantages of the death certificate as a source of mortality and morbidity data. The distribution of death will then be described, and this will form the basis for the later analysis of the relationship between death and the quality of the environment.

I. PROBLEMS ASSOCIATED WITH THE CERTIFICATE OF DEATH

The information available from the certificate of death. The death certificate used by the Kansas State Department of Health classes information into five categories: data about the deceased, description of the parents of the deceased, identification of the cause of death, information about the medical practitioner who certified death, and specification of the nature and place of the burial or cremation. The writer was most interested in the data about the deceased and the identification of the cause of death.

The contents of the category identifying and describing the demographic characteristics of the dead person includes the following information:

1. the date of death of the deceased;
2. the place of death of the deceased;

3. the sex, race, age, marital status, place of birth, occupation, and residence of the deceased.

The most valuable part of this information was that relating to the residence of the deceased, which allowed the precise location of each residence, which in turn facilitated the eventual mapping of the distribution of mortality. The exclusion of non-residents who died in Manhattan from the research was also made possible. This restricted the studied population to that of residents of the city, a necessary step in view of the fact that non-residents would not be likely to have had much interaction with the environment of Manhattan.

The category dealing with the cause of death is surrounded by considerable controversy, and constitutes the most critical part of the death certificate. In part I of this category of the certificate the first line is reserved for the immediate cause of death, and the following lines are then used for description of the conditions, if any, which gave rise to the immediate cause, stating the underlying cause last. Part II is allocated to other significant conditions prevalent at and contributing to death, but not related to the immediate cause. From this method of classification of cause of death problems arise, related to the variation among physicians in their methods of completing this part of the certificate.

The certification of cause of death by the medical practitioner. The highly subjective element in the manner in which an individual medical practitioner must diagnose such causes is responsible for variation among physicians in their final descriptions of cause of death. Benjamin summarizes this problem:

> It must be borne in mind that though medical students are given training in the practice and purpose of death certification, and though the majority of medical practitioners take considerable interest in maintaining a high standard of reporting, there are sources of inaccuracy in the certification of causes of death and therefore in the statistics based on these certificates; and it implies no criticism of medical practitioners to say that statisticians sustain a healthy reluctance to regard the recorded causes as precise statements. Essentially they are statements of medical opinion however well informed and even when based on post-mortem examination.[1]

The major sources of inaccuracy may be related to the factors of old age, training of physicians, and human fallibility. The conditions concurrent with old age may so obscure the true succession of events leading to death as to make it difficult to discern the underlying cause. The variation in the training of physicians can be held as partly responsible for different diagnostic practices which will result in inconsistencies in certification, despite attempts to provide uniformity in the methods and language of the classification of death.[2] Today the physician is expected to privde sufficient detail to allow the underlying cause of death to be assigned to its proper category in the International Statistical Classification of Diseases, Injuries, and Causes of Death.[3] The element of human fallibility is present in that the possibility of an error in diagnosis on the part of the medical practitioner is always present. Thus medical personnel can, in these several ways, affect the quality of mortality

---

[1] B. Benjamin, Health and Vital Statistics (London: George Allen and Unwin, 1968), p. 81.

[2] Ibid., pp. 81-91.

[3] Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death, 1965 Revision (Geneva: World Health Organization, 1967) 2 volumes.

reporting.

Despite these problems mortality data can be used to give a satisfactory measure of the health of a given population, even with the probability of a small degree of inconsistency. Thus Benjamin is still readily prepared to use death certificates as a source of health data.[4] A more enthusiastic endorsement of mortality statistics as measures of morbidity is that of Lerner and Anderson:

> Records of death have to be kept in any well-ordered society, and therefore routine registration within official agency has long been common practice in this country. The by-product, a series of national mortality statistics, has proved immensely valuable in measuring health levels and in delineating long-term trends in the level of health. Death is a clearly defined event, so that mortality statistics are precise; and registration in this country is now fairly complete, so that the derived statistics are believed to be reasonably accurate, at least as they relate to the demographic characteristics of the deceased.[5]

Thus, with an awareness of the existence of some limitations, but also a belief in the reasonable accuracy of the mortality data, the task of deriving and describing the incidence of death in Manhattan was initiated.

II. MORTALITY BY CAUSE IN MANHATTAN, 1965-1969

Methodology of the derivation of the mortality data. From the death certificates for the state of Kansas during the period 1965-1969 it was possible to identify 769 deaths of people who were recorded as residents

---

[4] B. Benjamin, op. cit., pp. 73-91.

[5] M. Lerner and O. W. Anderson, Health Progress in the United States (Chicago: University of Chicago Press, 1963), p. 1.

of Manhattan. Of these, 38 were fatalities of residents of institutions,[6] with the remaining 731 being deaths of persons recorded as domiciled in the city at the time of their death.

The deaths were first classified according to cause. The procedure used was that outlined in the International Statistical Classification, which suggests that death be classified by the general rule that the condition on the lowest used line of part I of the cause of death section be used. Thus in a simple example used for Manhattan the cause of death was described in part I as:

> (a) metastatic carcinoma of the left hip due to (or as a consequence of)
> (b) carcinoma of the breast

The death was thus classified as being caused by carcinoma of the breast. In many cases the selection procedure was not as simple as this, and recourse had to be made to the other three selection and eight modification rules provided by the manual. These allow evaluation of causes of mortality in the event of the identification of one or more sequences of disease, the lack of any sequences, the reporting of senility, the recording of the underlying cause as being an ill-defined or a trivial condition, the use of specific terms, and the late effects of previous conditions.[7]

Having identified the probable cause of death for all of the cases

---

[6] The institutions concerned were: Parkview Manor, Wharton Manor and the County Memorial Hospital, and St. Mary's Hospital.

[7] The International Statistical Classification, op. cit., pp. 413-436 should be referred to for a full discussion of the selection procedure. The additional rules and the examples of each are too complex to be discussed in detail here.

it was desired to aggregate these causes into a classification involving a manageable number of categories. This was done by locating the code number assigned to each cause by the International Statistical Classification, and then grouping the causes into classes on the basis of the code number and the special tabulation lists provided by the manual.[8] The results of this classification process may be identified in Table I.

Comparison of the Manhattan mortality statistics with other similar data. In an effort to make Table I more meaningful the non-institutional deaths were compared to the mortality statistics for the U. S. A. in the year 1965, and to those of the city of Topeka, Kansas, for the same period of 1965-1969 (see Tables II and III). The comparison should be treated with some reservations because of the probable differences in the methodology of classification of death between the three places, especially in the subjective matter of the coding of mortality and selection of classes of causes.

The comparison with the U. S. A. and Topeka data shows similarity in the ranking of the first seven causes of death. In all three instances the leading cause of death, accounting for over a third of all fatalities, is heart disease. Malignant neoplasms (cancer) and cerebrovascular disease (brain diseases) follow in order. Thereafter there are some changes in the rankings, although the same four causes fill the next four positions. In Manhattan accidental deaths are less important than in the other cases,

---

[8] Ibid., pp. 41-411 and 437-449.

TABLE I. DEATHS OF ALL RESIDENTS BY CAUSE IN MANHATTAN, KANSAS, 1965-1969[1]

I.S.C. Code	Cause of Death	Deaths of Residents: Non-institution No.	Non-institution %	All No.	All %
392-429	Diseases of the Heart	273	37.34	282	36.67
140-209	Malignant Neoplasms	116	15.87	117	15.21
430-438	Cerebrovascular Disease	55	7.52	61	7.93
470-486	Pneumonia and Influenza	52	7.11	58	7.54
440	General Arteriosclerosis	36	4.92	40	5.20
760-779	Certain Causes of Death in Early Infancy	25	3.42	25	3.25
E800-E949	Accidental Deaths	24	3.28	28	3.64
442-458	Other Diseases of the Circulatory System	24	3.28	28	3.64
490-493	Bronchitis, Emphysema, and Asthma	19	2.60	19	2.47
531-533	Peptic Ulcer	8	1.09	9	1.17
580-584	Nephritis and Nephrosis	8	1.09	9	1.17
E950-E959	Suicide	8	1.09	8	1.04
250	Diabetes Mellitus	7	0.95	7	0.91
320-358	Diseases of the Nervous System	7	0.95	7	0.91
441	Aortic Aneurysms	7	0.95	7	0.91
510-519	Other Diseases of the Respiratory System	7	0.95	7	0.91
571	Cirrhosis of the Liver	7	0.95	7	0.91
590-629	Other Diseases of the Genito-Urinary System	6	0.82	6	0.78
E960-E969	Homicide	4	0.55	4	0.52
	All Other Causes[2]	38	5.20	40	5.20
	Total	731	100.00	769	100.00

[1]Derived from the Kansas State Death Certificates and first summary tapes of the 1970 Census of Population.

[2]Includes: nutritional deficiency 3; congenital anomalies 3; diseases of the blood 3; alcoholism 3; other diseases of the liver 3; enteritis 2; diseases of the musculo-skeletal system 2; diseases of the intestine 2; disease of the pancreas 1; benign neoplasm 1; non-psychotic mental disorder other than alcoholism 1; symptoms and ill-defined conditions 2; cause of death not specified 12.

TABLE II. LEADING CAUSES OF DEATH IN THE UNITED STATES, 1965[1]

Cause of Death	Number of Deaths	% of all Deaths
Diseases of the Heart	712,087	38.95
Malignant Neoplasms	297,588	16.28
Cerebrovascular Disease	201,057	11.00
Accidental Deaths	108,004	5.91
Influenza and Pneumonia	61,903	3.39
Certain Diseases of Early Infancy	55,398	3.03
General Arteriosclerosis	38,102	2.08
Diabetes Mellitus	33,174	1.81
Other Diseases of the Circulatory System	27,279	1.49
Other Bronchopulmonic Disease	26,518	1.45
Cirrhosis of the Liver	24,715	1.35
Suicide	21,507	1.18
Congenital Malformations	19,512	1.06
All Other Causes	201,286	11.02
Total	1,828,136	100.00

[1]Derived from Lowell, Edwards and Palmer, Tuberculosis (Cambridge, Massachusetts: Harvard University Press, 1969), p. 71.

TABLE III. LEADING CAUSES OF DEATH IN TOPEKA, KANSAS, 1965-1969[1]

Cause of Death	Number of Deaths	% of all Deaths
Heart Disease	2083	36.82
Malignant Neoplasms	936	16.54
Cerebrovascular Disease	729	12.88
Accidental Deaths	311	5.85
Pneumonia and Influenza	185	3.27
General Arteriosclerosis	154	2.72
Certain Causes of Motrality in Early Infancy	133	2.35
Suicide	86	1.52
Diabetes Mellitus	79	1.40
Cirrhosis of the Liver	74	1.31
Homicide	34	0.60
Nephritis and Nephrosis	33	0.58
All Other Causes	801	14.16
Total	5658	100.00

[1]Derived from correspondence with the Research and Analysis Section, Registration and Health Statistics Services Division, Kansas State Department of Health, September 14, 1970.

possibly reflecting the lack of manufacturing industry and lower incidence of vehicular traffic in the city. Pneumonia and influenza, general arteriosclerosis, and certain causes of death in early infancy make up the rest of the leading seven causes in each case. Below the first seven positions the correspondence in terms of ranks becomes erratic, probably because of the low frequencies of these types of fatality in Manhattan.

To see if this apparent correspondence of the other two distributions with the Manhattan death statistics was valid statistically, two Kolmogorov-Smirnov tests were carried out between the Manhattan and the U. S. A. and the Manhattan and the Topeka distributions. The Kolmogorov-Smirnov test can be used for two sets of data which have been grouped into three or more categories. It compares the cumulative frequency distributions of the two data arrays, the test statistic D being the maximum difference between the two distributions. The nul hypothesis was that there was no significant difference between the distributions. In each case the Kolmogorov-Smirnov statistic proved to be not significant at the .01 level, thus allowing the null hypothesis to be accepted. Therefore it was concluded, on the basis of the qualitative and the quantitative information, that the Manhattan distribution of mortality by cause was not significantly different from the others, and that there were no disparities of any considerable magnitude. With this in mind the intra-urban pattern of death in Manhattan was investigated.

III. THE SPATIAL VARIATION OF MORTALITY IN MANHATTAN, 1965-1969

The mapping of the distribution of death. As stated previously, the certificate of death gives the address of the person who died.[9] The location of the home of each of the 731 residents who died in the period in question could thus be established. This was done by use of a lot map purchased from the City Engineer's Office in Manhattan, and with the aid of the City Directory, which listed all the inhabitants residences by street for 1969.[10] In order to represent the information about residence of the deceased it was decided to plot the occurrence of deaths by the 1970 Census of Population enumeration districts for the city.

The use of the enumeration districts had several advantages. There are thirty-two of them inside the city limits, and this was envisaged as a sufficient number to show spatial variation of mortality within the city without breaking down the city into units that were too small to contain a large enough number of deaths to be meaningful for analysis. The use of 1970 enumeration districts was also desirable because the analysis of the relationship between environment and mortality relies heavily on the 1970 Census of Population data which is made available in a form based on these districts.

With the aid of a map of the enumeration districts purchased from the

---

[9] See above, pp. 15-16.

[10] Manhattan City Directory 1969 (Kansas City: R. L. Polk, 1969).

Bureau of Census the homes of the deceased residents were assigned to their correct unit. This allowed the number of deaths to be calculated for each area, and from this the average annual death rate per 1000 people for the five year period for each enumeration district was computed. The results of these calculations are shown in Map I and Table IV.

The spatial variation of mortality in Manhattan, 1965-1969. Map I shows the variation of mortality by enumeration district in average annual rates for the period in question. On this map a distinctive and very interesting pattern of death emerges. Around the central business district in the southeast of the city are to be found the heaviest rates of above 12.00 per 1000 persons per annum.[11] There is a decline in intensity from this part of the city in a concentric fashion, until the lowest rates of less than 3.00 are found in the northwest, west, and north of the city. The concentric decline of the mortality rate outward from the city center is remarkably uniform, with very few interruptions to the pattern. This decline of death rates with distance from the city center is very interesting in view of the similar decline in many demographic, economic, and social elements, such as population density and income. But before proceeding to the analysis of how this distribution is related to environmental quality in particular, it remains to be seen whether individual causes of death vary spatially about the city in a like manner.

The spatial variation of death by cause. Table IV gives a breakdown

---

[11]In the rest of this chapter all rates of mortality will be measured in the same units of a rate per 1000 persons per annum.

THIS BOOK CONTAINS NUMEROUS PAGES WITH DIAGRAMS THAT ARE CROOKED COMPARED TO THE REST OF THE INFORMATION ON THE PAGE.

THIS IS AS RECEIVED FROM CUSTOMER.

MAP 1

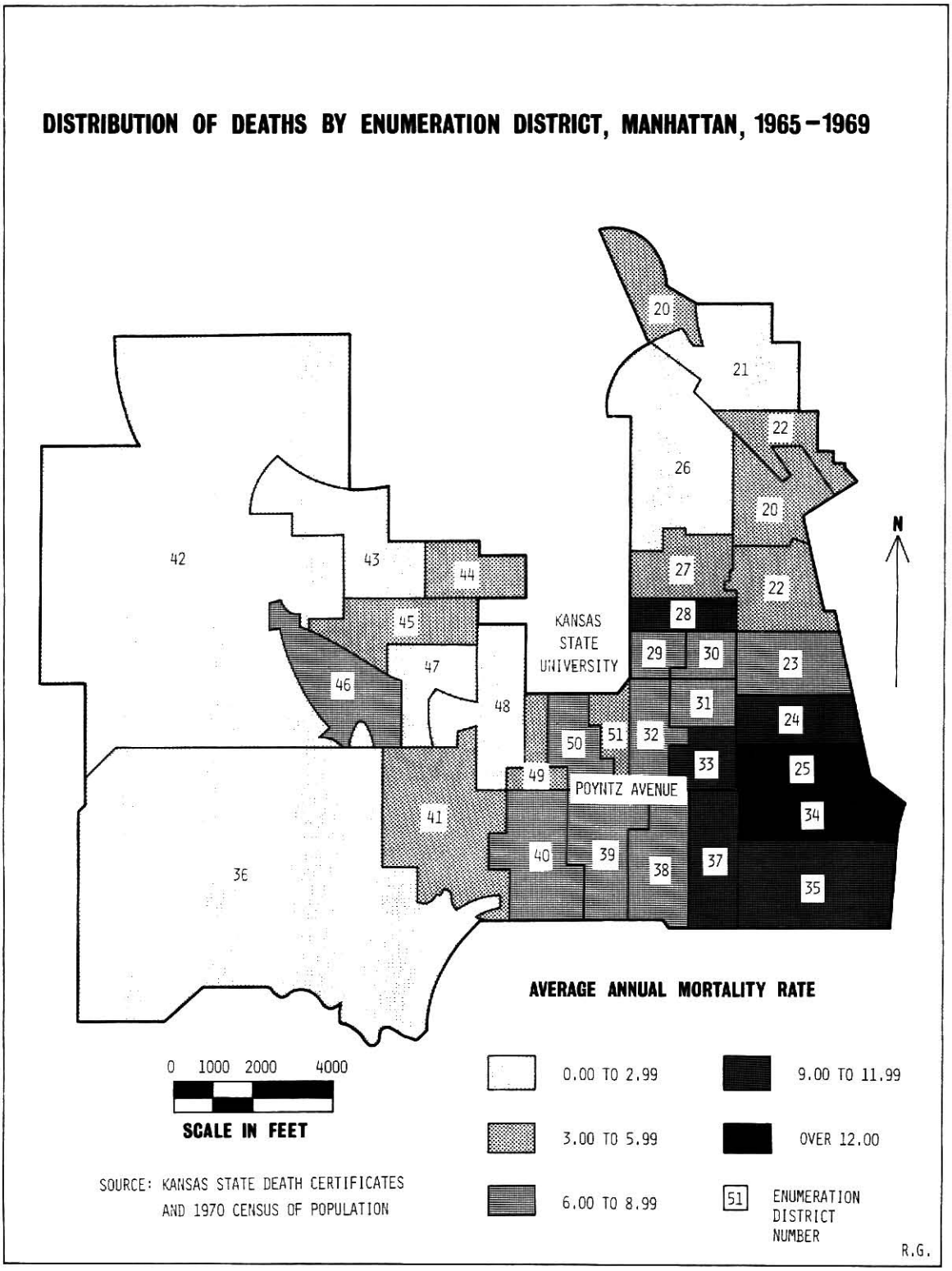

DISTRIBUTION OF DEATHS BY ENUMERATION DISTRICT, MANHATTAN, 1965–1969
20
21
22
26
20
42
43
44
27
22
45
KANSAS
STATE
UNIVERSITY
28
29
30
23
47
46
48
31
24
50
51
32
33
25
49
POYNTZ AVENUE
34
41
40
39
38
37
36
35
N
AVERAGE ANNUAL MORTALITY RATE
0 1000 2000 4000
SCALE IN FEET
0.00 TO 2.99
9.00 TO 11.99
3.00 TO 5.99
OVER 12.00
6.00 TO 8.99
51 ENUMERATION DISTRICT NUMBER
SOURCE: KANSAS STATE DEATH CERTIFICATES AND 1970 CENSUS OF POPULATION
R.G.

TABLE IV. DEATHS BY CAUSE, BY ENUMERATION DISTRICT, MANHATTAN, 1965-1969

	Cause of Death																				
ED	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	Rate/1000
20	8	1		1			1														11	4.56
21	1	5		1		2	1														10	1.72
22	3	3	1	1	1	2	2		1			1			1					2	18	4.38
23	10	6	4	2	1	1	1	1	2					1						2	31	8.17
24	14	6	2	5	2	2		1	1											1	34	11.55
25	13	2	4	3	2		1	3		1						1		1			31	18.40
26	2	1																			3	0.54
27	6	6		1	1	4		1	1												20	3.11
28	12	6	1	4			2	2	1			2								3	33	10.64
29	13	4	2	1	3		1		2			1						1	1	1	30	8.44
30	6	2	1	1	1	1	1	1						1							15	6.25
31	14	4	2	2				2					2	1	1						28	7.06
32	5	1	1	2	2	2	1				1				1					1	17	8.95
33	14	9	2	8	2		3	2	1		1	1				2			1	3	49	10.93
34	10	7	3	1	2				1	1								1			29	14.21
35	8	3	3	2	1				2	2					1	1					23	10.93
36	3			1									1								5	2.06
37	16		3	3	3			2	1		2			2		1	3		1	2	39	11.52
38	9	7	2	1	2	2	2		1				2	1				1	1	3	34	8.56
39	12	4	3		1		1			1	1		1			1				1	26	6.25
40	11	2	1	3		2	2	1	1	1	1	2						1		1	29	8.88
41	5		5			1		3	1		1				1		1			2	20	4.36
42	9	6	3	2		3	3	1									1				28	1.64
43	1		1		1	2			1											1	7	1.70
44	4	4	2	1							1									2	14	3.14
45	9	1			1			1							1		1				14	5.02
46	10	6	2															1		1	20	7.31
47	3	3		1						1										1	9	1.93
48	13	4		1	3		1										1			1	24	2.99
49	7	6	1	1		1	1													3	20	4.25
50	15	4	4	3	6	1	1	2	1			1		1	1					2	42	9.22
51	7	3	2		1			1					1			1				2	18	5.37
Total:	273	116	55	52	36	25	24	24	19	8	8	8	7	7	7	7	7	6	4	38	731	

of the incidence of the different causes of death by enumeration district. The only cause of death represented in all districts is not surprisingly diseases of the heart, although the other six leading causes are all well dispersed. The pattern of mortality from diseases of the heart does not appear to differ to any great extent from the overall pattern of death. A similar conclusion can be reached with malignant neoplasms, to which no fatalities were attributed in four districts. It is surprising that this cause is absent from enumeration district 37,[12] which has a high incidence of death. Enumeration district 46 is also anomalous, having a much higher rate of deaths from malignant neoplasms than surrounding areas. Cerebrovascular disease and pneumonia also show a tendency to concentrate in the southeast of Manhattan, in a similar pattern to total deaths. Deaths due to infant mortality and accidents show little pattern in their distribution about the city. For the less important causes of death there are too few frequencies of their occurrence to really justify suggestions of spatial

---

Key to Table IV: ED -- enumeration district.

The causes of death are as follows: a -- diseases of the heart; b -- malignant neoplasms; c -- cerebrovascular disease; d -- pneumonia and influenza; e -- general arteriosclerosis; f -- certain causes of death in early infancy; g -- accidental deaths; h -- other diseases of the circulatory system; i -- bronchitis, emphysema, and asthma; j -- peptic ulcer; k -- nephritis and nephrosis; l -- suicide; m -- diabetes mellitus; n -- diseases of the nervous system; o -- aortic aneurysms; p -- other diseases of the respiratory system; q -- cirrchsis of the liver; r -- other diseases of the genito-urinary system; s -- homicide; t -- all other causes of death.

u is used to denote the total number of deaths in each enumeration district for the five year period.

Rate/1000 refers to the average annual death rate per 1000 persons for each enumeration district.

[12]The enumeration districts can be located by their number on Map I.

pattern.

All the leading four causes of death thus bear a considerable resemblance to the distribution shown in Map I for total deaths. Whether this is borne out statistically will be one of the questions discussed in chapter IV, which will analyze the relationship of total deaths and the four leading causes of death to the quality of the environment. But first the variables by which the quality of the urban environment was measured are identified in chapter III.

CHAPTER III

THE QUALITY OF THE ENVIRONMENT IN MANHATTAN

In this chapter the variation in the quality of the environment in Manhattan will be discussed. The first consideration will be previous evidence that suggested the existence of such a variation. Then the main emphasis of the chapter, the identification and measurement of variables indicating how the environment varies spatially, will be considered. The variables chosen will then be used in the analysis in the next chapter.

I. EVIDENCE FOR SPATIAL VARIATION IN THE QUALITY OF THE MANHATTAN ENVIRONMENT

The need to assess previous evidence. The writer found it necessary at this point in his research to consider whether there was justification for supposing variation in the quality of the Manhattan environment. It seemed unwise to proceed with an analysis which included variation in environmental quality as an important factor if this presumed factor did not exist. In a small city of the size of Manhattan it would not be as safe to assume environmental differences spatially as it would be for a large metropolis, such as New York, Detroit, or Chicago. Therefore, it seemed in order to substantiate a personal awareness of differences in environmental conditions about the city, before going on to define specific variables measuring those difference.

The Neighborhood Analysis of Manhattan. The most comprehensive recent study of the urban environment of Manhattan is the neighborhood analysis of Oblinger-Smith.[1] This report states its purpose clearly:

> A neighborhood analysis is a study within the community designed to identify the nature, extent, and cause of blight on an area by area basis... The purposes of a neighborhood analysis are all directed toward an understanding of the physical condition of each of the neighborhoods... and of the social and economic conditions which are characteristic of these neighborhoods.[2]

Therefore Manhattan was divided into ten neighborhoods by the analysis. For these neighborhoods the analysis found a considerable variation in the condition of housing, which was used as the main index of the physical quality of the environment. Map II shows the areas with the highest proportion of blight are generally in that part of the city within the city limits of 1945,[3] with poor housing conditions most noticeable in the "South Side" of the city to the south of Poyntz Avenue. The neighborhood analysis also found that there was an apparent high correlation of the socio-economic factors of overcrowding, large families, non-white population, welfare cases, old aged people, incidence of crime, and inadequacies in public services with the incidence of blight. Conversely, the north, northwest, and west of the city are marked by an absence of blight and much lower incidence

---

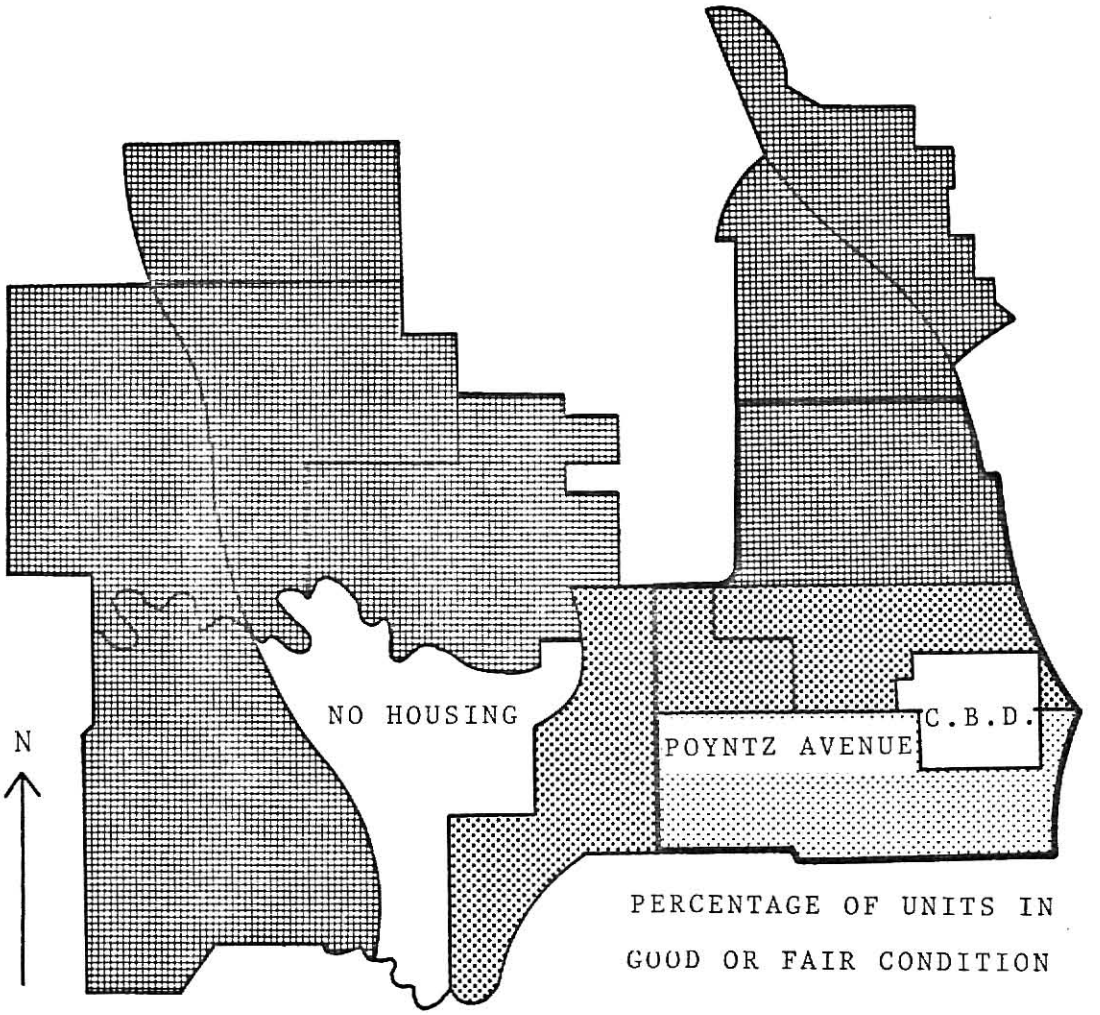
[1] Oblinger-Smith, Neighborhood Analysis of Manhattan, Kansas (Wichita, 1968).

[2] Ibid., pp. 1-2.

[3] Blight is defined here in the same terms as the analysis, in which blighted housing is that which is defined according to the U. S. Bureau of Census categories of deteriorating and dilapidated. Ibid., pp. 11-12.

MAP 2

CONDITION OF HOUSING BY NEIGHBORHOOD, MANHATTAN, 1965



SOURCE: OBLINGER-SMITH, pp. 33-53

R.G.

of adverse social and economic conditions.[4]

Thus Oblinger-Smith substantiates fairly comprehensively the notion of spatial differences in environmental conditions between the lower quality southeast and higher quality north and west of Manhattan. This reflects itself in the actions of the city government, for whom the neighborhood analysis was prepared.

Urban renewal. The government of the city has shown itself aware of the presence of spatial variation in the urban environment. The result of this awareness was the approval in January 1971 of an urban renewal project for 260 acres in the southeast of the city, including the central business district, for which the federal government gave a grant of $4,300,000. The city has also concerned itself with other schemes in the southeast of the city, including an attempt to provide some low cost housing for low income persons.[5]

It was thus decided the evidence for variation in the quality of the environment was substantial enough to warrant its investigation as a factor in the variation of death rates about the city. Therefore nine variables were selected from the 1970 Census of Population data to represent five elements thought of as indicative of environmental quality.

---

[4] Ibid., pp. 5-32.

[5] Manhattan Mercury, January 13 and 14, 1971.

II. THE CHOICE OF VARIABLES TO MEASURE ENVIRONMENTAL QUALITY

The nine variables selected were used to represent the following five elements: non-white population, marital status of the population, overcrowding in housing units, value of housing and the availability of plumbing facilities. Each of these elements will not be considered in turn.

Non-white population. The tendency for members of minority groups to be forced to live in a poor environment is a well-known feature of contemporary American (and Western) urban society. Many statements to this effect have been made; among those by geographers the most notable include those by Morrill and Rose.[6] The association of minority groups with the low quality environment suggested the use of the two variables of percentage non-white and percentage Negro as indicators of environmental quality. It was assumed that the higher the proportion of non-whites and Negroes per enumeration district the lower the environmental quality in that unit. The presence of a minority group population was also considered to be conducive to racial and social tension that would be likely to worsen the social environment in the area affected. Thus the percentage non-white variable was chosen. The percentage Negro variable was added because in

---

[6] R. L. Morrill, "The Negro Ghetto: Problems and Alternatives," Geographical Review, 55: 339-361, 1965; and H. M. Rose, "The Development of an Urban Subsystem: The Case of the Negro Ghetto," Annals, Association of American Geographers, 60: 1-17, 1970.

Manhattan non-white and Negro are not necessarily synonymous as is so often the case, the Blacks accounting for only 65% of the non-white population. Thus it was possible that the distribution of the Negro population was different to that of the non-white as a whole, and the Blacks were sufficient in number to have some impact on the social environment of their own accord.

The distribution of non-whites and Negroes about the city of Manhattan is summarized in Table V. With the use of Map I to locate enumeration districts, it can be seen that the non-white population as a whole is exceedingly concentrated in the southeast part of the city, and the Negro population is, if anything, even more concentrated here. This was the part of the city described by Oblinger-Smith as the most blighted and socio-economically the poorest environment in Manhattan.[7]

Marital status. The marital status of the population was used as a measurement of the quality of the social environment. This was based on the assumption that a departure from the norm of a family with a male head and wife with some or no children was likely to reflect family instability and thus a poor social environment as regards the home environment of the individuals involved in such a situation.[8] Two variables were thus selected to represent the non-normal marital condition. The percentage of the

---

[7] See above, pp. 32-34.

[8] The use of marital status in this context was suggested to the writer by G. M. Lewis in "Levels of Living in the Northeastern United States," Transactions, Institute of British Geographers, 45: 11-37, 1968.

TABLE V. DISTRIBUTION OF VARIABLES MEASURING ENVIRONMENTAL QUALITY IN MANHATTAN, 1970[1]

ED[2]	POP	NW	N	DS	FH	MP	OVER	VAL	R	PF
20	482	0.41	0.41	1.99	2.90	0.61	3.95	13338	73	0.00
21	1163	0.60	0.00	2.22	4.14	0.66	5.22	9882	92	0.00
22	821	0.97	0.00	2.39	6.28	0.64	5.22	16633	97	0.00
23	759	0.00	0.00	3.71	7.94	0.56	3.68	8917	79	1.25
24	589	2.38	2.04	5.66	9.81	0.53	3.28	9752	72	4.48
25	337	1.48	0.30	5.43	7.40	0.51	3.74	8221	70	3.00
26	1102	2.27	1.54	0.69	1.58	0.65	10.64	23939	119	0.35
27	1286	1.01	0.23	2.07	7.24	0.56	1.43	1550	118	0.46
28	620	1.94	0.16	3.02	9.94	0.50	3.44	11812	90	0.37
29	711	6.33	1.55	1.88	3.27	0.57	2.19	10396	87	17.63
30	480	5.21	1.46	4.93	12.31	0.52	4.50	10536	81	5.31
31	793	2.14	0.63	2.65	7.58	0.56	5.14	9373	79	4.46
32	380	5.53	1.05	3.50	6.67	0.56	5.44	8676	83	7.98
33	897	1.56	0.45	4.01	6.22	0.52	2.69	9639	66	8.01
34	394	3.30	2.54	8.97	15.38	0.59	9.27	12039	57	9.15
35	439	32.35	29.16	8.54	15.13	0.64	11.26	7628	55	9.15
36	486	8.85	7.41	3.06	3.68	0.57	5.44	28028	87	0.63
37	677	37.67	35.01	9.09	26.71	0.61	9.01	8770	59	6.17
38	794	21.03	19.14	6.84	15.38	0.64	9.52	7559	66	2.44
39	832	6.37	4.45	3.25	9.65	0.60	6.62	9804	82	3.40
40	653	1.68	1.23	4.27	9.64	0.54	4.44	15100	84	2.73
41	917	0.44	0.00	2.23	6.34	0.52	2.73	22925	100	0.00
42	3407	2.14	0.68	1.62	2.94	0.61	2.51	24236	143	0.09
43	824	1.21	0.00	2.46	9.18	0.64	4.05	18504	124	1.78
44	891	0.90	0.22	2.50	8.33	0.58	5.15	16583	117	2.55
45	558	1.97	0.36	2.11	8.66	0.51	2.34	19409	101	1.66
46	547	0.00	0.00	0.71	5.59	0.49	1.11	26537	117	1.11
47	931	1.61	0.00	0.86	5.61	0.52	2.90	24052	101	0.71
48	1604	0.81	0.44	0.75	4.14	0.55	3.89	22870	113	1.01
49	942	3.29	0.32	1.56	6.08	0.53	3.09	16919	96	9.74
50	911	2.85	0.22	1.65	6.25	0.53	5.06	14542	90	8.50
51	670	5.52	1.94	1.66	1.96	0.59	5.16	13542	75	19.03
City	26897	4.16	2.70	2.79	7.18	0.62	4.58	16580	94	3.54

[1]Derived from first summary tapes, 1970 Census of Population.

[2]For the location of enumeration districts see Map I.

population over fourteen divorced or separated was chosen because of the breaking up of a home and the resultant stress caused by this situation. The percentage of families with female heads was the other variable; the lack of a male head is often reflected in the reduced income of such a family with attendant adverse economic and social consequences.

Table V also shows the distribution of these two variables about Manhattan. Again they tend to concentrate in the southeast of the city, the concentration here of the families with female heads being particularly pronounced, as with the case of the non-white and Negro populations.

Overcrowding. Measures of overcrowding were selected because a crowded environment is assumed to be of lower quality. According to Senn, social tension and greater stress on physical facilities may result from an overcrowded environment. He suggests that the stress placed on the mental, behavioral, and social well-being of the individual by an overcrowded housing situation is more important today in considering health than the increased liability of transmission of communicable disease.[9]

Key to Table V: ED -- enumeration district; POP -- total population; NW -- percentage non-white; N -- percentage Negro; DS -- percentage of population over 14 who were divorced or separated; FH -- percentage of families with a female head; MP -- mean persons per room for occupied housing units; OVER -- percentage of occupied housing units with more than 1.00 persons per room; VAL -- mean value of owner-occupied housing units; R -- mean monthly rent of renter-occupied units; PF -- percentage of all housing units lacking one or more plumbing facilities.

[9]C. L. Senn, "Evaluation, Conclusions, and Recommendations," in Proceedings, First Invitational Conference on Health Research in Housing and its Environment, (Department of Health, Education, and Welfare: Washington, D. C., 1970).

Again two varialbes were selected to represent this element of the environment. The first was simply the mean number of persons per room for all occupied units. However, "overcrowding in housing is generally measured in terms of more than one person per room in a dwelling unit,"[10] and therefore this was also selected as a variable. From Table V it can be seen that the mean persons per room variable shows no readily observable pattern. This may possibly reflect the younger age of families in different parts of the city. Such families will have more children living at home and thus a greater number of persons per room. The younger age structure of the north, northwest, and west of the city in this case may be balanced by the presence of some student families in the south and east of the city, and also the possible overcrowding associated with student apartments in general. The percentage of units with more than 1.00 persons per room shows the same tendency as other variables to concentrate in the southeast of Manhattan. Enumeration district 26 emerges as an interesting anomaly, having a high score on this variable when its location in the city would not have led one to predict this. The high value is probably related to the large number ofapartments in this district.

Value of housing. This environmental element was used because the data as to housing condition as used by Oblinger-Smith was not available from the 1970 Census of Population. The value of the housing is assumed to generally reflect the physical condition of that housing. A high value unit is usually, in good condition and in a high quality environ-

---

[10]Oblinger-Smith, op. cit., p. 16.

mental location; the opposite is assumed to be true of low value housing. The value of the housing unit tends to reflect the income of the dwellers within, high income people occupying higher value housing and low income people lower value housing. This indication of income was useful because of the lack of income data in the first census tapes for 1970.[11]

The census data divided housing into the classes of owner-occupied and renter-occupied, with vacant units being categorized as for sale or for rent respectively. It was possible to calculate the mean value for owner-occupied housing (because the value of mobile homes was not included in the aggregate value sum, some interpolation had to be made if they were present),[12] and the mean monthly rent was calculated for rented units. The results appear in Table V. For both variables the higher values are generally found in the north, northwest, and west of the city, with the lower values in the south and east, a pattern suggesting a variation in environmental quality that is becoming familiar.

Plumbing facilities. The census data considers the availability of three types of plumbing facility -- a flush toilet, hot running water, and a bath or shower. The lack of any or all of these is likely to denote

---

[11] For further elaboration in considerable detail of the points made in this paragraph see R. F. Muth, "Urban Residential Land and Housing Markets," in H. S. Perloff and L. Wingo (eds), Issues in Urban Economics (Baltimore: Johns Hopkins, 1968), 285-333. The discussion by Downs that follows (pp. 423-426) is also useful.

[12] For the purpose of the calculation it was assumed that the proportion of owner-occupied mobile homes of the total number of mobile homes was the same as the proportion of all owner-occupied units of all housing units.

poor quality of the housing facilities and conditions detrimental to the physical condition of the individual(s) living in such a unit. This relates to health especially, because of the greater difficulty of maintaining personal hygiene in such a situation.

Thus the percentage of housing units without one or more of the plumbing facilities was calculated. The results in Table V show the same tendency to a concentration in the southeast of the city, although the two enumeration districts with the very high scores on this count are on the fringes of the older city near the university. Their high values may therefore have been due to the larger frequency of student apartments in these districts.

The relationship of the variables and elements selected with the concept of the environment adopted in this thesis. In chapter I it was stated that the environment was considered as the room for living of an individual, and would be construed therefore as a largely social environment.[13] At this point it seems pertinent to ask whether the measures of the environment chosen fit this concept.

The non-white and marital status variables are concerned with given groups of people as they are distributed in the space of the city. The overcrowding measures reflect the effects of the spatial differences in the social factor of density. The value of housing suggests the distribution of income groups about Manhattan. Only the plumbing facilities

---

[13] See above, pp. 7-9.

element is not related directly to a social factor, and this reflects the physical conditions of the environment. However, the physical conditions of the urban environment used in the context of this chapter are not syncnymous with the physical environment as it is usually defined, where it is regarded as the natural, non-human environment of the earth's surface. But the physical environment discussed here in man-made, and the construction of this environment by man is dictated to a great extent by social and economic forces.[14] Therefore it does not disagree with the concept of a social environment if the variables consider in part or in whole the condition of housing; these variables are likely to reflect social conditions in the urban environment.

---

[14]For further elaboration of this point see: Perloff and Wingo, op. cit., pp. 237-431; E. S. Mills, "The Value of Urban Land," in H. S. Perloff (ed.), The Quality of the Urban Environment (Resources for the Future: Baltimore, 1969), pp. 231-254; B. J. L. Berry, J. W. Simmons and R. J. Tennant, "Urban Population Densities: Structure and Change," Geographical Review, 53: 389-405, 1963.

CHAPTER IV

ANALYSIS

In this chapter the relationship between the distribution of mortality and the quality of the environment in Manhattan will be analyzed. Before the actual analysis it is necessary to consider certain control variables. Then the nature of the technique used will be outlined, so that its use in the analysis will be seen to be justified. The results of the analysis will then be discussed, concentrating on the relationship between total deaths and environmental quality, but with a brief final section considering the main causes of death and their relationship to total deaths and environmental quality.

I. CONTROL VARIABLES

Age. In analyzing the environment as a factor in the distribution of death it would be unwise to suggest that the environment is the only factor at work. Therefore to analyze the given relationship satisfactorily it becomes necessary to control for other probable factors. For death Benjamin suggests that besides the environment there are two other factors involved -- age and sex.[1]

In the case of age, Benjamin suggests that there is "a continuing

---

[1] B. Benjamin, Health and Vital Statistics (London: George Allen and Unwin, 1968), pp. 73-77.

increase in the risk of death as age advances."[2] This suggests that, the more older people in a given area, the greater the mortality rate is going to be. It thus became imperative to see how much the distribution of older people varied about Manhattan. For each of the enumeration districts the percentage of people aged fifty-five and over was calculated. The results are given in Table VI, and from here it can be seen that the distribution of persons fifty-five and over varied considerably throughout the city, with there being generally higher proportions of this age group in the south and east of the city. It was therefore decided to control for age in the analysis.

Sex. Males have a greater risk of mortality at all ages for most causes.[3] As related to age the importance of sex as a factor in death is clear. But there is no obvious spatial relationship between the variation in the sex structure of a population and differences in mortality. It seems to be impractical to assume that because the proportion of males in a given area is higher the mortality rate will also be higher. But because, as Table VI shows, there is considerable variation in the ratio of males to females about the city, and because it could not be stated with complete certainty that the spatial variation of the sexes had no

---

[2] Ibid., p. 75.

[3] Apart from those diseases and causes of death peculiar to women, the only cause of death that has a higher rate among females at a given age is diabetes mellitus. M. Lerner and O. W. Anderson, Health Progress in the United States 1900-1960 (Chicago: University of Chicago Press, 1963), p. 88.

TABLE VI. AGE AND SEX STRUCTURE OF THE POPULATION IN MANHATTAN, 1970[1]

Enumeration District	Percentage over 55	Males per 100 Females
20	9.13	105.98
21	6.71	110.31
22	10.35	108.38
23	21.61	93.62
24	37.52	89.39
25	23.74	84.15
26	5.90	95.39
27	12.36	84.51
28	25.97	96.83
29	12.94	120.81
30	25.83	99.17
31	15.89	105.97
32	9.21	156.76
33	27.20	90.04
34	21.32	98.99
35	20.50	97.75
36	7.20	108.58
37	20.09	83.97
38	18.89	96.53
39	19.23	103.42
40	21.59	96.10
41	12.76	129.25
42	5.84	112.94
43	8.25	114.03
44	13.58	68.43
45	21.32	101.44
46	19.93	91.93
47	16.65	144.36
48	13.78	165.12
49	13.48	98.73
50	12.84	120.58
51	8.36	121.12
City	14.14	106.77

[1]Derived from first summary tapes, 1970 Census of Population.

effect at all on the distribution of mortality, it was decided to control for sex.

II. METHODOLOGY OF ANALYSIS

Nature of the data to be analyzed. In selecting a quantitative technique for analysis it is necessary to consider the form of the data, and most specifically the level of measurement of the data. The mortality, age, sex, and environmental variables are all measured at an interval scale (see Tables IV, V, and VI). This means that the distributions are measured in such a manner that it is possible to define the exact distances between given observations. The data here are measured in units of percentages, ratios, rates, and dollars, which allow the computation of the quantitative distance between observations, and thus this is an interval scale of measurement.

Another important element of the data is the size of the sample (N). If N is equal to or less than thirty it is suggested that "there are not enough cases in the sample to indicate the form of the population distribution."[4] Because of this it is necessary with a sample of less than thirty to restrict analysis to certain tests, and the use of parametric techniques is inhibited. Fortunately the sample size was greater than thirty, and thus the use of parametric statistical techniques was possible.

---

[4] H. M. Blalock, Social Statistics (New York: McGraw-Hill, 1960), p. 142.

Parametric and nonparametric techniques. A parametric test requires the assumption of a normal distribution with its associated parameters, as opposed to nonparametric tests, which do not require this assumption.[5] But a parametric test is more powerful in that it incurs less risk of failing to reject a false hypothesis than does a nonparametric test. This is important because in scientific and statistical methodology one cannot prove a hypothesis to be correct, and must therefore rely on the progressive elimination of faulty hypotheses for the purposes of explanation. A hypothesis not rejected by a statistical test may be accepted as an explanation, assuming the research methodology that involved the test of the hypothesis is acceptable, until a better alternative hypothesis is found. It is therefore desirable to be able to be in a position to reject a hypothesis if it is false.[6]

Multiple regression analysis. The parametric statistical test that was appropriate for the proposed analysis was multiple regression. This technique attempts to predict a single dependent variable (Y) from any number of independent variables ($X_1 \ldots X_n$).[7] In this analysis the rate of death was the dependent variable Y, and other related variables were

---

[5] For a discussion of the normal distribution and its parameters see Ibid., pp. 76-90.

[6] Ibid., pp. 89-96; and M. Cohen, "Scientific Method," Encyclopedia of the Social Sciences 10: 389-395.

[7] For a discussion of the relationship between dependent and independent variables in correlation-regression analysis see Blalock, op. cit., pp. 274-285.

treated as independent (X_1 . . . X_{11}).

The interest of the analysis was not in prediction of the dependent variable, but in the measurement of the degree of relationship between the dependent and independent variables. The multiple regression model can be used to consider the relationship between the dependent variable and one of the independent variables, while controlling for the other independent variables. This type of procedure is known as partial correlation, in which the scores of the control variables are, in effect, held constant at their mean, which is possible assuming a multivariate normal distribution. Because of the nature of this technique it was possible to treat age and sex, for the effects of which it was desired to control in the overall analysis, simply as two of the independent variables. This allowed both for the effects of these two variables to be held constant in the analysis of the death-environment relationship, and for the actual effects of these two variables on the distribution of mortality to be studied.

Methods used in computation. As the brief discussion above would suggest, the multiple regression and partial correlation techniques are complex in their claculation, and it would be very difficult and unnecessarily time consuming to use them manually. Therefore use was made of the BMD 03R program of the Biomedical Center of the University of California at Los Angeles, which is a computer program designed specifically for the multiple regression model. This program is flexible, so that the overall death rate, and those for diseases of the heart, malignant neoplasms, cerebrovascular disease, and pneumonia could in turn be treated as

the dependent variable. Age (percentage of persons over 55), sex (males per 1000 females), and the nine environmental variables were in each case treated as the independent variables. The results of the analysis between these independent and the dependent variables will now be discussed in order.

III. ANALYSIS: TOTAL DEATHS AND THE ENVIRONMENT

From the output provided in the BMD 03R program there was considerable information of value to the analysis. This included the mean and standard deviation of the distributions of the variables, a matrix of correlation coefficients between all the variables in the analysis, the coefficient of determination (R^2) with the proportion of the variation in Y attributable to each of the independent variables, an analysis of variance for the coefficient of determination, the partial correlation coefficient for each of the dependent variables, and a table of residuals. The value of each of these and the information derived will be considered in turn.

The mean and standard deviation of the distributions of each variable. The mean and standard deviation can be used as a guide to see if each distribution is normal. If the standard deviation is less than the mean it suggests the distribution tends toward normality, with no extreme values, in either a positive or a negative direction from the mean, affecting the mean and the standard deviation to produce a skewed distribution. In Table VII the values of the mean and standard deviation for all of the variables considered are given. Of the dependent variables it can be

seen that the main three, total deaths, diseases of the heart, and malignant neoplasms, are essentially normally distributed, whereas the other two had a slight tendency to skew. Of the independent variables only the non-white and Negro populations and plumbing facilities indicate skewed distributions, and all the rest were judged to be essentially normal on the evidence of Table VII. Thus it seemed justifiable to assume a multivariate normal distribution. This is a necessary assumption for mutiple regression analysis, particularly with the use of partial correlation analysis, in which the values of the controlled independent variables are held constant at their means.[8]

The matrix of the coefficients of correlation. This gives the correlation coefficient (r) for the correlation between each pair of variables that is considered, whether dependent or independent. The value of r varies between +1.0 (a perfect positive relationship between the two variables) and -1.0 (a perfect inverse relationship), with 0.0 designating no relationship. From the matrix, which is reproduced in Table VIII, two features were of interest: the relationship of the independent variables to the dependent variable of total deaths, and the relationships between the independent variables themselves. In dealing with these associations it should be stressed that this is simple correlation between pairs of of variables with no control for any of the other variables.

---

[8] For further discussion of the assumption of a multivariate normal distribution see Blalock, op. cit., pp. 297-298.

TABLE VII. MEAN AND STANDARD DEVIATION FOR THE DISTRIBUTION OF EACH OF THE VARIABLES CONSIDERED IN THE ANALYSIS[1]

Variable	Mean	Standard Deviation
Total Deaths	6.69	4.17
Diseases of the Heart	2.55	1.70
Malignant Neoplasms	0.98	0.78
Cerebrovascular Disease	0.52	0.53
Pneumonia	0.49	0.53
Percentage over 55	16.25	7.42
Males per 100 Females	106.09	20.66
Percentage Non-white	5.11	8.78
Percentage Negro	3.53	8.32
Percentage Divorced or Separated	3.32	2.33
Percentage of Families with Female Heads	7.94	4.99
Mean Persons per Room	0.57	0.05
Percentage of Units with more than 1.00 Persons per Room	4.82	2.59
Dollar Value of Owner Occupied Housing	14872.84	6205.39
Mean Monthly Rent of Renter Occupied Housing	89.78	21.37
Percentage of Units Lacking One or More Plumbing Facilities	4.16	4.88

[1]Derived from output of BMD 03R computer program used for the multiple regression analysis.

All of the variables describing the environment, except two, were measured in such a way that the higher the value of an observation for that variable the lower the quality of the environment. For the two exceptions, the housing value variables, a high score was taken to indicate a high quality environment. It was hypothesized that the total deaths would be inversely related to the quality of the environment. Therefore it would be expected to find a positive coefficient of correlation between total deaths and all the environmental variables except for the value of owner-occupied housing and mean monthly rent of renter-occupied housing, which would be inversely related to total deaths. These expected results were found in all cases except one; the variable of mean persons per room was found to be inversely correlated with total deaths. In this case the unexpected result seems to be more a result of a poor measure of the environment rather than negating the hypothesis, in the light of the fact that the other associations were as expected.

For age and sex, the two control variables of the analysis, contrasting results were obtained. Age showed an expected positive correlation with death, but sex, as indicated by the proportion of males in the population, did not. This further lessened confidence in any expectation that the known relationship between death and sex through time would also be found spatially. If a higher proportion of males is spatially associated with a higher incidence of death a positive correlation would have been expected.

The coefficients of correlation were then tested for significance. A null hypothesis of no relationship between the two variables (i.e. $r = 0$)

TABLE VIII. MATRIX OF COEFFICIENTS OF CORRELATION

Variable	01	02	03	04	05	06	07	08	09	10	11	12
01. Total Deaths	1.00	.69	-.30	.30	.29	.70	.51	-.31	.18	-.65	-.73	.36
02. Age	.69	1.00	.40	.13	.14	.53	.50	-.53	-.67	-.44	-.50	.55
03. Sex	-.30	-.40	1.00	-.17	-.22	-.39	-.38	-.35	-.14	.27	.18	.99
04. Non-white	.30	.13	-.17	1.00	.99	.70	.74	.37	.66	-.35	-.50	.27
05. Negro	.29	.14	-.22	.99	1.00	.71	.76	.39	.67	-.32	-.49	.17
06. Divorced and separated	.70	.53	-.39	.70	.71	1.00	.85	.19	.60	-.61	-.74	.21
07. Female heads	.51	.50	-.38	.74	.76	.85	1.00	.10	.50	-.48	-.52	.61
08. Mean persons per room	-.31	-.53	-.35	.37	.39	.19	.10	1.00	.65	-.14	-.93	-.29
09. 1.00 persons or more	.18	-.67	-.14	.66	.67	.60	.50	.65	1.00	-.30	-.45	.13
10. Value of housing	-.65	-.44	.27	-.35	-.32	-.61	-.48	-.14	-.30	1.00	.74	-.41
11. Rent of housing	-.73	-.50	.18	-.50	-.49	-.74	-.52	-.93	-.45	.74	1.00	-.44
12. Plumbing facilities	.36	.55	.99	.27	.17	.21	.61	-.29	.13	-.41	-.44	1.00

was tested in each case. At the .001 level of significance the associations of total deaths and the four variables of age, the percentage of persons divorced or separated, the value of owner-occupied housing, and the mean monthly rent proved to be significant.[9] At the .01 level the percentage of female heads proved to be significant, and at the .05 level the plumbing facilities variable proved likewise. For the other independent variables of sex, non-white and Negro populations, mean persons per room, and 1.00 or more persons per room it was not possible to reject the null hypothesis of no relationship, even at the .05 level of significance. This suggests two possible conclusions: the incidence of death will be more strongly associated with some of the environmental factors and not so strongly with others, or else the variables chosen differ in their ability to measure environmental quality.

The correlations between the independent variables were examined to see how many of these variables were so strongly associated as to act jointly, rather than independently, in the relationship with total deaths. Four of the five elements used in the analysis were each measured by two variables, and not surprisingly the component variables of each of these four pairs were strongly related with each other. This association within each of these four pairs was found to be significant at the .001 level. The actual coefficients varied from a very strong association between non-white and Negro, with weaker coefficients between the marital

---

[9] This means that 999 times out of 1000 such a relationship would not have been obtained by chance. Unless stated otherwise, the .001 level of significance will be used throughout this analysis.

status variables (r = .85), and between the housing value variables (r = .74). The weakest association for any of these four pairs was between the two overcrowding variables (r = .65), but this was still a significant degree of association. It was therefore concluded that each of these pairs would act jointly to a considerable degree in explaining statistically the spatial variation in total deaths around the city of Manhattan.

Outside of these four pairs there were also other strong and significant associations among the independent variables. The two variables for non-white population were both significantly and positively related with the divorced and separated, female heads of families, and 1.00 persons or more variables. The divorced and separated variable was inversely related to both housing value variables, which is in keeping with the probable lessened ability of such persons to afford high cost housing. The mean persons per room variable was very strongly and inversely related to the value of renter occupied housing, suggesting a not too surprising greater crowding in the less expensive rental housing. The 1.00 or more persons per room variable was inversely related to age, which gives further support to the earlier suggestion that because younger families tend to be larger they are more likely to be associated with a high density of persons per room.[10] Finally, the plumbing facilities variable was strongly related to the distribution of males and to the female heads variable. The first of these associations has no apparent logical basis and may

---

[10] See above, p. 39.

therefore be meaningless; the second is more logical in that it implies a relationship between families with female heads and poor quality housing, a not improbable association.

Therefore a considerable degree of interrelationship between the independent variables can be established.[11] This means that most of the variables in the analysis can be considered as acting jointly, and not independently, in terms of explaining statistically the spatial variation of death around Manhattan. But the degree of association of all the variables is far from perfect. The mean persons per room, plumbing facilities, and sex variables are significantly associated with two or less of the other variables. Thus, although most of the variables are so interrelated as to work jointly as a group explaining the variation in death to a considerable degree, some act relatively more independently of the others in this statistical explanation.

The coefficient of determination and analysis of variance. The concept of variance ordinarily denotes deviations from a fixed measure of central tendency, the mean, with the deviations divided into the explained and unexplained variation.[12] In multiple regression analysis the concept of variance is generalized. The explained variation becomes the deviation about the mean predicted by the regression plane, and the unexplained

---

[11] All the described relationships were significant at the .001 level. See above, p. 54, footnote 9.

[12] Blalock, op. cit., 244-248.

variation is the deviation of the values of Y about the regression plane.[13] The coefficient of determination (R^2) derived from the multiple regression analysis can be shown to be the ratio of the explained sum of squares to the total sum of squares.[14] Therefore the higher the value of R^2 the greater the variation about the mean of Y (here total deaths) that can be explained. The R^2 obtained for this analysis was .84; this means that the variables used statistically explained 84% of the spatial variation of mortality in Manhattan. This allows for considerable confidence in the explanatory power of the variables chosen to account for spatial differences in death.

To see if this confidence in the explanatory power of the independent variables was justified statistically, R^2 was tested for significance. A statistic F is computed from the ratio of the explained to the unexplained sum of squares, and used to test the null hypothesis of no relationship between the dependent and all of the independent variables acting together (i.e. $R^2 = 0.0$). The computed F statistic was 11.2852, which proved to be significant at the .001 level.[15] This means the computed R^2 would be obtained by chance only one time in a thousand if there was no relationship between the variables, and thus the test of significance further justifies

---

[13] "Explained" does not infer causal explanation, but is used in the sense that the knowledge of the association of Y and $X_1 \ldots X_n$ has helped reduce the amount of error in prediction.

[14] Blalock, op. cit., pp. 297-298. The sum of squares of the deviation about the mean is used to calculate the variation.

[15] The table F statistic to allow significance at this level was approximately 5.28.

the confidence that the independent variables acting together did explain a large proportion of the spatial variation in the dependent variable, the total occurrence of death about Manhattan, Kansas.

The sum of squares attributable to regression, and therefore the coefficient of determination, can be broken down in such a way as to enable the proportion of the variation attributable to each independent variable to be calculated. This is done by the computation of the ratio of the sum of squares for each independent variable to the total sum of squares. This is essentially the same way that R^2 was derived, except that now the variation is that explained by one rather than all of the variables.

In Table IX it can be seen that nearly 48% of the total variation was attributable to age, and nearly 36% to the environmental variables. Sex proved to be insignificant in that it explained less than 1% of the variation. The amount of variation attributable statistically to the environmental variables is sufficient to justify the conclusion that the environment is an important factor in expalining death, and therefore health, in an urban situation. Further refinement of the variables used to measure environmental quality would quite probably increase the power of the environment as an explanatory factor. The importance of the environment in explaining death in the city strengthens the validity and importance of the hypothesized inverse relationship between death and environmental quality.

The importance of the individual environmental variables in statistically explaining the spatial variation in mortality is difficult to assess

TABLE IX. PROPORTION OF THE VARIATION EXPLAINED AND PARTIAL CORRELATION COEFFICIENTS FOR THE INDEPENDENT VARIABLES[1]

Variable	Proportion of the Variation Explained (per cent)	Partial Correlation Coefficient
Percentage over 55	47.73	-0.13
Males per 100 Females	0.07	0.04
Percentage Non-white	4.53	-0.33
Percentage Negro	2.41	0.33
Percentage Divorced or Separated	14.18	0.58
Percentage of Families with Female Heads	1.77	-0.33
Mean Persons per Room	4.57	-0.60
More than 1.00 Persons per Room	0.09	0.24
Value of Owner-occupied Housing	7.14	-0.52
Mean Monthly Rent of Renter-Occupied Housing	0.16	0.19
Units Lacking One or More Plumbing Facilities	1.59	0.30

[1]Derived from output of BMD 03R computer program used to compute the multiple regression analysis.

because of the previously suggested close association of the independent variables. Assuming that each pair tends to explain the variation jointly, it might be most valid to suggest that marital status explained 15.94%, housing value 7.30%, non-white population 6.93%, overcrowding 4.64%, and plumbing facilities 1.59% of the spatial variation of deaths in Manhattan.

The partial correlation coefficients. The output of the program included a partial correlation coefficient between the dependent variable of total deaths and each independent variable. It is difficult to extend the analysis too far from these coefficients because in the real world the independent variables act together upon the dependent variable in a manner not exclusive of each other, as was suggested by the considerable interrelatedness among the independent variables in the matrix of correlation coefficients. Table IX shows the partial correlation coefficients that were obtained. Perhaps the most interesting of the results is the negative relationship between total deaths and age obtained when all the other variables are held constant.

The table of residuals. Residuals represent the degree of error in the prediction of Y by the regression plane. The table of residuals (see Table X) therefore permits the analysis of the deviation of the observed values of Y (total deaths) around the regression plane, so that it is possible to identify where the degree of overprediction and underprediction was most marked. The residuals were stated in the output of the program in the metric of the dependent variable, which here was rate of death. This makes it difficult to assess the degree of error unless the

TABLE X. TABLE OF RESIDUALS[1]

Enumeration District	Y value[2]	Y Estimate[3]	Residual	Rdl/S_{yx}[4]
20	4.56	4.30	0.26	0.13
21	1.72	3.15	-1.43	-0.70
22	4.38	1.16	3.22	1.56
23	8.17	9.44	-1.27	-0.62
24	11.55	12.31	-0.76	-0.37
25	18.40	14.25	4.15	2.01
26	0.54	-0.07	0.61	0.30
27	3.11	5.40	-2.29	-1.11
28	10.64	8.79	1.85	0.90
29	8.44	6.49	1.95	0.94
30	6.25	9.36	-3.11	-1.51
31	7.06	7.98	-0.93	-0.45
32	8.95	8.73	0.22	0.11
33	10.93	11.75	-0.82	-0.40
34	14.21	13.79	0.42	0.21
35	10.93	11.96	-1.03	-0.50
36	2.06	3.23	-1.17	-0.56
37	11.52	10.36	1.16	0.56
38	8.56	8.62	-0.06	-0.03
39	6.25	5.65	0.60	0.29
40	8.88	9.13	-0.25	-0.12
41	4.36	5.78	-1.42	-0.69
42	1.64	1.04	0.60	0.29
43	1.70	0.65	1.05	0.51
44	3.14	5.38	-2.24	-1.08
45	5.02	5.31	-0.29	-0.14
46	7.31	4.44	2.87	1.39
47	1.93	2.60	-0.67	-0.32
48	2.99	3.70	-0.71	-0.35
49	4.25	5.78	-1.53	-0.74
50	9.22	7.29	1.93	0.94
51	5.37	6.29	-0.92	-0.45

[1]Derived from output of BMD 03R computer program used to compute the multiple regression analysis.

[2]Rate per 1000 for all deaths.

[3]Rate per 1000 for all deaths predicted by the regression plane.

[4]The residual divided by the standard error of the estimate.

metric is converted to a standardized form. Therefore the residuals were divided by the standard error of the estimate (S_{yx}).[16] In this case the S_{yx} was 2.06, and the residuals ranged from 4.15 in enumeration district 25 (Residual/S_{yx} = 2.01) to -3.11 in enumeration district 30 (Residual/S_{yx} = -1.51).

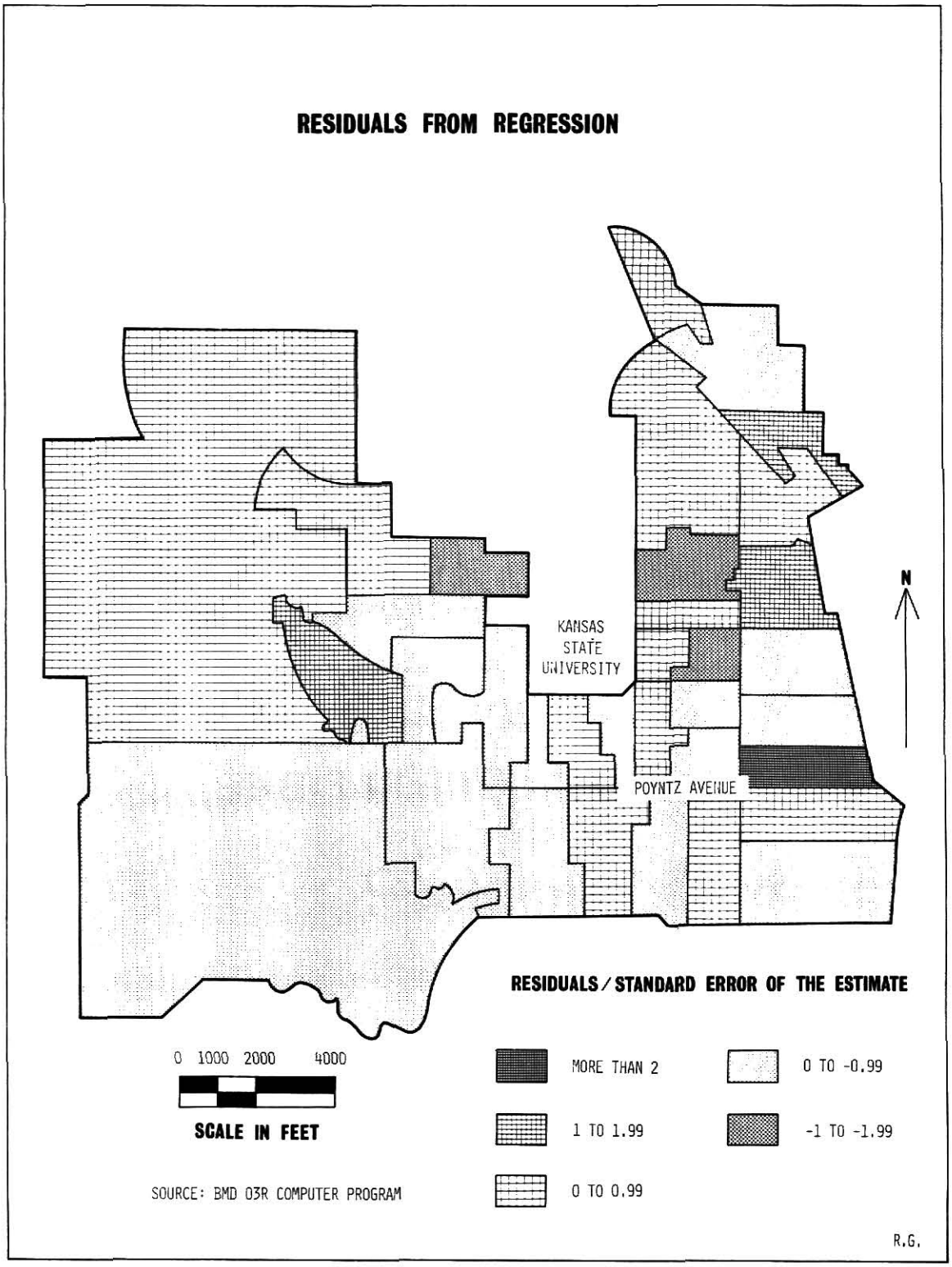
If the residuals are distributed normally about the regression plane 68% of the values of Y will be within one standard error of the estimate, and 95% within two standard errors. As Table X reveals, twenty-six of the thirty-two observations (81%) for total death were within one standard error, and only one Y value was more than two standard errors from the regression plane (i.e. 97% were thus within two standard errors). This suggests the observations for Y were as a whole close to the regression plane, and this suggests the prediction of the death rates by the independent variables was accurate to a considerable degree. This allows for further confidence in the ability of the chosen independent variables to explain the spatial distribution of death in Manhattan.

The residuals for total death were then mapped, to see where the greatest errors in prediction had been made. A positive residual denotes underprediction, because the value specified by the regression plane was less than the observed value, and therefore a negative residual denotes overprediction. The pattern of prediction shown on Map III shows no

---

[16] The standard error of the estimate expresses the standard deviation of the observed values of Y from the regression plane. The lower the value of the standard error, the closer the observed values of Y to the plane.

MAP 3


RESIDUALS FROM REGRESSION
KANSAS STATE UNIVERSITY
POYNTZ AVENUE
N
RESIDUALS / STANDARD ERROR OF THE ESTIMATE
MORE THAN 2
1 TO 1.99
0 TO 0.99
0 TO -0.99
-1 TO -1.99
0 1000 2000 4000
SCALE IN FEET
SOURCE: BMD 03R COMPUTER PROGRAM
R.G.

distinct association of negative or positive residuals spatially, and it was concluded that the pattern was thus essentially random, with no marked tendency to overpredict or underpredict in any given area.

The greatest degree of underprediction was in enumeration district 25, where the death rate was the highest for the city as a whole. This district includes part of the central business district, has the smallest population of any of the enumeration districts, and 9.8% of the population is 75 years or over, a high proportion relative to the rest of the city. No other prediction was more than two standard errors away from the estimate. The greatest overprediction was in enumeration district 30, where the total death rate was only 6.25 per 1000. Here there were many persons over 55, and a higher death rate would have been expected. The people ready to die may move on, or the high proportion of females in the over 55 age group may suggest greater longevity of older persons in this district.

Brief summary of the results of the analysis between total deaths and the environment. From investigation of the means and standard deviations it was possible to assume a nultivariate normal distribution. The matrix of coefficients of correlation suggested an inverse relationship between death and environmental quality. The independent variables were found to act together, in statistically explaining the variation of death, to a considerable degree. They were found to explain 84% of the variation of mortality, with the environmental variables explaining 36% of this variation. The coefficient of determination was found to be significant at the .001 level of significance, allowing confidence in the independent variables

chosen to explain the spatial variation of death. A brief analysis of the table of residuals inferred that the prediction by the independent variables was generally accurate.

IV. ANALYSIS: SPECIFIC CAUSES, TOTAL DEATHS, AND THE ENVIRONMENT

Specific causes and total deaths. From the matrix of correlation coefficients it was possible to determine how closely the distribution of individual causes of death were associated with the distribution of total deaths. The coefficients suggested a high degree of correlation for heart disease, cerebrovascular disease, and pneumonia with total deaths ($r = .92$, .81, and .77 respectively). But the correlation coefficient for malignant neoplasms is much lower than for the others, although still significant at the .001 level of significance ($r = .59$). The coefficients matrix does not allow a satisfactory explanation for this to be derived, although malignant neoplasms do have an interestingly high relationship with the percentage non-white in the population. In order to see whether this could be accounted for, the degree of explanation attributable to the independent variables was investigated.

Specific causes and the independent variables. For malignant neoplasms the degree of success in explaining the variation in mortality due to that cause was not great. The R^2 was only .52, and thus only 52% of the spatial differences in deaths from malignant neoplasms were explained. Of this age explained 34%, sex 1%, and the environmental variables only 17%. The

F statistic was 2.02, which was not significant at the .05 level. It would therefore seem the spatial variation in mortality due to malignant neoplasms may be explained by another factor not considered in the analysis.

Of the other causes of death heart diseases were the most similar to total deaths, not surprisingly in light of the fact that they account for over one-third of all deaths. The R^2 was almost the same (.836), and the breakdown of the explained variation attributable to the independent variables had age accounting for 45% and the environmental variables 39%. For cerebrovascular disease R^2 was only .59, which was barely significant at the .05 level. The proportion of the variation attributable to age was only 17%, which was a very surprising result for a disease strongly associated with age. With only 55 deaths due to this cause spread over 32 enumeration districts it is probable that there are not enough cases to leave confidence in the results of an analysis this precise. The same comment can be made for pneumonia, although the R^2 here was .74 (this is significant at the .01 level), with 42% of the variation attributable to age and 30% to the environmental variables.

Therefore the individual causes of death were found to vary to some degree in their relationships with total deaths and the independent variables. Heart disease showed the closest correspondence to the distribution for total deaths, and the others were less strongly associated with this variable. Generally age was the best variable in terms of explaining the cause of death, with the exception of cerebrovascular disease, where the most important of the variables in the statistical explanation of the

variation was marital status. The choice of independent variables was found to be most satisfactory in explaining the spatial distribution of total deaths, heart disease, and pneumonia; they were less successful with malignant neoplasms and cerebrovascular disease.

CHAPTER V

SUMMARY AND CONCLUSIONS

Before making the concluding remarks that are the primary concern of this chapter, suggestions for further research will be made, and the findings of the research will briefly be summarized.

I. SUGGESTIONS FOR FURTHER RESEARCH

The main concern of this thesis has been with the spatial association between death and the quality of the environment. If the need for health care reflects the quality of the environment, it should be asked how adequate the distribution of health services is in terms of the spatial association of the two elements. From this question two possiblities for further research emerge.

In Manhattan future research could consider whether the present pattern of health services is meeting the needs of the population of the city. Previous research indicated that the physicians have been moving away from the southeast of the city, which is the poorest quality environment of the city, and may thus be supposed to have the greatest need for medical services of any part of the city. Further research should emphasize the problems of the patients in traveling to the medical facilities available, and should ascertain how aware people are of the range of medical services available to

them.[1]

The methods developed in this thesis could also be applied in a larger city, where the environment is likely to be more varied spatially than in Manhattan, and the provision of medical care is more extensive in terms of the range of services provided. This research could more fully integrate the investigation of the relationship between health and death and the environment with the investigation of the adequacy of health services. This approach could broaden by a considerable amount the narrow perspective of the study of urban health services, on the part of geographers, in the Chicago Regional Hospital Study. Such research could also have policy implications, suggesting possible improvements in the system of health care of the city concerned.

II. SUMMARY

It was discovered that research into health has been approached in two ways. Most medical geographers have concentrated on describing patterns of health and their relationship to the environment. In contrast, urban geographers have discussed the provision of urban health services from the narrow viewpoint of the availability of hospital care, but have done so in a rigorous quantitative fashion.

This thesis concentrated on health, as measured by mortality, and its

---

[1]For a more lengthy discussion of the points in this paragraph see K. Piercy, The Changing Pattern of Health Services in Manhattan (Unpublished term paper, Department of Geography, Kansas State University, December 1970).

relationship to the urban environment of Manhattan, Kansas. It was found that mortality showed much spatial variation in its distribution about the city. Generally the area of highest death rates was in the south and east of the city, around the central business district. The incidence of death tended to decline in concentric fashion away from the downtown area.

To measure the relationship of the distribution of death to the quality of the urban environment nine variables were selected as indicative of the spatial variation in environmental quality. A multiple regression analysis was then used to test for the existence of such a relationship, controlling for age and sex. The results of this analysis confirmed the existence of an inverse relationship between the spatial variation in the quality of the environment and the spatial variation of death. The analysis also stated that the independent variables together explained 84% of the variation in death, age accounting for 48% and the environmental variables 36%.

III. CONCLUSIONS

The research hypothesis of this thesis was that there was an inverse relationship between the spatial variation of death and the quality of the urban environment. The evidence developed by the research tended to support this hypothesis. Therefore it is possible to conclude that death is related inversely to the environment, and that this relationship is strong enough to explain statistically 36% of the variation in death about Manhattan.

The poorest quality environment and the highest incidence of death

are both to be found in the southeast of Manhattan. Using death as an indicator of health, and assuming environmental quality to be inversely related to death, it is concluded that the average personal need for health care is more likely to be greater in this part of the city. In the west, northwest, and north of Manhattan the environmental variables and death rates suggest there is, on the average, less need for health care.

The importance of age as a consideration in the spatial association between death and the environment must be stressed. Age explained a large proportion of the variation in death, and it is likely it would be an important factor in the analysis of morbidity as well, because of the increased risk of disease with age. For the second control variable, sex, there was no evidence to suggest it exerted any influence spatially on the distribution of death.

The environmental variables differed considerably in their association with and ability to explain death. This may have been because some are better measures of the environment than others, or because death may respond in varying degrees to different environmental factors. A need for more sophisticated indicators of environmental quality is indicated, and the use of some measures which were not available to the research, such as income, is also called for.

The need for more reliable, detailed, and informative health statistics in such a form that geographers, among others, can study morbidity proper on an areal basis is pressing. Mortality is a useful substitute, but is more restricted in the aspects of health it covers.

The concentric pattern of mortality in Manhattan may be repeated in other cities. This should be studied further, and it should especially be related to the other variables that vary concentrically about the city center to see if the association is significant. This would provide an excellent point of departure for the more complete integration of the study of urban health problems into urban geography as a whole.

BIBLIOGRAPHY

Abrams, A., J. H. Gagon and J. J. Levin, "Psychosocial Aspects of Addiction," American Journal of Public Health, 58: 2142-2155, 1968.

Benjamin, B., Health and Vital Statistics (London: George Allen and Unwin, 1968).

Blalock, H. M., Social Statistics (New York: McGraw-Hill, 1960).

Chorley, R. J., and P. Haggett (eds), Models in Geography, (London: Methuen, 1967).

De Vise, Pierre, et al., Slum Medicine: Chicago's Apartheid Health System Report No. 6 of the Interuniversity Social Research Committee -- Metropolitan Chicago Area, (University of Chicago, 1969).

Ford, T. R., Health and Demography in Kentucky (University of Kentucky Press, 1964).

Freeman, H. F., S. Levine and L. G. Reeder, Handbook of Medical Sociology (New Jersey: Prentice-Hall, 1963).

Greenes, R. A., and V. W. Sidel, "The Use of Computer Mapping in Health Research," Health Services Research, 2: 243-258, 1967.

Health is a Community Affair Report of the National Commission on Community Health Services, (Cambridge, Massachusetts: Harvard University Press, 1966).

Hochstim, J. R., D. A. Athanasopoulos, and J. H. Larkins, "Poverty Area under the Microscope," American Journal of Public Health, 58: 1815-1827, 1968.

Howe, G. M., National Atlas of Disease Mortality in the United Kingdom (Royal Geographical Society, London: Nelson, 1963).

Kupchik, G. J., "Environmental Health in the Ghetto," American Journal of Public Health, 59: 220-225, 1969.

Learmonth, A. T. A., "Medical Geography in India and Pakistan," Geographical Journal, 127: 10-20, 1961.

Lepper, M. H., et a., "Approaches to Meeting Health Needs of Large Poverty Populations," American Journal of Public Health, 57: 1153-1157, 1967.

Lerner, M., and O. W. Anderson, Health Progress in the United States 1900-1960 (Chicago: University of Chicago Press, 1963).

McGlashan, N. D., "The Nature of Medical Geography," Pacific Viewpoint, 10 (2), 60-65, 1969.

Manhattan City Directory 1969, (Kansas City: R. L. Polk, 1969).

Manhattan Mercury, January 13 and 14, 1971.

Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death, 1965 Revision (Geneva: World Health Organization, 1967) 2 volumes.

Markush, R. E., and D. G. Seigel, "Prevalence at Death," American Journal of Public Health, 58: 544-556 and 772-776, 1968.

May, J. M., "Medical Geography: Its Methods and Objectives," Geographical Review, 40: 9-41, 1950.

May, J. M., "Medical Geography," in P. E. James and C. Jones American Geography: Inventory and Prospect (Syracuse University Press, 1954), pp. 452-468.

May, J. M., The Ecology of Human Disease (New York: MD Publications, 1958).

Morrill, R. L., and R. J. Earickson, "Variation in the Character and Use of Chicago Area Hospitals," Health Services Research, 3: 224-238, 1968.

Morrill, R. L., and R. J. Earickson, "Hospital Variation and Patient Travel Distances," Inquiry, 5: 1-9, 1968.

Morrill, R. L., and R. J. Earickson, "Locational Efficiency of Chicago Hospitals: An Experimental Model," Health Services Research, 4: 128-141, 1969.

Morrill, R. L., and R. J. Earickson, "Problems in Modeling Interaction: The Case of Hospital Care," in K. R. Cox and R. G. Golledge (eds), Behavioral Problems in Geography: A Symposium (Northwestern University, 1969), pp. 254-276.

Morrill, R. L., R. J. Earickson, and P. Rees, "Factors Influencing Distances Traveled to Hospitals," Economic Geography, 46: 161-171, 1970.

Morrill, R. L., and P. Kelley, "Optimum Allocation of Hospital Services in Chicago," The Annals of Regional Science, 3: 55-66, 1969.

Morris, R., "The City of the Future and Planning for Health," American Journal of Public Health, 58: 13-22, 1968.

Muller, C., "Income and the Receipt of Medical Care," American Journal of Public Health, 55: 510-521, 1965.

Murray, M. A., "The Geography of Death in the United States and the United Kingdom," Annals, Association of American Geographers, 57: 301-314, 1967.

Oblinger-Smith, Neighborhood Analysis of Manhattan, Kansas (Wichita, 1968).

Reinke, W. A., and T. D. Baker, "Measuring the Effects of Demographic Variables on Health Services Utilization," Health Services Research, 2: 61-75, 1967.

Stamp, L. D., The Geography of Life and Death (Cornell University Press, 1964).

Suchman, E. A., Sociomedical Variations among Ethnic Groups," American Journal of Sociology, 70: 319-331, 1964.

Terris, M. "Epidemiology of Cirrhosis of the Liver: National Mortality," American Journal of Public Health, 57: 2076-2088, 1967.

Terris, M., "A Social Policy for Health," American Journal of Public Health, 58: 5-12, 1968.

THE GEOGRAPHY OF HEALTH AND DEATH AND THE QUALITY OF THE ENVIRONMENT IN MANHATTAN, KANSAS, 1965-1969

by

KEITH PIERCY

B.A., Sheffield University, 1969

AN ABSTRACT OF A MASTER'S THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER OF ARTS

Department of Geography

KANSAS STATE UNIVERSITY
Manhattan, Kansas

1971

Geographers have considered the problems of health through two different approaches. The more traditional method has been that of medical geography, which has concentrated on the description of the distributions of various diseases and of death. Within this descriptive framework geographers have stressed the understanding of the symbiosis between man and his environment in understanding the distribution of disease. More recently urban geographers have applied techniques based on central place theory to the analysis of the availability of urban health services. But in general they have not gone on to consider the spatial aspects of the distribution of death and disease in urban areas.

This thesis uses the symbiosis of man and his environment, as employed in medical geography, to explain one particular urban health problem, the distribution of death. This concept was applied in the city of Manhattan, Kansas, using the hypothesis that the incidence of death was inversely related to the quality of the urban environment.

In describing the distribution of death it was found that the seven major causes of death were heart disease, malignant neoplasms, cerebrovascular disease, pneumonia, general arteriosclerosis, accidental deaths, and infant mortality. These are the same major causes as for the U.S.A. as a whole. Mortality rates were highest around the central business district of the city, declining in concentric fashion away from the downtown area. The lowest death rates were therefore found on the outskirts of Manhattan in the west, northwest, and north. The concentric pattern of death is interesting in view of the other socio-economic variables such

as income and population density that vary in this manner.

To indicate the quality of the environment nine variables were selected from the 1970 Census of Population to measure five environmental elements. These variables were: percentage non-white, percentage Negro, percentage of the persons over 14 who are divorced or separated, percentage of occupied housing with more than 1.00 persons per room, mean monthly rent of renter-occupied housing, and the percentage of housing units lacking one or more plumbing facilities. These variables indicated that generally environmental quality was lowest in the southeast of the city, and highest in the north and west.

To analyze the relationship between mortality rates and the nine environmental variables, a multiple regression analysis was used. This technique made it possible to control for age and sex, which were other factors that could influence the spatial variation in mortality. The analysis tended to confirm the hypothesized inverse relationship between the incidence of death and the quality of the environment. It also stated that the environmental variables statistically explained 36% of the variation in mortality, thus testifying to the importance of the hypothesized relationship. Of the control variables age explained 48% of the variation (statistically), but sex found to be insignificant.

It was therefore concluded that there was an inverse relationship between death and the environment, and that this relationship played an important part in explaining the spatial variation of death. Further research extending this observed relationship into the analysis of the adequacy of health services would seem to be called for.