71-23,737

DEVER, Alan George, 1941-
INTRA-URBAN VARIATION OF LEUKEMIA OCCURRENCES.

The University of Michigan, Ph.D., 1971
Geography

University Microfilms, A XEROX Company, Ann Arbor, Michigan

THIS DISSERTATION HAS BEEN MICROFILMED EXACTLY AS RECEIVED

INTRA-URBAN VARIATION OF

LEUKEMIA OCCURRENCES

by
Alan George Dever

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Geography)
in The University of Michigan
1971

Doctoral Committee:

Associate Professor Gunnar Olsson, Chairman
Assistant Professor James D. Clarkson
Professor L. A. Peter Gosling
Associate Professor Stanley H. Schuman

ABSTRACT

INTRA-URBAN VARIATION OF LEUKEMIA OCCURRENCES

by
Alan George Dever

Chairman: Gunnar Olsson

This paper is concerned with the spatial aspects of housing, demography and environment in relation to the intra-urban variation of leukemia occurrences. Emerging from previous theory and existing thought on leukemia is the intuitive idea that the intra-urban variation of leukemia might possibly be linked to an infectious process. To explore and test this intuitive idea a regression model was formulated and the empirical parameters were estimated. While testing the relationship of leukemia to the three factors the scale was varied to determine at which scale the most readily interpretable patterns occur. Specifically, four scales were selected. Namely, the census tract, the block, the tax districts and the individual housing unit. It was the tax district scale which suggested that leukemia was occurring in "spatial high risk areas" and "spatial low risk areas." It appears that these high risk areas were a result of variables that suggest a viral transmission. Thus, the evidence found in this paper gives support to the viral etiology theory although it has not been conclusive in this regard.

ACKNOWLEDGEMENTS

This dissertation represents the consistent effort of many individuals. I am indebted to Gunnar Olsson who, while executing his own research in Thailand during the summer of 1970, gathered the time to evaluate and criticize each draft. It was through his unfailing efforts, his constant encouragement and excellent advice that I was able to complete this work. I would also like to thank a fellow graduate student, Dierk Rhynsburger, without whose first-rate assistance I would have encountered time-consuming obstacles. To all my other colleagues at the University of Michigan, for their valued support and counsel, a special thank you. And of course, last but not least, my dear wife Georgie for whom this project represents the culmination of seven years hard, painstaking and often turbulent periods during which she kept the completion of this dissertation economically feasible and emotionally plausible.

TABLE OF CONTENTS

Chapter		Page
I	Introduction	1
	Objectives	3
	Study Area	3
	Scale	4
	Organization	4
II	Medical and Spatial Aspects of Leukemia	7
	Medical Aspects	7
	What is Leukemia?	7
	What causes Leukemia?	8
	Who contracts Leukemia?	10
	The Spatial Aspects of Leukemia	10
	World Scale	11
	National Scale (U.S.A.)	11
	Local Scale (Buffalo, N.Y.)	13
III	Data and Operational Definitions	17
	Introduction	17
	Dependent Variable: Leukemia Incidence Rate (Y)	18
	Independent Variables IA: Housing; Census tract and Blocks (X_1)	19
	Condition of the Housing Unit (x_{101},x_{102})	20
	Plumbing Facilities (x_{103})	21
	Occupancy (x_{104} through (x_{112})	22
	Summary	23
	Independent Variables: IB; Housing unit	25
	Independent Variables: II; Demography (X_2)	28
	Independent Variables: III; Environment (X_3)	29
IV	Leukemia and Housing	33
	Introduction	33
	Housing-Health and/or Disease	33
	Leukemia and Housing	35
	The Analyses of Housing with Various Scales	37
	The Census Tract Scale	37
	The Block Scale	42
	An Aggregated Scale (Tax Districts)	48
	The Micro Scale (the Housing Unit)	54
	Summary	61

Chapter		Page
V	Leukemia and Demography	65
	Introduction	65
	Male/Female Ratio	65
	Leukemia and Age	67
	Leukemia-Duration and Age	69
	All Ages: Acute and Chronic	69
	Age group 14 and under	71
	Age group 15-49	73
	Age groups 50-64 and over 65	74
	Summary of All Age groups	77
	Leukemia and Income	78
	Leukemia and Occupation	80
VI	Leukemia-Seasonal Onset and Air Pollution	83
	Introduction	83
	Seasonal Incidence of Leukemia	83
	Leukemia and Air Pollution	85
	Summary	87
VII	Summary and Conclusions	89
Appendix		95
Bibliography		103

LIST OF TABLES

Table		Page
1	Characteristics of Existing Leukemia Studies	2
2	Hypothesized Relationship of Housing Variables to Leukemia	24
3	The Weighed Relationship Between the Population and Housing Unit Sample	27
4	Hypothesized Relationship of Demographic Variables to Leukemia	30
5	Hypothesized Relationship of Environmental Variables to Leukemia	31
6	Census Tract - Correlation Matrix	38
7	Census Tract Scale - Correlation and Regression Analysis	39
8	Census Tract Scale - Standardized Variables Correlation Matrix	41
9	Census Tract Scale - Standardized Variables Correlation and Regression Analysis	42
10	Block Scale - Correlation Matrix	43
11	Block Scale - Correlation and Regression analysis	44
12	Block Scale - Standardized Variables Correlation Matrix	46
13	Block Scale - Standardized Variables Correlation and Regression Analysis	47
14	Tax Districts - Maximum Partial Correlation Coefficients	51
15	Class of Dwelling - Smirnov Test	56
16	Type of Foundation - Smirnov Test	56
17	Type of Basement in the Home - Smirnov Test	57
18	Type of Construction of the Home - Smirnov Test	57
19	Type of Exterior Construction on the Home - Smirnov Test	58
20	Type of Roof on the Home - Smirnov Test	58

Table		Page
21	Roof Construction - Smirnov Test	59
22	Type of Outside Trim - Smirnov Test	59
23	Type of Heating - Smirnov Test	60
24	Source of Household Power - Smirnov Test	60
25	Condition of Housing Unit - Smirnov Test	60
26	Observed Relationships of The Tested Variables Housing and Leukemia	62
27	Sex Ratios for Leukemia in Buffalo, N. Y. 1959-1966	66
28	Age-Specific Incidence Rates of Leukemia - Buffalo, N.Y. 1959-66	68
29	Regression and Correlation Analysis - Acute and Chronic Leukemia Cases - All Ages	70
30	Regression and Correlation Analysis - Acute and Chronic Leukemia Cases - Ages 14 and under and 15-49.	72
31	Regression and Correlation Analysis - Acute and Chronic Leukemia Cases - Ages 50-64	75
32	Regression and Correlation Analysis - Acute and Chronic Leukemia Cases - 65 and over	76
33	Observed Relationship of the Tested Variables Leukemia and Demography	79
34	Leukemia and Median Income	80
35	Leukemia and Occupation	81
36	Seasonal Variation of Leukemia By Month of Onset 1959-1966	84
37	Leukemia and Air Pollution - 1963-1966	87
38	Observed Relationship of the Tested Variables Leukemia and Environment	87
39	Summary of Significant Findings	90

LIST OF ILLUSTRATIONS

Figure		Page
1	World Distribution of Leukemia.	12
2	Distribution of Leukemia - United States.	14
3	Distribution of Leukemia - Buffalo, N.Y..	15
4	Buffalo Tax Sections.	49
5	Leukemia Intra-urban Risk Areas.	53
6	Distribution of Leukemia 1963-66 and Air Pollution...	86

CHAPTER I

INTRODUCTION

Diseases, like other phenomena, vary spatially over the surface of the earth. For diseases like multiple sclerosis these variations may depend on the distribution of environmental factors (Dean,1970). In cases like the Asian Flu, on the other hand, the spatial distributions can be better understood through analyses of existing linkages in time and in space. It is such spatial and temporal linkages that have to be seriously considered if further hypotheses on the geographical variations of disease are to be formulated.

The temporal-spatial variation of leukemia occurrences has been viewed sometimes as a random and sometimes as a clustered phenomenon (Dever,1969; Ederer,1965; Fraumeni,1966). These two conflicting views are a question of considerable epidemiologic interest and the main studies concerned with this topic are listed in Table 1. For this reason, it is the purpose of this dissertation to analyze the relationships between the intra-urban pattern of leukemia and of several demographic and environmental factors.

Epidemiologists have long been interested in the geography of diseases. In general, however, they have considered only spatial variations on the least detailed spatial scale such as broad geographic areas, urban-rural differences, and inter-city variations. Furthermore, epidemiologists have usually considered urban areas as homogeneous

TABLE 1

CHARACTERISTICS OF EXISTING LEUKEMIA STUDIES

Author	Year of Study	Temporal Limit	Spatial Unit
Pinkel and Nefzger	1959	7 - 2 year periods[a]	4,779 3-dimension cells
Wood, E. E.	1960	5 years[b]	Cornwall, England
Heath, C. W. et al.	1962	3 years[b]	2 square miles
Knox, G.	1964	10 years[b]	3100 square miles
Ederer, F. et al.	1964	3 - 5 year periods[a] (by year of occurrence)	Connecticut town size
Fraumeni, J. F. et al.	1966	3 - 5 year periods[a] (by year of birth)	417 cities and towns in New York State
Dever, G. E. A.	1968	8 years[b]	Square and radial spatial units in Buffalo, New York

[a] Temporal limit represents discrete intervals with no overlap.

[b] Temporal limit represents duration of the study in years.

entities thereby neglecting spatial intra-urban analyses. In contrast, geographers' interest in general intra-urban studies are immense although only a handful of geographers have contributed to intra-urban medical research.

Objectives

The general objective of this study is to determine whether incidences of leukemia can be associated with variations in selected housing, demographic, and environmental factors. Provided the analysis indicates that a specific variable or variables are highly correlated to the disease, then the next step is to undertake a medical pilot study.

Within the context of this general objective, the specific objectives of this study are: (1) to determine the intra-urban spatial distribution of leukemia incidence in Buffalo, New York for the years 1959-1966; (2) to determine the relationship between the spatial incidence of leukemia and selected housing characteristics; (3) to determine the relationship between the spatial incidence of leukemia and selected demographic and environmental characteristics; and (4) to determine if varying the scale of analysis will suggest more reliable interpretations of the patterns of the disease.

Study Area

The area selected for the analysis is the city of Buffalo, New York. Buffalo represents a heterogeneous conglomeration of population groups living under different housing and environmental conditions. It is possible, therefore, that the variation of these variables may be

statistically associated with high or low risk areas within the city. Analogous high and low risk areas have been delineated for other phenomena; for instance, Anderson showed evidence of high pregnancy risk areas in Buffalo (Anderson,1965), while Graham has indicated differences in the socio-economic distributions of various cancer sites in Buffalo, New York (Graham,1960).

Scale

In formulating and testing new theories on the etiology of disease the spatial scale of the analysis forms an intricate part of the theory itself. In particular, by varying the units of analysis it may be possible to determine at which scale the most readily interpretable patterns occur.

The hypothesis that leukemia incidence is an infectious disease suggests that the analysis should be on the micro level. Indeed, since leukemia is a relatively rare disease, it may even be suggested that the cause, like the disease, may represent a rare occurrence and, as a consequence, would not be noticed if the data were aggregrated. Nevertheless, the study of aggregrated data can sometimes provide important clues for a subsequent macro level analysis.

Organization

The present study consists of seven chapters, entitled (1) Introduction; (2) Medical and Spatial Aspects of Leukemia; (3) Data and Operational Defintions; (4) Leukemia and Housing; (5) Leukemia and Demography; (6) Leukemia - Seasonal Onset and Air Pollution; and (7) Summary and Conclusions.

Chapter two was prompted by the particular problems encountered in studies devoted to disciplines with which the reader is not usually acquainted. The case of medicine and geography is noted, and some medical information about leukemia is provided. Finally, the spatial distribution of leukemia is pictured for the world, national (U.S.A.) and the intra-urban (Buffalo, New York) levels.

Chapter three details the data sources employed and attention is drawn to the use of the regression model and the Kolmogorov-Smirnov tests which will be used in the subsequent analysis.

In chapters four, five, and six the relationship between leukemia incidence and selected housing, environmental and demographic aspects are analyzed. The chapters describe and test the hypotheses relative to each of these three categories. The housing variables analyzed are the following: (1) plumbing (water supply); (2) condition of the housing unit; (3) occupied, i.e., renter or owner; (4) persons per room; (5) average number of rooms; (6) average value of the housing unit; (7) average contract rent; (8) type of heating; and (9) other varibles directly related to the construction of the house. The demographic variables employed are (1) age; (2) sex; (3) occupation; and (4) median income, while the analyzed environmental variables include (1) temperature; and (2) suspended particulate data.

Various aspects of housing, demography and environment have had a marked impact on the cause, type and prevalence of disease (Hochstim,1968; Gilbertson,1964; Suchman,1963; Graham,1963). If the theory is limited to the hypothesis of single cause, then housing, demographic and environmental factors cannot be viewed as directly causing disease (Suchman,1963). If, on the other hand, the theory of multiple causation

is accepted, then the same factors may be viewed as one of the direct causes of disease (Suchman,1963). Especially in chapter seven, it is noted that leukemia may be viewed within this context of multiple causation. However, the present study searching for causation seeks association. It follows that if association is established then further research is warranted; those variables for which association is noted should provide the basis for the design of a pilot-study in an intra-urban area.

Chapter seven reviews what has been demonstrated in the previous chapters.

CHAPTER II

MEDICAL AND SPATIAL ASPECTS OF LEUKEMIA

Medical Aspects

The relationships between leukemia incidence and location has received even recent attention from researchers in the medical field. The emphasis has usually been on noting the significance of temporal-spatial clusters. For instance, this was the case in the famous study from Niles, Illinois (Heath and Hasterlik,1963) in which references to the concept of clustering were used to support the viral etiology hypothesis.

The characteristics of leukemia are well documented in some instances and rather vague in others. For example, its diagnosis is normally straight forward but the causes of leukemia are not known. The individuals who contract leukemia frequently share common demographic characteristics. But, this is not always the case. To shed further light on those issues the following three questions will be discussed. These questions are: (a) what is leukemia? (b) what causes leukemia? and (c) who contracts leukemia?

What is Leukemia?

The word "leukemia" stems from a Greek term meaning "white blood" (Gr. leukos-white, Gr. aimia-blood). It is a disease of the blood forming tissues characterized by an abnormal proliferation in the

number of leukocytes. These abnormal cells infiltrate other organs and tissues of the body and cause extensive damage. Any of the different types of blood cells may be involved in the process. The disease can be classified either as chronic, with the cells showing some maturation in development and with the disease progressing slowly, or as acute, with the cells showing primitiveness or "youngness" and with the disease progressing rapidly.[1]

Some types of leukemia occur more frequently than others as exemplified by the varieties Granulocytic (Myelogenous), Lymphocytic, Monocytic and Stem cell leukemia (Primitive white cells). Some less common varieties are Eosinophilic, Basophilic, Megakaryoctic, Reticulum cell, Plasma cell, and Erythroleukemia.

The characteristics and definitions of leukemia are relatively concise. However, the cause of leukemia is subject to extensive discussion and there are many plausible theories.

What causes Leukemia?

It is known that radiation and some chemicals can cause leukemia in humans and that viruses can cause leukemia in animals such as rats, mice and chickens (Beard,1963; Lewis,1957; March,1950; Wald,1958). Building on these latter findings are a growing number of scientists who are leaning towards the view that viruses may play a role in some, or possibly all, cases of leukemia-like diseases in human (Editorial,1963; Grace,1964).

As already mentioned, the classical study of a leukemia out-

---

[1]For a more detailed discussion on leukemia, refer to Maile, J. B., Laboratory Medicine-Hematology, The C.V. Mosby Company, 1962.

break is that from Niles, Illinois, a suburb of Chicago (Heath and Hasterlik,1963; Schwartz and Greenspan,1963). This outbreak provided the first opportunity to investigate a cluster of leukemia cases in a relatively small population. The 13 children that contracted the disease lived within a radius of two miles of the parochial school, and almost all had moved to Niles from Chicago. Nothing unusual was noted about their homes, including natural radiation from the soil and buildings. In fact, the only characteristic shared by the 13 cases was the spatio-temporal co-incidence. For this reason, it is firmly believed that the outbreak may have been caused by a virus. This suspicion is further supported by the fact that in cases of viral infections, one of two things may happen to an infected person: (1) antibodies may form and destroy the virus; or (2) antibodies do not form and the disease will take its course. None of the 13 cases possessed any antibodies against leukemia.

The Niles outbreak may be a chance or happenstance occurrence. However, there are several other examples where clusters of outbreaks have been found (Dever,1968; Fraumeni,1966; Knox,1965; Mustacchi,1965; Pinkel and Nefzger,1959; Wood,1960). But, the evidence of clustering in an area is not nearly as shocking as clustering of several leukemia cases of unrelated individuals living in the same house (Gilmore and Zelesnick,1962; McPhedran and Heath,1969). For instance, in a small town in the southern United States, there were four leukemia victims between 1950 and 1967 that had been associated with the same house. As in the Niles cluster, the house associated with the multiple cases was not in any way peculiar when compared to its immediate surroudings. This suggests that other etiologic factors may have been operating in the

home environment (Gilmore and Zelesnick,1962; McPhedran and Heath,1969).

In such situations, further studies of the personal characteristics of those individuals who contract leukemia may lead to alternative causal hypotheses.

Who contracts Leukemia?

Leukemia is most common in two age groups. Thus, children under 15 years of age and adults over 65 years have the highest likelihood of contracting the disease. Furthermore, it is more prevalent in the white population than in Negroes. The same holds for males as compared to females; the ratio is almost three males to every two females. Jews display a greater prevalence of leukemia but have lower rates of other forms of cancer. These examples show that there is no one demographic aspect which predisposes an individual to leukemia. At the same time, it is clear, that the probability of contracting the disease is subject to variation in age, sex and race of the individual.

The Spatial Aspects of Leukemia

The preceding discussion was focused on leukemia as discussed in the medical literature. In comparison, the existing geographical studies have been more concerned with the analysis of temporal-spatial clusters. Two such studies have been completed in Buffalo, and both indicate a clustered distribution of outbreaks (Dever,1968; Pinkel and Nefzger,1959). However, the concept of spatial clustering is inherently dependent on the selected geographic scale. For this reason, it is the intent of this section to discuss the spatial distribution of leukemia at various scales.

World Scale

The world scale (Figure 1) of leukemia incidence displays an interesting pattern. Thus, the age-standardized mortality rates per 100,000 at this scale vary from a high of 10,1 in Canada to a low of 2.2 in Colombia. A contiguous area of high mortality exists in North America. In addition to Canada and the United States, Israel, Denmark and Peru are leukemia stricken countries. The generally high incidence rates among the Jewish population are reflected in the high Israelian rates. In contrast, the high mortality rates in Denmark and Peru cannot be accounted for. The fact is that we do not know why higher rates exist in these countries.

Absence of the disease or at least low mortality rates associated with it is reported for Colombia and Japan. At present, hypotheses related to this pattern are relatively rare. The mortality rates for the other countries vary from 4.0 to 7.9 per 100,000 but there is a large number of countries with no data available. As Bailar notes, that substantial geographic variation in completeness of diagnosis is a relevant consideration in the interpretation of the pattern (Bailar,1967). In conclusion, it is evident that spatial variation exists on the world scale but the variation is based only on a few reporting countries.

National Scale (U.S.A.)

The spatial distribution of leukemia in the United States has been reported for all states except Alaska and Hawaii. On map 2 the state mortality rates have been divided into three categories. The highest rates generally occur in the northern states, such as New York, Wisconsin, Minnesota, Iowa, Montana, Idaho, and Nebraska. Exceptions

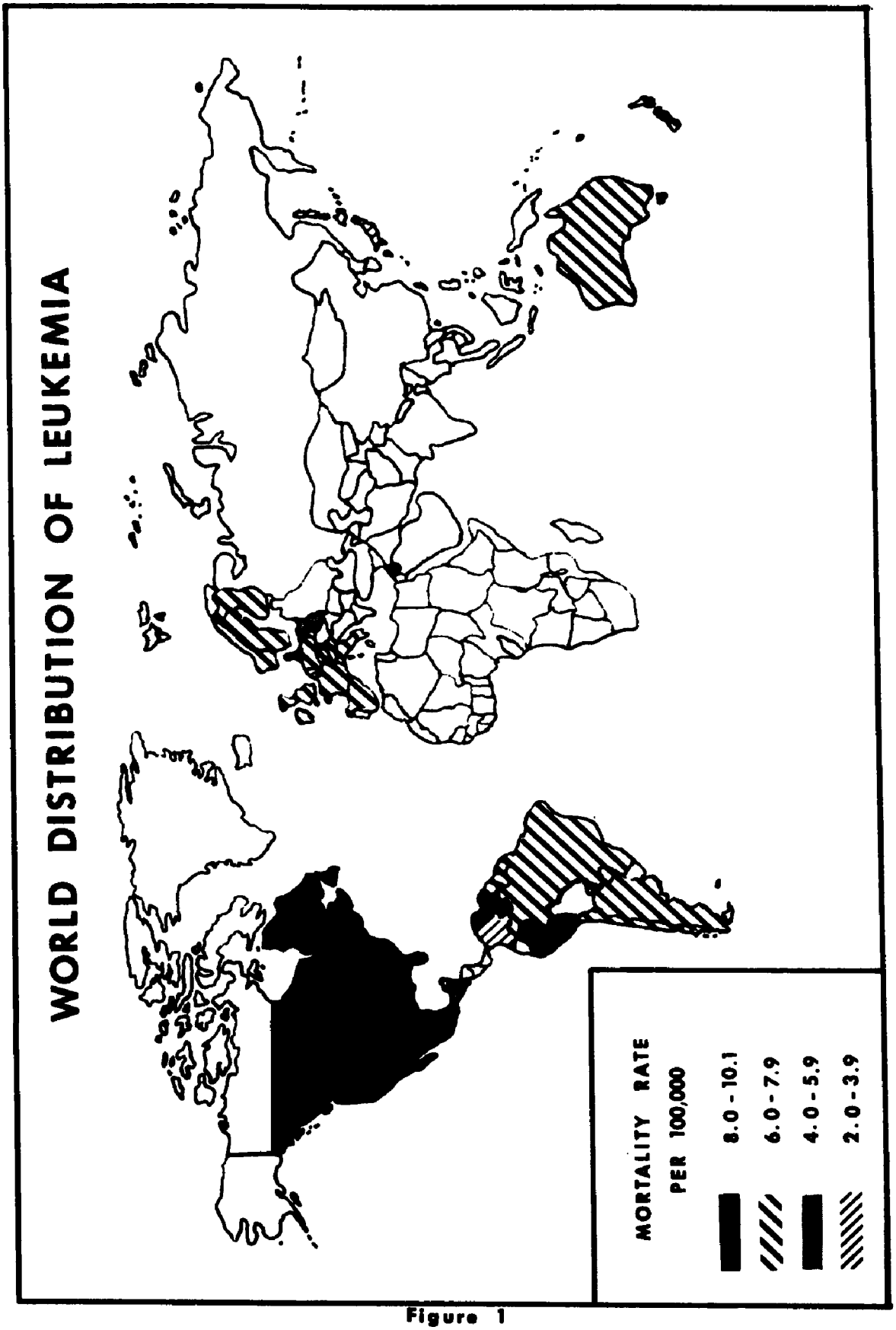

WORLD DISTRIBUTION OF LEUKEMIA
MORTALITY RATE
PER 100,000
8.0 - 10.1
6.0 - 7.9
4.0 - 5.9
2.0 - 3.9

Figure 1

to this are the high incidences in the two states of Kansas and California. The lowest rates occur mostly along the eastern seaboard and stretch into the southern states. The map suggests that leukemia rates in the United States may be a function of latitude, although, the pattern is not distinct enough to provide conclusive evidence. It is necessary, therefore, to pursue the question of spatial variation also on the local level.

Local Scale (Buffalo, N.Y.)

The leukemia distribution for Buffalo, New York, is portrayed by the block diagram in figure 3. The block diagram was produced by an IBM 7094 computer using an offline ink-on-paper Calcomp plotter. The statistical surface is vertically exaggerated in proportion to the incidence rates per 100,000 individuals as determined for each grid point. The surface indicates the relative difference between the high and low areas where the peaks represent high rates and the valleys represent low rates. As seen from the block diagram, high rates occur in the north-east, south-west and central areas of the city. In contrast, the lowest rates are found in the east-central area. The extremely low flat area in the south-western section of the diagram is Lake Erie and not an area of low incidence. If no intra-urban variation existed the statistical surface would be flat.

Evaluating the porperties of this intra-urban variation will be the primary concern of the subsequent chapters. In so doing we will draw on the medical and geographical properties of leukemia particularly in their relation to existing theories of a viral etiology as discussed in the present chapter. It is self-evident that if the process

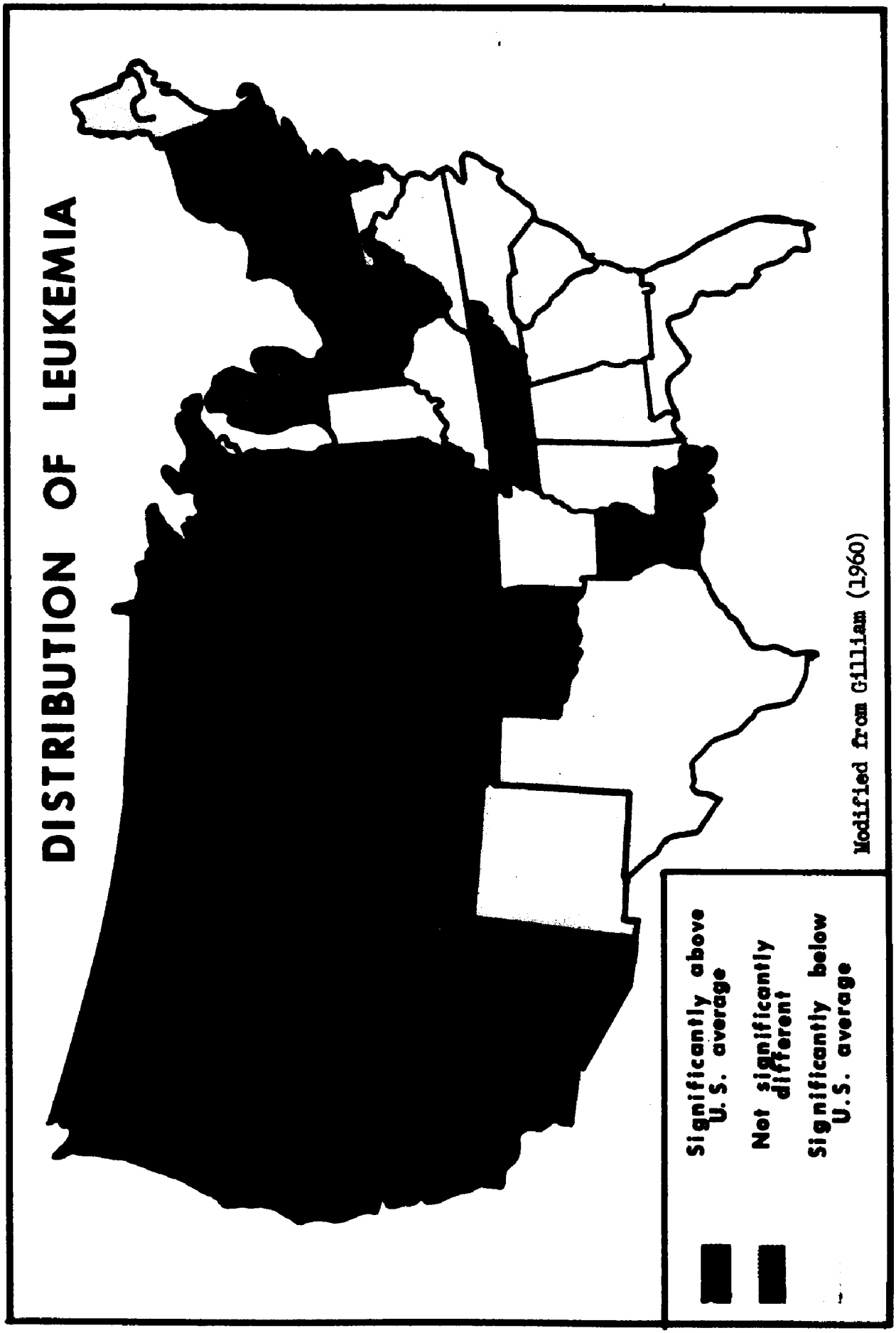

DISTRIBUTION OF LEUKEMIA
Modified from Gilliam (1960)
Significantly above U.S. average
Not significantly different
Significantly below U.S. average

Figure 2

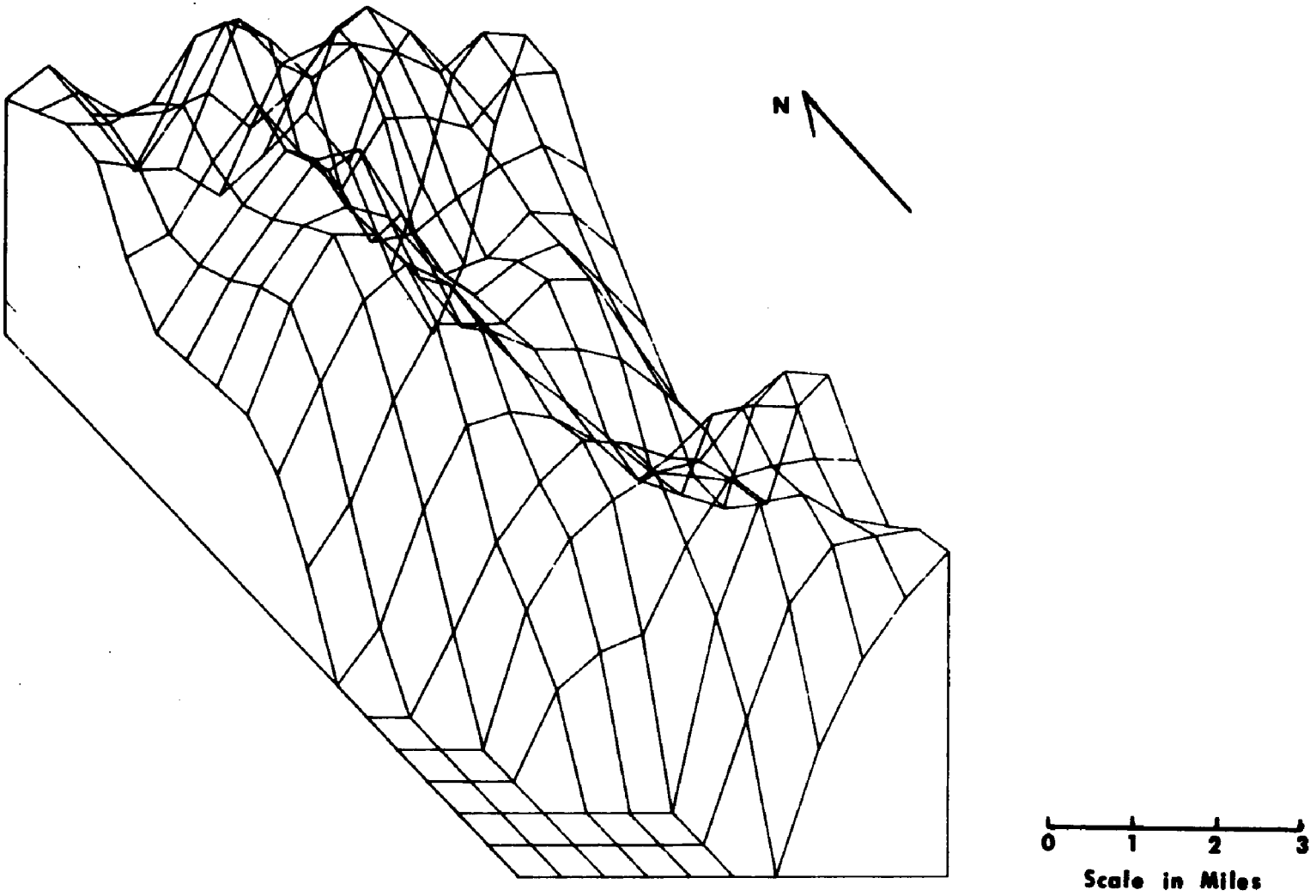


Figure 3. A Block diagram showing the distributies of leukemia is Buffalo, New York, 1959-1966. The exaggerated surface represents incidence rates per 100,000.

that triggers leukemia is infectious then the spatial distribution of the disease should exhibit the clustering properties common to infectious diseases.

Emerging from the several points discussed thus far is the intuitive idea that the intra-urban variation of leukemia might possibly be linked to an infectious process. To explore this idea further it will be necessary to formulate more exact hypotheses through which the relationships between incidence rates and the spatial aspects of demography, physical environment, and housing can be tested. To test these intuitive thoughts, it is necessary to collect data on the particular aspects stated. This will be done through the formulation and empirical parameter estimation of a regression model. In the next chapter, both the model and the employed data will be specified.

CHAPTER III

DATA AND OPERATIONAL DEFINITIONS

Introduction

In the previous chapter it was noted that the incidence rates of leukemia vary geographically at the world, national, and local scales. The local variation of leukemia may be further analyzed at even more detailed scales, although there seem to be other ways in which intra-urban high or low risk areas may be determined. For example, the population-at-risk for each specific spatial area within the intra-urban setting could be considered. Since many diseases are age or sex specific, such analyses could include aspects like age and sex. Unfortunately, relevant data on the population structure of small spatial areas are not always available and other approaches will therefore be sought.

It will be recalled that the simple map comparison in chapter two led to the intuitive suggestion that the spatial variations in leukemia can be linked to an infectious process. In the present chapter, this idea will be further explored but now in the immediate connection with causal factors measured in terms of demography, housing, and environmental characteristics. For instance, if the disease were infectious then certain properties of each factor should provide positive evidence of a viral transmission. On the other hand, characteristics like adequate and clean water, lack of crowding in a home and high socio-

economic status would be suggestive of an non-infectious transmission.

To investigate the importance of the suggested relationships, it will be necessary to treat the intra-urban variation of leukemia as a functional relationship of the form $Y=f(X_1,X_2,X_3)$, where Y stands for the intra-urban variation of leukemia incidence; while X_1, X_2, X_3 represent characteristics of housing, demography and environment. This basic expression will subsequently be operationalized in terms of a regression model whereby the intuitive feeling of an infectious process can be statistically evaluated.

Dependent Variable: Leukemia Incidence Rate (Y)

To investigate the spatial intra-urban variation of leukemia, it has been necessary to collect some very specific data. The data on leukemia for Buffalo, New York were obtained from Albany, New York because all cancer reports, including leukemia, in New York state are sent to the state's Bureau of Cancer Control.[2] The data on all leukemia cases were collected from different files (the live file; the dead file; and the death certificate).

To exerpt data on any disease, one must know the code number. According to the Eighth Revision of the International Classification of Diseases, leukemia occupies a specific site code. Under the Eighth Revision the site codes of 200 through 207 include the neoplasms of Lymphatic and Hematopoietic Tissue. However, under the Seventh Revision the site codes are 200 through 205. This change affects leukemia only

---

[2]Appreciation is expressed to William S. Burnett, M.D. Assistant Director, Bureau of Cancer Control, State of New York Department of Health., 855 Central Avenue, Albany, N.Y. for assistance in the collection of the data.

minimally in this investigation and both ICD Revisions are listed in Appendix 1 and 2 respectively.

From the appropriate data files and using the above mentioned codes the exact street address of every leukemia case was noted. To locate the leukemia death as accurately as possible in each block and further in each census tract, a street map of the city of Buffalo and the Polk city directory were used. The directory from each year was employed for the exact determination of the place of residence in the specific year of death. The 1960 total population for each census tract and each block in which a leukemia case occurred was then observed. Since this investigation involves an eight year period (1959-1966), the 1960 population figure was multiplied by eight to give a more correct population-at-risk for the study period. On the basis of this information the incidence rates were calculated for each tract and block in the following manner:

$$Y = \frac{\text{number of leukemia occurrences}}{8 \times 1960 \text{ population}} * 100{,}000 .$$

It is the incidence rate calculated as above that will be used as the dependent variable in the subsequent analysis.

Independent Variables IA: Housing; Census Tract and Blocks (X_1)

There are three levels of aggregration for which information exists on housing characteristics, namely (1) census tract; (2) block; and (3) the housing unit itself. Data for tax districts can be aggregrated from the block statistics. It is the intention to use data for these scales in attempts to determine intra-urban risk areas.

The information pertaining to census tracts and blocks was extracted from the 1960 census and relate to the following properties: (1) total population; (2) condition and plumbing of the housing unit; (3) occupied housing unit; (4) color of the head of the household; (5) persons per room; (6) average number of rooms; (7) average value of the housing unit; and (8) average contract rent.

Condition of the Housing Unit (x_{101}, x_{102})

According to the 1960 Census, the condition of a housing unit is classified as being either sound, deteriorating or dilapidated.[3]

Sound housing is that which has no defects, or only slight defects which are normally corrected during the course of regular maintenance. Examples of slight defects are slight damage to porch or steps, small cracks in walls, plaster, or chimney, broken gutters or drainspouts, and slight wear on floors or doorsills.

Deteriorating housing needs more repairs than would be provided in the course of regular maintenance. It has one or more defects of an intermediate nature that must be corrected if the unit is to continue to provide safe and adequate shelter. Examples of intermediate defects include shaky or unsafe porch or steps, rotted window sills or frames, and broken or loose stair treads or missing baluster. Signs of neglect such as the above defects may lead to serious structural deterioration or damage if not corrected.

Dilapidated housing, finally, does not provide safe and healthy shelter. It has one or more critical defects, or has a combination of

---

[3] The U.S. Bureau of the Census, U.S. Census of Housing: 1960. Series HC(3)-262. City Blocks. Buffalo, N.Y. Table 2 pages 1-32 was used for variables x_{101} through x_{112}.

intermediate defects in sufficient number to require extensive repair or rebuilding, or is of inadequate original construction. Examples of critical defects include; holes, open cracks or missing materials over a large area of the floors, walls, roof, or other parts of the structure. Inadequate original construction includes structures built of makeshift materials and inadequately converted cellars, sheds, or garages not orginally intended as living quarters.

In the subsequent analysis two variables will be used to describe the condition of the housing unit. The first of these variables (x_{101}) will be operationally defined as the percentage of sound housing units per census tract and block. The second variable (x_{102}) will be operationally defined as the percentage of the combined number of deteriorating and dilapidated housing units per census tract and block.

Plumbing Facilities (x_{103})

In terms of plumbing categories the 1960 Census distinguishes between homes with all plumbing facilities and homes which lack some or all facilities.

The category "with all plumbing facilities" consists of units which have hot and cold piped water inside the structure, flush toilet and bathtub or shower inside the structure for the exclusive use of the occupants of the unit. Houses lacking hot water, toilet, bathtub or shower fall into the second category of plumbing facilities. Also included in this category are units whose occupants share toilet or bathing facilities with the occupants of another housing unit.

In the following analysis the variable related to plumbing facilities (x_{103}) will be operationally defined as the percentage of hous-

ing units per census tract and block which possess all plumbing facilities.

Occupancy (x_{104} through x_{112})

Again, according to the 1960 census the occupied housing unit may be classified as owner occupied (x_{104}) or renter occupied (x_{105}). The statistics for owner occupied units include information on the average value of the unit in dollars (x_{106}) and the average number of rooms per housing unit (x_{110}). Renter occupied units include data on the average contract rent of the unit in dollars (x_{107}) and the average number of rooms per renter occupied unit (x_{111}).

The average value of the occupied unit represents an estimate by the respondent of how much the property would bring on the market. The average rent is total rent payed excluding utilities or other extras. The average number of rooms is computed by totalling the number of rooms for each respective category and dividing by the appropriate number of units. Rooms included in the enumeration were whole rooms used for living purposes, such as living rooms, recreation rooms, bedrooms, dining rooms, kitchens, finished attic or basement rooms, lodgers' rooms, and rooms used for offices by a person living in the unit. Not included were bathrooms, halls, foyers or vestibules, closets, alcoves, pantries, etc.

Occupied housing units were classified as to the color of the head to the house hold (x_{108}) and this category will also be entered into the analysis as an independent variable. It should be noted that Mexicans were classed as non-white.

The final data set related to occupancy of the house provides

a measure of the number of persons per room (x_{109}) computed as the number of persons divided by the number of rooms in the unit. This means, of course, that values larger than 1.01 signify an average of more than one person per room. Also, the same statistic was calculated for number of persons per house (x_{112}).

Operationally, the variables related to the occupancy of the housing unit were defined as follows: x_{104} is equal to the percentage of owner-occupied housing units per census tract and block; x_{105} is equal to the percentage of renter-occupied housing units per census tract and block; x_{106} is equal to the average value of the housing unit in dollars per census tract and block; x_{107} is equal to the average monthly rent in dollars per census tract and block; x_{108} is equal to the percentage of non-white individuals per census tract and block; x_{109} is equal to the percentage of 1.01 or more persons per room per census tract and block; x_{110} is equal to the average number of rooms in the owner-occupied units per census tract and block; and x_{112} is equal to the average number of persons per house per census tract and block.

Summary

Each of the variables described about the census tract and block data represent the independent variables (x_{101} through x_{112}) in the regression and correlation models. The proposed relationship between these variables and leukemia are summarized in Table 2. It deserves notice that the chosen housing characteristics represent the two spatial scales of the census tract and the block. Data for tax districts can be aggregrated from the block data.

TABLE 2

HYPOTHESIZED RELATIONSHIP OF HOUSING VARIABLES TO LEUKEMIA

Factor	Variable	Exact definition of the variable	Hypothesized relationship with Y
Housing (X_1)	x_{101}	Percent of Housing units sound	Inverse
	x_{102}	Percent of Housing units Deteriorating	Direct
	x_{103}	Percent of Housing units with plumbing	Inverse
	x_{104}	Percent of Housing units owner-occupied	Inverse
	x_{105}	Percent of Housing units renter-occupied	Direct
	x_{106}	Average value of Housing units in dollars	Direct
	x_{107}	Average rent in dollars per month	Direct
	x_{108}	Percent of Non-white	Inverse
	x_{109}	Percent 1,01 or more persons per room	Direct
	x_{110}	Average no. of rooms owner-occupied units	Inverse
	x_{111}	Average no. of rooms renter-occupied units	Inverse
	x_{112}	Persons per house	Direct

Independent Variables: IB; Housing Unit

Data on the characteristics of the individual houses were obtained from the City Assessor's office in Buffalo, New York.[4] The records are divided into 14 enumeration districts from which the data were gathered. Each district corresponds to about five census tracts. The specific characteristics listed for each home are specified on the field sheet reproduced in Appendix 3.

The field sheets include the following information: (1) class of dwelling; (2) type of foundation; (3) type of basement; (4) type of construction; (5) type of exterior; (6) type of roof; (7) type of outside trim; (8) type of heating; (9) type of plumbing; (10) type of lighting; (11) repair; and (12) number of rooms.

The reason for investigating the influence of the listed variables is that our present knowledge is too meagre to allow a positive choice. This represents a detective approach where all lines of investigation are pursued in an effort to find clues as to the spatial variation of the disease.

The number of homes of diseased for 1959-1966 was 469. For several reasons, however, it was not possible to collect housing data for more than 418 of the 469 cases. If, for example, the house had been destroyed through urban renewal, then the field sheet had been destroyed by the assessor's office. In other instances, the information contained on the field sheets would include only the address, owner's name, and sometimes the size of the lot, while the balance of the field sheet

---

[4] Special thanks to L.W. Pauly, Principal Assessor, Buffalo, New York., for his assistance during this project.

would be blank. For these reasons, the data represent only 89 percent of the total cases. It should be noted, though, that this holds only for the individual housing characteristics and not for the block or the census tract data; in the latter instance the data represent 100 percent, leaving in mind, of course, that some may have been missed due to faulty diagnosis or simply failure of the case report to reach the Cancer Bureau.

To allow comparison between the characteristics of the houses of diseased and the homes of non-diseased, a sample was drawn from the same source as the leukemia housing data, i.e. from the city assessor's office. According to the 1960 Census, Buffalo had 177,224 housing units.[5] The drawing of a sample that would be representative of the population at the 95 percent level of confidence with a plus of minus 3 percent reliability, one would need a sample of 350 (Arkin and Colton,1968). However, since leukemia is contracted by the individual, the sample must be stratified according to the population distribution and not according to the spatial distribution of the housing units. This has been done by calculating the percentage of the total population from each district. The calculated values are reproduced in Table 3.

To explore the intuitive idea that the housing characteristics are associated with leukemia, it is necessary to know the length of time that the diseased had been living in the house. This is a complicating proposition, however, because the average length of time from first symptoms to death may vary from a few days for acute cases to many years

---

[5] U. S. Bureau of the Census, U.S. Census of Housing: 1960, Series HC(3)-262. City Blocks. Buffalo, New York. Table 1, p. 1.

TABLE 3

THE WEIGHED RELATIONSHIP BETWEEN THE POPULATION AND HOUSING UNIT SAMPLE

Tax District	Percent of total population in each tax district	Weighed value of Housing unit sampled in each tax district
1	3.3	12
2	9.8	35
3	6.6	23
4	8.4	30
5	6.6	23
6	6.7	24
7	7.5	27
8	6.6	23
9	8.3	29
10	12.2	43
11	5.7	20
12	7.4	26
13	2.5	9
14	8.4	30
Total	100.0	

for chronic cases. Nevertheless, it is necessary to verify the diseased person's residence prior to onset of symptoms. For example, if onset of illness was in 1959, the 1958 Polk city directory was used to determine if the diseased had lived in the same house prior to onset of the disease. This was checked for all cases and for at least one year before onset.

In the case of children under two years of age it was even more important to obtain this information then for adults; by doing so we are able to exclude the syngeristic effects which may result in persons of an older age or those persons who had lived in various homes while diseased. If the diseased individuals had not lived in the same house one year prior to onset of the disease they were not included in the analysis. On this basis, only five cases had to be excluded. The data on the characteristics of the housing units will not be analyzed via regression or correlation models but rather via Kolomorgorov-Smirnov tests.

Independent Variables: II; Demography (X_2)

The housing properties discussed in the two preceding sections of this chapter represent only one factor that can be related to the spatial aspects of leukemia. As to the general hypothesis that leukemia is transmitted by a virus, aspects of demography and the environment may provide more exact clues. For this reason, demographic factors will be discussed in the present section, while the next section will be devoted to a discussion of environmental factors. Most of the demographic data on diseased individuals were obtained from the Cancer Bureau in Albany. The information collected included: (1) age; (2)

sex; (3) birth date; (4) date of first symptom by month; and (5) occupation. Although the data from Albany occassionally gave the occupation of the deceased, it was frequently necessary to secure this information from the Polk city directory of Buffalo. Finally, information on median yearly earnings for each census tract was collected from the 1960 census.[6] Table 4 specifies the operational definitions of the chosen variables as well as the hypothesized relationships between the demographic variables and leukemia.

The demographic variables provide a description of the diseased individuals. However, some of these variables permit sub-categorization for further analyses. The relationship between leukemia and demographic characteristics will be discussed in chapter five. Part of the analysis will be via a regression model while other parts will be on the basis of tables.

Independent Variables: III; Environment (X_3)

The relationship between environmental characteristics and the disease are difficult to analyze. One reason for this is that most environmental data in a city are usually represented by weather stations located at discrete points. This means the data for continuous areas are obtained through interpolation. For this reason, great care has to be excerised in the establishment of relationships between the total point pattern of leukemia incidence and the areal distribution of environmental factors. Table 5 suggests possible relationships of air

---

[6] U. S. Bureau of the Census, U. S. Census of Housing: 1960. Series HC(1). Census Tracts. Buffalo, New York. Table P-1. pp. 16-21.

TABLE 4

HYPOTHESIZED RELATIONSHIP OF DEMOGRAPHIC VARIABLES TO LEUKEMIA

Factor	Exact Definition of the Variable	Hypothesized Relationship with Y
Demography (X_2)	Age Groups of the Deceased 14 and under 15-44 45-64 65 and over	Direct for Chronic cases; Inverse for Acute cases.
	Deceased Male/Female Ratio	Direct
	Yearly Median Income Groups per Census Tract <4000 4001-6000 6001-8000 >8000	Inverse
	Occupation of the Deceased	Direct with the amount of "people contact"

pollution and seasonal onset to leukemia.

TABLE 5

HYPOTHESIZED RELATIONSHIP OF ENVIRONMENTAL VARIABLES TO LEUKEMIA

Factor	Exact Definition of the Variable	Hypothesized Relationship with Y
Environment (X_3)	Average temperature per season	Inverse
	Air Pollution Micrograms of particulate per cubic metre	Direct

The environmental data obtained include: (1) average temperature per season; and (2) the amount of suspended particulates. The data on temperature were obtained from the U. S. Weather Bureau, Climatological Data for the yegrs 1959-1966 (U. S. Census,1959-1966), while the suspended particulate data were collected from Erie County Department of Health, Division of Air Pollution Control.[7]

There are three monitoring systems within the city of Buffalo. They are located at the Museum of Science at the center of the city, public school number 66 which is in the north-east area of the city, and public school number 26 situated in the south-east. The stations

[7] I would like to thank D. Tamol, Assistant Engineer, for his kind and generous assistance.

are along a straight north-south line.

Table 5 specifies the operational definitions of the employed environmental variables and their hypothesized relationship to leukemia incidence. The data will be analyzed in chapter six primarily on the basis of computed tables and maps.

CHAPTER IV

LEUKEMIA AND HOUSING

Introduction

> "on rare occasions leukemia occurs in two or more occupants of the same house. The significance of such "leukemia houses" is hard to assess. While they may represent chance events, it is possible that residence in certain houses somehow predisposes the occupants to leukemia."
>
> (Heath,1969.p.2021)

It has been hypothesized that various housing characteristics have been associated with the intra-urban variation of leukemia incidence in Buffalo. Specifically, it was suggested an analysis of housing characteristics might indicate whether leukemia is an infectious disease. This idea will now be tested via regression models in which the various housing characteristics will be entered as independent variables. The dependent varaible will be the rate of leukemia incidence per areal unit. In the course of this testing procedure, the scale of investigation will be changed in an effort to isolate the most interpretable spatial pattern; as recalled from chapter one, the concept of scale and scale change is of prime importance in any theory about the etiology of a disease.

Housing - Health and/or Disease

It is well known that the total environment has a predominant

influence in determining the physical, mental and social well-being of man. Housing is a part of this total environment and therefore to some extent responsible for the status of man's health and well-being (Nelson, 1945; Wilner,1959; Bernstein,1957). Still, it has been difficult to delineate clearly and concisely the magnitude of this influence.

The most conclusive relationships between housing and disease were established by Wilner, who found that poor housing is positively correlated with poor health (Wilner,1962). In the case of tuberculosis, for instance, he noted a positive relationship between the degree of crowding in a home and the disease incidence. Other diseases such as plague, typhus, and malaria may be transmitted in the house via indirect modes like rats, mice, and mosquitoes. Generally, the houses in which this type of infestation occurs are poverty stricken or located in slum areas. Conversely, M'Gonigle found higher death rates among a number of publicly rehouses families than in a group of similar slum-dwelling families (M'Gonigle,1933).

Already these selected references suggest that housing characteristics may play a significant role in the health or disease of an area. Hence, the spatial aspects of housing in an intra-urban area are important to investigate. However, there are several methodological difficulties inherent in such investigations, particularly in separating statistical association from real causation;[8] established associations can give clues and form the basis of a detective approach, while the establishment of causation requires the isolation of true and invariant

---

[8] The problems inherent with association and causation have been discussed particularly by Blalock, H.M. Causal Inferences in Non-Experimental Research and by Brown, R. Explanation in Social Science.

cause and effect relationships.

Many other reasons account for the difficult problems of establishing the relationship of housing to disease and/or health. For instance, it is difficult to determine what spatial unit of analysis is most meaningful to employ.

A related difficulty is that the housing unit is only part of the total environmental influence. Each aspect of the house may influence health or disease in a separate manner. For example, poor plumbing facilities, poor construction, and insect infested wood may all produce viral or bacterial infections. It is even possible that each of these aspects may produce different diseases. Thus, the residential environment may affect man's physical health in at least one of three ways: (1) it may concern the transmission of communicable diseases; (2) it may not fulfill man's physiological needs; and (3) it may cause injury (W.H.O. 1961). It is evident that those who believe in the viral etiology hypothesis of leukemia are concerned only with the first of these three affects.

Leukemia and Housing

Diseased leukemia patients obviously lived in housing units of some kind. It is possible, therefore, that the housing unit may be a factor which enhanced the transmission of the disease. Furthermore, if multiple cases of leukemia were found in unrelated individuals living in the same housing unit, then the hypothesis of a relationship between leukemia and housing would be strengthened. Likewise, if leukemia occurred in victims who had lived next door to each other, the same hypothesis would be substantiated. Given the stated etiology hypothesis

it is interesting to note that Heath (McPhedran and Heath,1969) has drawn attention to a case where four persons associated with the same house had contracted leukemia over the eighteen year period between 1950-1967. The house is a single dwelling constructed of concrete blocks and wood and located in a residential section near the center of town. It has three bedrooms, a living room, kitchen and a bath. It has city water and is heated by space heaters. As Heath points out "dwellings such as the one described may well be logical places to look for possible associations between the house and leukemia" (Heath, 1969.p.2025).

In addition to Heath's study, three cases of Hodgkin's disease and three cases of leukemia have been reported in two adjoining houses (A and B) in a town in south central Pennsylvania (Gilmore and Zelesnick, 1962). House A was a brick house and house B was a frame house. Both houses have two stories, an attic and a basement. House B had hot air from a coal furnace but switched to gas hot air. House A was divided into two units which had heat by coal hot air in one unit (this unit developed the leukemia case) while the other was heated by gas hot air. Both had excellent sanitary facilities. Gilmore concluded that "environmental aspects appear to be the most logical explanation" (Gilmore, 1962.p.1049).

These two studies point up a major area for research which can be specified in the question: "Are the housing units in which leukemia occur any different from those homes of a non-diseased population?" Since it cannot be determined from two studies of three homes if there are common factors to each house, a study of leukemia patients and their homes will now be undertaken. The study will also attempt to

determine the proper spatial scale for the identification of risk areas.

The Analyses of Housing with Various Scales

In the remainder of this chapter the relationships of leukemia incidence and several housing variables will be analyzed. It can be expected a priori that these relationships will vary depending on the scale of investigation. For instance, a variable may show significant relationships on one scale but not on another scale. The four scales to be investigated are: (1) the census tract; (2) the block; (3) the tax district; and (4) the individual housing unit.

The Census Tract Scale

The census tract analysis uses data from 75 different tracts. Table 6 reproduces the simple correlation matrix, while Table 7 gives the results of the simple and multiple regression and correlation analyses.

The tables indicate that all nine variables are extremely weakly correlated with the leukemia incidence rates even though four of the variables are statistically significant. Variable x_{107}, average rent has a correlation coeffieient of .3792 in the simple equation, while it is .3318 in the multiple case. Both of these values are significant at the .01 level. This correlation suggests two interpretations namely either that the rent paid by an individual is a substitute measure for his socio-economic status or that the apartment is in good condition. In order to choose between these two interpretations it is necessary to discuss the influence of other variables. Thus, it should be noted that

TABLE 6

CENSUS TRACT SCALE
CORRELATION MATRIX

Variable	Variable Number	Y	(x_{101})	(x_{103})	(x_{104})	(x_{106})	(x_{107})	(x_{108})	(x_{109})	(x_{110})	(x_{111})
Incidence Rate of Leukemia	Y	1.0000									
Percent of Housing units sound	x_{101}	.2177	1.0000								
Percent of Housing units with plumbing	x_{103}	.1346	.6374	1.0000							
Percent of Housing units owner-occupied	x_{104}	.0684	.5273	.7506	1.0000						
Average Value of the Housing unit	x_{106}	.3557	.6535	.3130	.1865	1.0000					
Average Rent	x_{107}	.3792	.6343	.4144	.2578	.7485	1.0000				
Percent Non-white	x_{108}	-.0555	-.4090	-.2693	-.4678	-.2026	-.4829	1.0000			
Percent of 1.01 or more persons per room	x_{109}	-.1853	-.6248	-.5398	-.5259	-.5114	-.3792	.5616	1.0000		
Average No. of rooms owner-occupied units	x_{110}	.2400	.8791	.3162	.0980	.5943	.5222	-.0671	-.5024	1.0000	
Average No. of rooms renter-occupied units	x_{111}	.0993	.6634	.5894	-.0670	-.1182	-.0860	-.2568	.2239	.0939	1.0000

TABLE 7

CENSUS TRACT SCALE
CORRELATION AND REGRESSION ANALYSIS

Variable	Variable Number	Simple Regression (b)	Simple Regression (r)	Multiple Regression (b)	Multiple Regression (r)
Percent of Housing units sound	x_{101}	.0696	.2177	.0135	.0203
Percent of Housing units with plumbing	x_{103}	.1346	.1356	-.2047	-.1758
Percent of Housing units owner-occupied	x_{104}	.0222	.0684	-.0671	-.1297
Average value of the Housing unit	x_{106}	.4105	.3557*	.2106	.1004
Average Rent	x_{107}	.1519	.3792*	.2320	.3318*
Percent Non-white	x_{108}	-.0105	-.0555	-.0188	-.0594
Percent of 1.01 or more persons per room	x_{109}	-.2542	-.1853	-.0034	-.0016
Average No. of rooms of owner-occupied units	x_{110}	2.1503	.2400**	-1.5553	-.1072
Average No. of rooms of renter-occupied units	x_{111}	.0801	.0993	.4761	.3320*

*significant at .01 level ** significant at .05 level

the simple equation shows both average value of the home and average number of rooms in the owner-occupied units to be positively and significantly realted to the dependent variable.

The analysis indicates that an increase in the number of rooms per unit is associated with an increased incidence of leukemia. This can be interpreted as contradicting the hypothesis that a viral disease is best transmitted in a crowded home or under crowded conditions. One may note that variable x_{109}, which indicates the degree of crowding in the house, is not significantly correlated with the incidence of leukemia although it has a negative regression coefficient. Furthermore, the positive sign of the partial correlation coefficient for variable x_{111}, the average number of rooms in renter-occupied units, suggests that increased leukemia rates are associated with more rooms per unit.

To achieve a standard for comparison of five of the variables (x_{106}, x_{107}, x_{109}, x_{110}, x_{111}) a new analysis was run in which the "beta weights" were used. The results are given in Table 9, which indicates that none of the variables were significantly related to leukemia incidence. Table 8 shows the correlation matrix of the tested standardized variables. The multiple correlation coefficients of both estimating equations are significant at the .05 level. The original nine variables have a multiple correlation coefficient of .4984 while the standardized form of the data has a multiple correlation coefficient of .4284. It may therefore be concluded that the analysis of the census tract data did not provide substantive results. It is apparent that the results from the census tract scale do not explain the spatial pattern of leukemia in Buffalo. For this reason, the scale of analysis will now be changed to the block level to investigate whether a more detailed

TABLE 8

CENSUS TRACT SCALE
STANDARDIZED VARIABLES
CORRELATION MATRIX

Variable	Variable Number	(Y)	(x_{106})	(x_{107})	(x_{109})	(x_{110})	(x_{111})
Incicence Rate of Leukemia	Y	1.0000					
Average Value of the housing unit	x_{106}	.3558	1.0000				
Average rent	x_{107}	.3793	.7486	1.0000			
Percent of 1.01 or more persons per room	x_{109}	-.1853	-.5113	-.4829	1.0000		
Average No. of rooms of owner-occupied units	x_{110}	.2400	.5943	.5222	-.5024	1.0000	
Average No. of rooms of renter-occupied units	x_{111}	.0993	-.0670	-.1182	-.2568	.2239	1.0000

TABLE 9

CENSUS TRACT SCALE
STANDARDIZED VARIABLES
CORRELATION AND REGRESSION ANALYSIS

Variable	Variable Number	(b)	(r)
Average value of the housing unit	x_{106}	.2063	.1366
Average rent	x_{107}	-.3150	-.2162
Percent of 1.01 or more persons per room	x_{109}	.1007	.0863
Average no. of rooms of owner-occupied units	x_{110}	-.0379	-.0308
Average no. of rooms of renter-occupied units	x_{111}	.1847	.1784

micro-level approach will porduce more conclusive results.

The Block Scale

The block scale provides a more detailed spatial analysis of the relationships between housing and leukemia. To allow comparison, however, the procedure used for the census tract scale will also be used in this instance. The only exception is the introduction of three new variables, x_{102} percent of housing units deteriorating, x_{105} percent of housing units renter-occupied, and x_{112} persons per house.

Table 10 gives the correlation matrix while Table 11 presents the regression and correlation analyses. Three of the variables are

TABLE 10

BLOCK SCALE
CORRELATION MATRIX

Variable	Variable Number	(Y)	(x_{101})	(x_{102})	(x_{103})	(x_{104})	(x_{105})	(x_{106})	(x_{107})	(x_{108})	(x_{109})	(x_{110})	(x_{111})	(x_{112})
Incidence Rate of Leukemia	Y	1.0000												
Percent of Housing units sound	x_{101}	.1363	1.0000											
Percent of Housing units deteriorating	x_{102}	.1295	.5441	1.0000										
Percent of Housing units with plumbing	x_{103}	-.0365	.6682	.2152	1.0000									
Percent of Housing units owner-occupied	x_{104}	.1517	.7458	.2804	.6957	1.0000								
Percent of Housing units renter-occupied	x_{105}	-.0565	.1924	.2281	.3197	.0387	1.0000							
Average value of the Housing unit	x_{106}	-.0325	.4769	.0707	.3871	.4819	.0261	1.0000						
Average rent	x_{107}	-.1147	.0487	.0230	.0166	-.0701	.4458	.0408	1.0000					
Percent of Non-white	x_{108}	-.0004	-.1085	.1726	-.0359	-.0723	.3038	-.1692	.5209	1.0000				
Percent of 1.01 or more persons per room	x_{109}	-.0565	-.1005	.1303	-.0271	-.1095	.2151	-.1958	.3409	.4829	1.0000			
Average no. of rooms of owner-occupied units	x_{110}	-.3867	.1099	-.0379	.1073	.1517	-.1649	.5488	.0714	-.1089	-.1841	1.0000		
Average no. of rooms of renter-occupied units	x_{111}	-.3290	.0078	.0136	.0837	-.0460	.0648	-.1312	.1601	.0652	.0805	.2110	1.0000	
Persons per house	x_{112}	-.1161	-.0383	-.0025	.0359	-.1526	-.0377	-.0014	-.1525	.2696	.0435	.1508	-.2158	1.0000

TABLE 11

CORRELATION AND REGRESSION ANALYSIS
BLOCK SCALE

Variables	Variable Number	Simple Regression (b)	Simple Regression (r)	Multiple Regression (b)	Multiple Regression (r)
Percent of Housing units sound	x_{101}	.1189	.1363*	.0972	.0649
Percent of housing units deteriorating	x_{102}	.1205	.1295**	.0498	.0486
Percent of housing units with plumbing	x_{103}	-.0307	-.0365	-.1949	-.1470*
Percent of housing units owner-occupied	x_{104}	.1781	.1517*	.2163	.1213**
Percent of housing units renter-occupied	x_{105}	-.0504	-.0565	-.0966	-.1042**
Average value of the housing unit	x_{106}	-.0220	-.0325	.0992	.1141**
Average rent	x_{107}	-.3370	-.1147**	-.0647	-.0544
Percent of Non-white	x_{108}	.0118	.0004	.1187	.1043**
Percent 1.01 of more persons per room	x_{109}	-.2577	-.0565	-.4457	-.0868
Average no. of rooms of owner-occupied units	x_{110}	-12.3658	-.3867*	-13.9873	-.3791*
Average no. of rooms of renter-occupied units	x_{111}	-12.7398	-.3290*	-6.2874	-.1809*
Persons per house	x_{112}	-11.3445	-.1161**	.8525	.0103

*Significant at .01 level **Significant at .05 level

significant in both the simple and multiple equations. The first of these, percent of housing units owner-occupied, has a correlation coefficients of .1512 and .1213 in the simple and the multiple correlation analysis respectively. The regression coefficient is positive, suggesting that increased rates of leukemia are directly associated with increases in home ownership. However, the correlation coefficients are extremely low.

The second variable to be significantly related to leukemia incidence is the average number of rooms in both the owner and renter-occupied units. The negative regression coefficient is in direct contrast to the results obtained from the same two variables in the census tract scale indicating that increased rates of leukemia are associated with fewer rooms in the unit and not with more rooms as in the census tract analysis. Thus, on the block scale the hypothesis that leukemia is an infectious disease more likely to be transmitted in crowded conditions is supported. At the same time, the negative regression coefficient associated with variable x_{112} (persons per house) suggests that increased rates of leukemia occur when there are fewer people in the home.

The importance of the remaining variables fluctuate depending on the equation cited. For example, average rent is significant at the .01 level in the simple analysis but not significant at all in the multiple analysis. It is important to note that the results from the block scale frequently are in direct contrast to those obtained from the census tract scale.

The results of the analysis of the standardized variables are shown in Tables 12 and 13. Table 13 gives the results of the six stand-

TABLE 12

BLOCK SCALE
STANDARDIZED VARIABLES
CORRELATION MATRIX

Variable	Variable Number	(Y)	(x_{106})	(x_{107})	(x_{109})	(x_{110})	(x_{111})	(x_{112})
Incidence Rate of Leukemia	Y	1.0000						
Average value of the housing unit	x_{106}	-.0799	1.0000					
Average rent	x_{107}	-.3220	.2324	1.0000				
Percent of 1.01 or more persons per room	x_{109}	-.0307	-.3402	-.1396	1.0000			
Average no. of rooms of owner-occupied units	x_{110}	-.4022	.5786	.2392	-.1859	1.0000		
Average no. of rooms of renter-occupied units	x_{111}	-.3574	-.1589	.5256	.0817	.1187	1.0000	
Persons per house	x_{112}	.0172	-.0088	-.1756	.4547	.1572	.2324	1.0000

TABLE 13

BLOCK SCALE
STANDARDIZED VARIABLES
CORRELATION AND REGRESSION ANALYSIS

Variable	Variable Number	(b)	(r)
Average value of the housing unit	x_{106}	.1157	.0985
Average rent	x_{107}	-.0664	-.0576
Percent of 1.01 or more persons per room	x_{109}	-.1712	-.1663*
Average no. of rooms of owner-occupied units	x_{110}	-.4870	-.4162*
Average no. of rooms of renter-occupied units	x_{111}	-.2851	-.2381*
Persons per house	x_{112}	.2273	.2064*

* Significant at the .01 level.

ardized variables as obtained from the multiple regression correlation analysis. Average rent and average value of the home are not significantly related to the dependent variable. However, as in the original data, both average number of rooms in the owner and renter-occupied units are significantly related at the .01 level. It is also noticable that the degree of relationship has remained approximately the same as in the analysis of the original data. Again, the regression coefficients are negative.

Variables x_{109} and x_{112} which measure crowding in a home are

both significant at the .01 level. Variable x_{109}, percent 1.01 or more persons per room, has a negative regression coefficient and an extremely low correlation coefficient. On the other hand, variable x_{112}, persons per house, is positivley related to the dependent variable. The multiple coefficient of correlation for the original data is .5583 and for the standardized data it is .5649. Both are significant at the .01 level. It is apparent from the census tract and block scale that the selected standardized variables (x_{106}, x_{107}, x_{109}, x_{110}, x_{111}, x_{112}) account for most of the explanation obtained by the estimating equation. Thus, the forthcoming analysis will utilize only six standardized variables.

The interrelationships of the standardized variables are shown in Table 12. It is evident from the table that the correlations between the variables are rather weak. This may also be said of the entire analysis up to this point. Having evaluated the census tract and block scale, with no apparent substantive results, except for the average number of rooms in the owner and renter units, the analysis will therefore have to turn to still another scale represented by the tax districts.

An Aggregated Scale (Tax Districts)

At this point of the analysis the scale can be changed in two ways. Either one may analyze each home of the diseased individual or one may aggregate the data. Both approaches will be used but the latter will be discussed first.

The data obatined for the individuals can easily be aggregated into 14 tax districts shown in Figure 4. The tax districts contain approximately five census tracts each. Again, all cases that occurred in a block within a particular district are aggregated and tested via

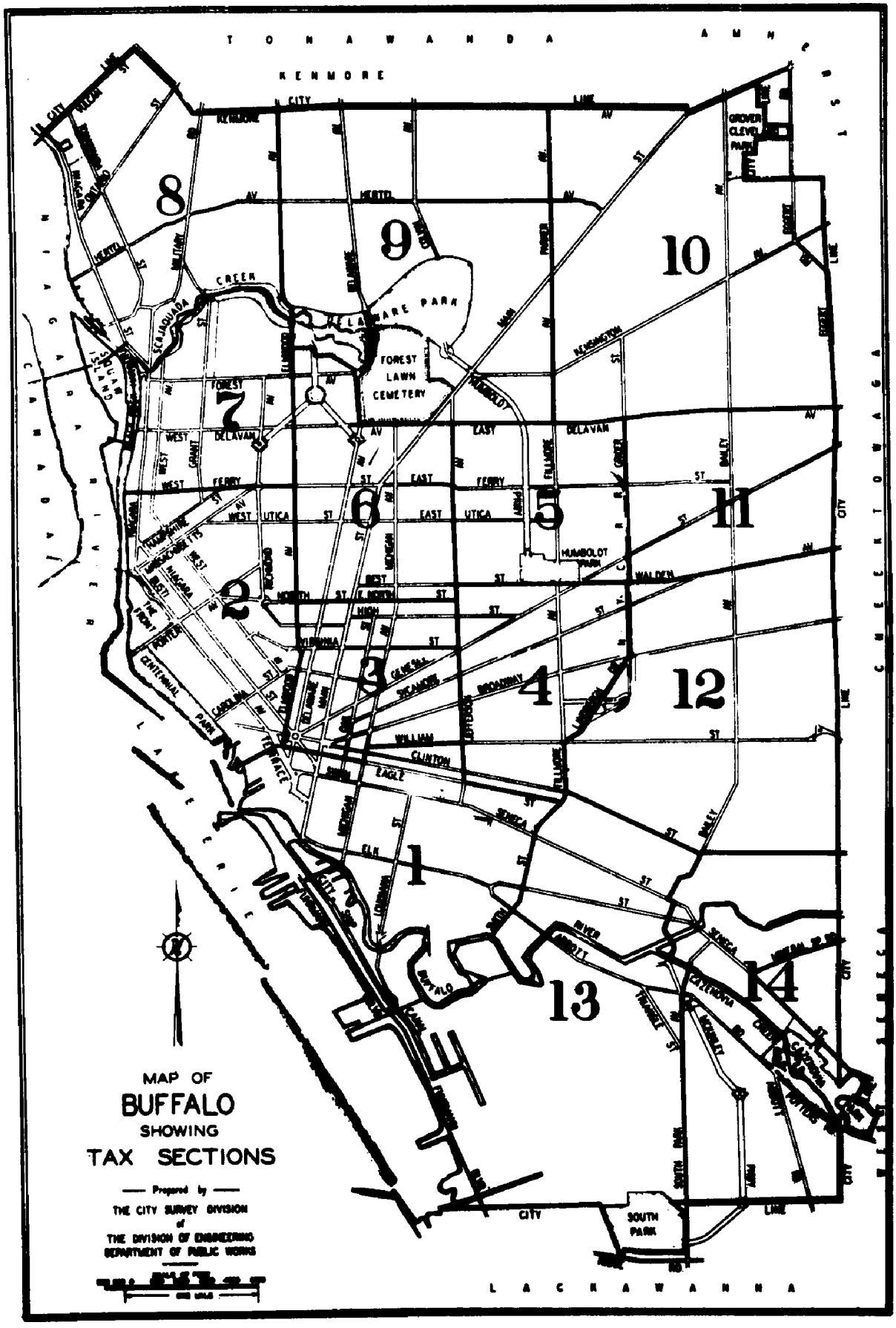

T O N A W A N D A
K E N M O R E
L A K E E R I E
N I A G A R A R I V E R
C A N A D A
L A C K A W A N N A
C H E E K T O W A G A
8
9
10
7
6
5
11
2
3
4
12
1
13
14
FOREST LAWN CEMETERY
DELAWARE PARK
HUMBOLDT PARK
SOUTH PARK
GROVER CLEVELAND PARK
MAP OF
BUFFALO
SHOWING
TAX SECTIONS
— Prepared by —
THE CITY SURVEY DIVISION
of
THE DIVISION OF ENGINEERING
DEPARTMENT OF PUBLIC WORKS

Figure 4

regression correlation analysis.

The variables analyzed are only the six that were standardized and that the previous analyses have indicated to be most important.

Table 14 gives the maximum partial correlation coefficients for the six variables of the 14 tax dsitricts. Of the 14 areas, the multiple correlation coefficient is significant for seven of the areas. The significance level is .01 for all seven. As shown in the table the correaltion coefficients are high, ranging from .9827 for district four to .6914 for district ten. Of the variables tested this indicates that they explain the incidence of leukemia in those areas with some degree of reliability even though the small number of observations entered into the analysis in some of the areas causes serious problems. Figure 5 shows the resulting high risk areas within the city of Buffalo. One may note there is almost a concentric zone around the core of spatially high risk areas. These areas may be considered low risk areas.

More detailed examination of the variables indicate that variable x_{106}, the average value of the home, shows no significant relationship in any of the 14 districts. Variable x_{110}, the average number of rooms in the owner-occupied unit is significant in five of the seven spatially high risk areas. The level of significance for each may be seen in Table 14. The regression analysis suggests that increased rates of leukemia are associated with a decreased number of rooms in the owner-occupied units. This is very suggestive of a crowding condition and may support the theory of a viral etiology.

Variable x_{111}, average number of rooms in a renter-occupied unit, is also significantly related to the dependent variable in four of the seven high risk areas. The relationship is negative as for variable

TABLE 14

TAX DISTRICTS
MAXIMUM PARTIAL CORRELATION COEFFICIENTS

Variables	Variable Number	Districts (1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Average value of the housing unit	x_{106}	-.8743	.2960	-.1195	-.2219	.1504	.1982	.0088	-.3824
Average rent	x_{107}	-.8629	-.4386**	-.3998**	-.4797	.1900	.3992**	-.1169	.6065
Percent of 1.01 or more persons per room	x_{109}	.7622	-.4354**	-.1018	-.0448	-.1670	.3883**	-.1556	-.3234
Average no. of rooms of owner-occupied units	x_{110}	.4820	-.3308	-.4662**	-.1976	-.2344	-.5364*	.0298	-.7891*
Average no. of rooms of renter-occupied units	x_{111}	.8403	-.8042*	.4524**	-.8543*	-.0767	-.3755	-.3049	-.5561**
Persons per house	x_{112}	-.6584	.8215*	-.2184	.7394*	-.3454	.2503	.0565	.6969*
Multiple Correlation Coefficients		.9450	.8691*	.8248*	.9827*	.6185	.8740*	.7164	.9691*

*Significant at the .01 level ** Significant at the .05 level

TABLE 14—continued

TAX DISTRICTS
MAXIMUM PARTIAL CORRELATION COEFFICIENTS

Variables	Variable Number	Districts (9)	(10)	(11)	(12)	(13)	(14)
Average value of the housing unit	x_{106}	.2866	.2374	.3058	.2839	-.8728	.0819
Average rent	x_{107}	-.2072	-.2157	-.0067	-.3025	-.5734	-.0831
Percent of 1.01 or more persons per room	x_{109}	.2228	-.1870	-.2252	-.2919	-.6657	-.3017
Average no. of rooms of owner-occupied units	x_{110}	.2149	-.2786*	.1422	-.7581*	-.8278	-.4167**
Average no. of rooms of renter-occupied units	x_{111}	.0444	-.1335	-.4091	-.1280	-.8636	-.3565
Persons per house	x_{112}	.0035	-.1462	.2252	.1962	-.6153	.2472
Multiple Correlation Coefficients		.6403	.6904*	.7455	.9430	.9982	.7754*

*Significant at the .01 level **Significant at the .05 level

LEUKEMIA INTRA-URBAN RISK AREAS

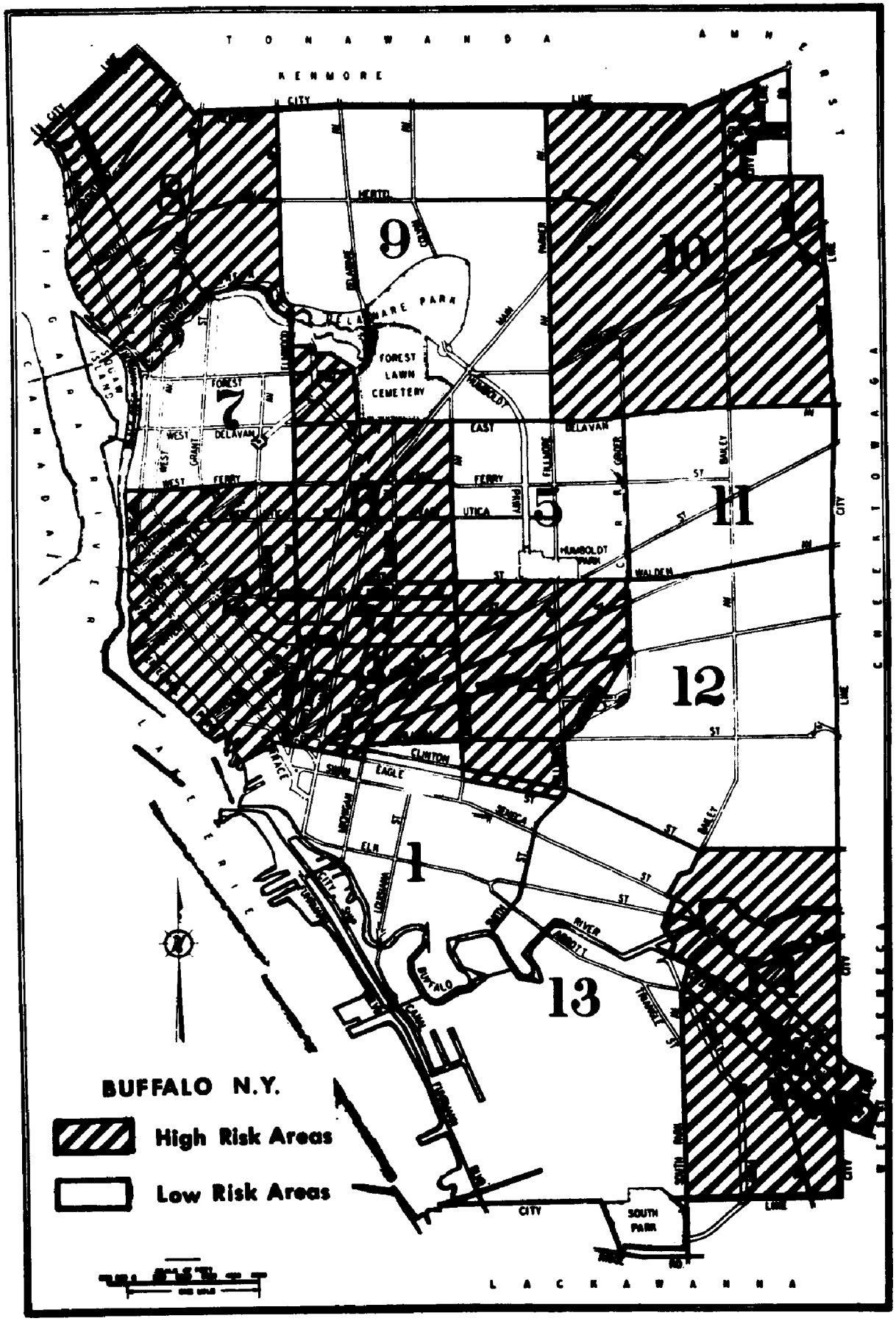


Figure 5

x_{110}, average number of rooms in the owner-occupied units. Again, increased rates of leukemia are associated with decreased number of rooms in the renter-occupied units. This further supports the viral etiology hypothesis.

Table 14 further specifies that variable x_{109}, the percent of 1.01 or more persons per room, is significant in only two to the spatially high risk areas. Variable x_{112}, average number of persons per house is positively related to leukemia with correlation coefficients equal to .8215, .7394 and .6969 in the spatially high risk areas of two, four and eight. The relationship in this instance indicates that increased rates of leukemia are associated with more peronse in the house. This is a definite condition related to crowding. Moreover, considering variables x_{110} and x_{111}, (which are the average number of rooms in the owner and renter-occupied units) in conjunction with variable x_{112}, one may further justified in suggesting that crowding is associated with leukemia.

Variable x_{107}, average rent, is significantly negatively related to the incidence of leukemia in four of the seven areas. However, in areas six and eight the relationship is positive.

Of the three scales tested, the aggregation of the block data into tax districts seems most relevant. Using this scale it was possible to identify spatially high and low risk areas. However, further associations of housing and leukemia might possibly be found in the homes of the diseased persons's. A specific analysis of each diseased person's home may therefore result in the recognition of other associations.

The Micro Scale (The Housing Unit)

In an effort to determine if the characteristics of the houses

are associated with the diseased individuals, a random sample of homes that did not have leukemia individuals were selected. Each classification of the data will be tested by application of the Kolmogorov-Smirnov test, which measures the maximum difference between two cumulative distributions.

Table 15 gives the results of the Kolmogorov-Smirnov test for the class of the dwelling.[9] It may be noted that a maximum difference occurs in the apartment class. The maximum difference of .132 is large enough to be significant at the .05 level. This significance level indicates that there is indeed a difference in the type of dwelling inhabited by the leukemia individuals. The difference is attributed to more leukemia victims living in apartments than in other types of dwellings.

Table 16 gives the test results on the type of foundation on which a home is built. The cumulative distributions of the two samples do not have a difference large enough to be significant. In fact, Tables 16 through 25 indicate that for none of the employed categories is there a difference between the housing units of the leukemia and non-leukemia individuals.

However, three additional variables are tested for variations via a difference of porportions test. These variables are: (1) the existence of fireplaces; (2) availabilty of laundry facilities in the home; and (3) average number of rooms in the home. The frequency of laundry

---

[9] The Kolmogorov-Smirnov test and the Difference of Proportions test were used to assess the statistical significance of the various characteristics of the housing unit. Both tests are outlined in Blalock, H. M. Social Statistics, McGraw-Hill, New York. 1960.

TABLE 15

CLASS OF DWELLINGS
SMIRNOV TEST

Class of Dwelling	Leukemia	Non-Leukemia	Difference
Hotel	.005	.000	.005
Duplex	.041	.011	.030
Apartment	.160	.028	.132*
Single	.576	.621	.039
Double	1.000	1.000	.000

* Maximum Difference (Significant at the .05 level)

TABLE 16

TYPE OF FOUNDATION
SMIRNOV TEST

Type of Foundation	Leukemia	Non-Leukemia	Difference
Brick	.019	.006	.013
Posts	.044	.078	-.034
Concrete	.379	.468	-.089*
Stone	1.000	1.000	.000

* Maximum Difference (would have to be at least .409 for .05 significance level)

TABLE 17

TYPE OF BASEMENT IN THE HOME
SMIRNOV TEST

Type of Basement	Leukemia	Non-Leukemia	Difference
No Basement	.037	.126	-.089
Quarter Basement	.080	.166	-.086
Half Basement	.139	.233	-.094*
Full Basement	1.000	1.000	.000

*Maximum Difference (would have to be at least .126 for .05 significance level)

TABLE 18

TYPE OF CONSTRUCTION OF THE HOME
SMIRNOV TEST

Tile	.003	.009	-.006
Concrete Stone	.011	.014	-.003
Brick	.137	.091	.046*
Frame	1.000	1.000	.000

*Maximum Difference (would have to be at least .126 for .05 significace level)

TABLE 19

TYPE OF EXTERIOR CONSTRUCTION ON THE HOME
SMIRNOV TEST

Exterior Construction	Leukemia	Non-Leukemia	Difference
Paper	.006	.000	.006
Pressed Brick	.015	.021	-.006
Stucco	.044	.043	.001
Combination Brick	.120	.055	.065*
Shingles	.257	.268	.011
Siding	1.000	1.000	.000

*Maximum Difference (would have to be at least .126 for .05 significance level)

TABLE 20

TYPE OF ROOF ON THE HOME
SMIRNOV TEST

Roof Type	Leukemia	Non-Leukemia	Difference
Board	.000	.000	.000
Iron	.0029	.003	.0001
Asbestos	.0089	.0091	-.0002
Tile	.027	.012	.015
Gravel	.050	.015	.035*
Slate	.085	.054	.031
Prepared Shingle	.414	.402	.012
Shingle	1.000	1.000	.000

*Maximum Difference (would have to be at least .126 for .05 significance level)

TABLE 21

ROOF CONSTRUCTION
SMIRNOV TEST

Roof Construction	Leukemia	Non-Leukemia	Difference
Gambrel	.0059	.003	.0029
Dormers	.0298	.0555	-.0257*
Flat	.065	.067	-.002
Hip	.134	.148	-.014
Gable	1.000	1.000	.000

*Maximum Difference (would have to be at least .126 for .05 significance level)

TABLE 22

TYPE OF OUTSIDE TRIM
SMIRNOV TEST

Outside Trim	Leukemia	Non-Leukemia	Difference
Galvanized Iron	.003	.003	.000
Terra Cotta	.010	.003	.007
Concrete	.020	.003	.017
Stone	.055	.020	.035*
Wood	1.000	1.000	.000

*Maximum Difference (would have to be at least .135 for .05 significance level)

TABLE 23

TYPE OF HEATING
SMIRNOV TEST

Type of Heating	Leukemia	Non-Leukemia	Difference
Central Station	.000	.000	.000
Steam	.040	.015	.025
Hot Water	.173	.131	.042*
Stove	.431	.408	.023
Furnace	1.000	1.000	.000

*Maximum Difference (would have to be at least .125 for .05 significance level)

TABLE 24

SOURCE OF HOUSEHOLD POWER
SMIRNOV TEST

Household Power	Leukemia	Non-Leukemia	Difference
Electric	.387	.432	.045*
Gas & Electric	1.000	1.000	.000

*Maximum Difference (would have to be at least .130 for .05 significance level)

TABLE 25

CONDITION OF HOUSING UNIT
SMIRNOV TEST

Condition	Leukemia	Non-Leukemia	Difference
Poor	.038	.058	.020
Fair	.239	.320	.090*
Good	1.000	1.000	.000

*Maximum Difference (would have to be at least .139 for .05 significance level)

facilities in the homes with leukemia is .6141 and the frequency in the homes without leukemia is .6081. The difference of proportion test yields a result of .157 which when entered into a normal curve does not indicate a signifciant difference between the two distributions.

The frequency of fireplaces was tested by the same method for the homes in which leukemia occurred, the fireplace frequency is .42847 and while the homes of the non-leukemia individuals had a frequency of .32012. The difference of proportions test yielded a result of 2.923 which gives a significant difference. In a two-tailed test the probability of obtaining a value as large or larger than 2.923 is .0032. This is significant at the .01 level which means that there is a difference between the two distributions.

Finally, the average number of rooms were found to be significantly different between the two samples. The frequency of the average number of rooms for the leukemia homes is .49492 while for the homes of non-leukemia individuals the frequency is .63271. The difference of proportions test yielded a result of 4.675. In a two-tailed test the probability of obtaining a value as large or larger than 4.675 is .00008. This is significant at the .001 level which means that there is a difference between the two distributions.

Summary

This chapter has presented material on housing variables for four different scales. There were only a few of the variables and only at one scale that substantive results were obtained. The observed relationships of the variables are presented in Table 26.

The census tract scale suggested possible relationships but the correlations were very weak. Average rent, average value of the home,

TABLE 26

OBSERVED RELATIONSHIP OF THE TESTED VARIABLES
LEUKEMIA AND HOUSING

Exact Definition of the Variable	Hypothesized relationship with Y	Observed relationship with Y
Percent of Housing units sound	Inverse	Direct
Percent of Housing units deteriorating	Direct	Direct
Percent of Housing units with plumbing	Inverse	Inverse
Percent of Housing units owner-occupied	Inverse	Direct
Percent of Housing units renter-occupied	Direct	Inverse
Average value of housing units in dollars	Direct	Direct
Average rent in dollars per month	Direct	Inverse
Percent of non-white	Inverse	Inverse
Percent 1.01 or more persons per room	Direct	Inverse
Average no. of rooms owner-occupied units	Inverse	Inverse
Average no. of rooms renter-occupied units	Inverse	Inverse
Persons per house	Direct	Direct

and average number of rooms in the owner and renter-occupied units were significant at the .01 level. However, the correlation coefficients were very low and none exceeded .4500.

The block scale produced similarly poor results. Thus, there were significant correlations of several variables but the correlation coefficients were well below accepted standards. Interestingly, the average number of rooms in the owner and renter-occupied units were significant as in the previous census tract analysis, but the direction of the relationship had changed from positive to negative. The standardized variables for the block scale also proved to be significant but the correlations were low. For the census tract scale the relationship to the number of rooms was positive, while for the block scale the reverse was the case. Comparing the two scales, of the census tract and the block it may be concluded that both fail to estimate the incidence with any relaibility.

In contrast to the census tract and block analysis, the results obtained from the aggregation of block data into tax districts were remarkable. In particular, it was possilbe to identify spatially high and low risk areas. The multiple correlation coefficients were high varying from .6914 to .9827. The variables which were significant most times were average number of rooms in the owner and renter-occupied units. The relationship was usually negative.

Characteristics of individual houses provided the data for a more detailed micro-analysis. Several aspects were tested but failure to reject the null hypothesis prevailed. Only the presence of fireplaces in the house and the number of rooms in the home were found to be significantly different between houses with leukemia and without leukemia

victims. Indeed, increased rates of leukemia existed where there were fewer rooms in the home. As to the fireplaces, it is possible the socio-economic status is higher in the individuals who get leukemia and this may reflect the presence of fireplaces in the leukemia homes. If this were the case, however, one would expect more rooms in a home and more than likely less crowding. On the other hand, a fireplace is in fact an open hole in the home and it is a possible route for various vectors. Also, burning of certain products in the fireplace may produce by-products which could cause leukemia. Obviously, the relationships between leukemia and fireplaces is open to wide conjecture.

The significant finding of this chapter is the identification of spatially high and low risk areas in the intra-urban setting of Buffalo. It is evident that causation has not been established but association was definitley determined. At the same time, housing is only of several factors that were intuitively suggested as being related to leukemia incidence. In addition, demographic and environmental factors were postulated to have an influence on leukemia. Consequently, these two aspects will be elaborated upon in the subsequent chapters.

CHAPTER V

LEUKEMIA AND DEMOGRAPHY

Introduction

It is the intention of the present chapter to explore those characteristics of the individual which were intuitively taken to indicate an infectious process. Again, as in the chapter on housing, demography represents a rather complex set of specific variables, which therefore will have to be specified in more detail.

Male/Female Ratio

Generally more males than females contract leukemia. A recent study by Shimkin noted that the sex ratio for leukemia in the United States white population was approximately 1.5 (Shimkin,1965). Goldenberg, on the other hand, noted that there was no significant variation of the ratio for leukemia incidence (Goldenberg,1967). It appears that this ratio flucuates depending on the geographical area analyzed. For instance, Denmark has a high sex ratio of 1.7, while Israel has a ratio of only 1.1 (Shimkin,1965). Other studies, Gilliam for example, found that death from all forms of leukemia was higher in the male with a ratio of 1.5 (Gilliam,1953). Finally, a report involving leukemia occurrences in Winston-Salem, North Carolina, found that the male to female ratio to be 1.45 (Hayes,1960).

Table 27 gives the male/female ratio for the years 1959 through

TABLE 27

SEX RATIOS FOR LEUKEMIA
IN BUFFALO, NEW YORK
1959 - 1966

Year	Males	Females	Male/female Ratio
1959	45	28	1.61
1960	35	26	1.34
1961	40	18	2.22
1962	47	27	1.74
1963	22	31	.71
1964	33	20	1.65
1965	26	24	1.08
1966	33	24	1.37
1959-66	281	198	1.42

1966 for Buffalo. As seen from the table, it was only in 1963 that females contracted leukemia more often than males. As indicated in the previous studies and further in this present analysis it is this trend of males contracting leukemia more often than females that appears to be typical. This observed relationship is noted in Table 33. However, the male being favored in contraction of disease is not limited to leukemia.

Leukemia and Age

Another important demographic aspect of leukemia is represented by the particular age group in which it occurs. Generally, leukemia incidence affects two separate age groups. Children under 15 years of age usually have a high incidence rate. Also, adults over 45 tend to experience high incidence rates.

The age-specific rates for leukemia in Buffalo are given in Table 28. The age-specific rates portray the usual bimodal distribution with high rates for the 0-4 and 5-9 age groups and for the adults over 44. The rate increases steadily from 5.4 at age 45 to a very high 70.2 per 100,000 for age 75 and over.

A report by the State of New York records leukemia incidence rates of 5.31 and 4.49 for the under 5 and 5-9 age groups respectively (Cancer in New York,1962). Also, the report gives the rate for the 75 and over age group as 21 per 100,000 in 1941-43 and about 51 per 100,000 in 1958-60. The indication form these rates and the incidence rates observed for Buffalo is that leukemia incidence has been steadily increasing. Certainly, this is true in the older age group of 75 and over.

TABLE 28

AGE-SPECIFIC INCIDENCE RATES OF LEUKEMIA
BUFFALO, N.Y.
1959-66

Age Group	Cases	Population	8-year Population	Mean annual age-specific Incidence Rate/100,000
0 - 4	24	54,934	439,472	5.5
5 - 9	21	46,339	370,712	5.7
10 - 14	7	41,434	331,472	2.1
15 - 19	8	36,246	289,968	2.8
20 - 24	6	31,535	252,280	2.4
25 - 29	9	32,138	257,104	3.5
30 - 34	16	33,827	270,616	5.9
35 - 39	10	34,973	279,808	3.6
40 - 44	11	35,179	281,432	3.9
45 - 49	15	34,601	276,808	5.4
50 - 54	19	31,597	252,776	7.5
55 - 59	33	30,272	242,176	13.6
60 - 64	44	28,806	230,448	19.1
65 - 69	66	24,968	199,744	33.0
70 - 74	75	18,395	147,160	51.0
75 +	103	18,328	146,624	70.2
All Ages	469			11.0

Leukemia-Duration and Age

Another significant aspect to consider is the classification of leukemia into acute or chronic cases. Acute cases usually occur in young children and are more suggestive of an infectious process. The onset is fast and the disease runs it course in a rapid manner, usually resulting in death after six months. Chronic cases, on the other hand, usually exhibit an insidious onset with a lenghty period from the onset to death. Victims are known to have survived ten to fifteen years after onset. This is not typical for infectious diseases but rather suggests that synergistic effects be hypothesized as a cause. For this reason, the following analysis will divide the cases into chronic and acute. Further, age will be divided into four distinct groups.

All Ages: Acute and Chronic

Four age groups were arbritrarily selected for this analysis, namely under 14; 15-49; 50-64; and over 65. Table 29 gives the results for all ages sub-divided into acute and chronic cases. The multiple correlation coefficients for both the acute and chronic analysis are significant at the .01 level. The acute analysis results in significant values for all variables except persons per house; while the chronic analysis shows significant values for all the variables except the average value of the home. The variable with the highest maximum partial correlation for the acute cases is the average number of rooms in the owner-occupied unit. The relationship is negative. The same is true for the chronic cases except that the partial correlation is considerably lower. Average rent for both analyses has approximately the same partial correlations but the relationship is negative for the acute

TABLE 29

REGRESSION AND CORRELATION ANALYSIS
ACUTE AND CHRONIC LEUKEMIA CASES
ALL AGES

Variables	Variable Number	Acute (b)	Acute (r)	Chronic (b)	Chronic (r)
Average value of the housing unit	x_{106}	.267	.220*	-.045	-.043
Average rent	x_{107}	-.200	-.188*	.282	.197*
Percent of 1.01 or more persons per house	x_{109}	-.174	-.166**	-.171	-.183*
Average no. of rooms of owner-occupied units	x_{110}	-.656	-.505*	-.300	-.308*
Average no. of rooms of renter-occupied units	x_{111}	-.221	-.190*	-.464	-.372*
Persons per house	x_{112}	.131	.114	.383	.372*
Multiple Correlation Coefficients			.643*		.542*

* Significant at the .01 level ** Significant at the .05 level

cases and positive for the chronic. This suggests that acute cases occur where lower rents are paid while chronic cases occur where higher rents are paid. However, if the home is owned then a high average value of the home is associated with increased rates in the acute cases. Thus, it seems that if one rents, then the occurrence of leukemia increases with decreases in rents. But, if one owns then the occurrences of leukemia increases with increases in average value of the home.

One would expect that homes with a high average value in dollars would have more rooms. However, the average number of rooms in the owner-occupied unit follows a pattern of high value and fewer numbers of rooms related to leukemia increases. Finally, variable x_{109}, percent of 1.01 or more persons per room, has low significant correlations with negative relationships. This indicates that crowding is not a factor and is not typical of infectious diseases. But, it may be noted that the chronic cases seem to occur where there is an increase in the number of persons per house.

Age Group 14 and Under

Table 30 presents the results for the youngest age group and for the 15-49 age group. For the 14 and under age group the 40 cases were not divided into acute and chronic since only a few of the cases were chronic. It appears from the high multiple correlation coefficient (.828 and significant at the .01 level) that all six variables aid in explaining the incidence of acute leukemia for the 14 and under age group.

Two specific variables, average number of rooms in the renter-occupied units (x_{111}) and percent 1.01 or more persons per room (x_{109}),

TABLE 30

REGRESSION AND CORRELATION ANALYSIS
ACUTE AND CHRONIC LEUKEMIA CASES
AGES 14 AND UNDER AND 15-49

Variables	Variable Number	Acute and Chronic Under 14 (b)	Acute and Chronic Under 14 (r)	Acute 15 to 49 (b)	Acute 15 to 49 (r)	Chronic 15 to 49 (b)	Chronic 15 to 49 (r)
Average value of the housing unit	x_{106}	.021	.055	.733	.478**	.217	.270
Average rent	x_{107}	-.085	-.190	-.809	-.612*	-.468	-.431
Percent of 1.01 or more persons per room	x_{109}	-.149	-.377*	.195	.187	.044	.105
Average no. of rooms of owner-occupied units	x_{110}	-.057	-.102	-.698	-.602*	.283	.172
Average no. of rooms of renter-occupied units	x_{111}	-.358	-.603*	.912	.491**	.340	.221
Persons per house	x_{112}	.100	.170	-.470	-.406**	-.301	-.342
Multiple Correlation Coefficient			.828*		.768*		.527

* Significant at the .01 level ** Significant at the .05 level

provide the highest degree of explanation. The partial correlation for variable x_{111} is .603 and it has a negative regression coefficient, which means that increased rates of leukemia are associated with decreases in the average number of rooms. The regression coefficient suggests that for every unit decrease of the average number of rooms in the renter-occupied units there is a .358 increase in leukemia incidence. Percent 1.01 or more persons per room, variable x_{109}, has a partial correlation coefficient of -.377. Nevertheless, the multiple correlation coefficient is high and significant.

Age Group 15-49

The results for the age group 15-49 are also given in Table 30. The chronic cases in this age group show no significant relationship either in the partial correlations or in the multiple correlation coefficient. However, the acute leukemia cases show a multiple correlation coefficient of .768 which is significant at the .01 level. Five of the six variables are significantly related to leukemia incidence.

Two of the variables, x_{106} and x_{111}, are positively related to the dependent variable with partial correlation coefficients of .478 and .491 respectively. Accordingly, the leukemia incidence is increased in association with average value of the house and the average number of rooms in the renter-occupied units. Both coefficients indicate that increased incidence is associated with a higher level of socio-economic status. Notably, more rooms in a renter-occupied unit possibly suggests larger apartments, hence higher costs. But in contrast to this, the average rent paid by the diseased individuals is negativley associated with leukemia. Likewise, the average number of rooms in the owner-occu-

pied unit has a negative relationship. These two variables indicate a low socio-economic level. Finally, the average number of persons per house is a significant variable indicating a crowding condition. In conclusion, this means that the analysis thus far has been more successful for the acute cases than for the chronic cases.

Age Groups 50-64 and Over 65

Table 31 presents the results for the 50-64 age group while Table 32 presents the results for the over 65 age group. It appears that as age increases the variables which explain the chronic cases become more significant. Thus, from table 30 it is clear that the chronic cases are not significant in the 15 to 49 age group. In the 50-64 age group, on the other hand, the chronic cases reach a significance level of.05 and in the over 65 age group the significance level is .01. However, there is no trend in the correlation coefficients between the three age groups.

There are only two variables that have significant partial correlations in the 50-64 chronic age group. They are average number of persons per house which is positively related to leukemia incidence and average number of rooms in the renter-occupied unit which is negatively related to leukemia incidence. This indicates that fewer rooms and more persons per house are positively related to leukemia. The two previous aspects are identical for the chronic cases in the 65 and over age group. However, crowding is noted in the house for this age group, which contradicts this tentative conclusion. Conversely, an increase in leukemia is related to a decrease of the percent of 1.01 or more persons per room. Obviously, it is not suggestive of crowding.

TABLE 31

REGRESSION AND CORRELATION ANALYSIS
ACUTE AND CHRONIC LEUKEMIA CASES
AGES 50-64

Variables	Variable Number	Acute (b)	Acute (r)	Chronic (b)	Chronic (r)
Average value of the housing unit	x_{106}	.513	.374	.068	.059
Average rent	x_{107}	-.462	-.471*	.386	.249
Percent of 1.01 or more persons per room	x_{109}	-.032	-.076	.228	.195
Average no. of rooms of owner-occupied units	x_{110}	-.739	-.610*	-.196	-.176
Average no. of rooms of renter-occupied units	x_{111}	-.015	-.013	-.424	-.310**
Persons per house	x_{112}	-.246	-.173	.325	.294**
Multiple Correlation Coefficient			.766*		.695**

* Significant at the .01 level ** Significant at the .05 level

TABLE 32

REGRESSION AND CORRELATION ANALYSIS
ACUTE AND CHRONIC LEUKEMIA CASES
65 AND OVER

Variables	Variable Number	Acute (b)	Acute (r)	Chronic (b)	Chronic (r)
Average value of the housing unit	x_{106}	.196	.159	-.108	-.092
Average rent	x_{107}	.095	.083	.257	.178
Percent of 1.01 or more persons per room	x_{109}	-.077	-.054	-.427	-.407*
Average no. of rooms of owner-occupied units	x_{110}	-.725	-.508*	-.278	-.285
Average no. of rooms of renter-occupied units	x_{111}	-.460	-.377*	-.456	-.375*
Persons per house	x_{112}	.357	.297*	.380	.404*
Multiple Correlation Coefficient			.657*		.594*

* Significant at the .01 level

Three variables are significant for the 50-64 acute age group and four are significant for the 65 and over acute age group. Variables x_{106} and x_{110} have the same relationship for both age groups. That is, that both high value of the home and fewer rooms in the owner-occupied units are related to increased rates of leukemia. Conversely, in the 50-64 age group, variable x_{107} suggests lower socio-economic status. Variable x_{111} in the over 65 age group suggests crowding conditions particularly when coupled with variable x_{112} which also suggests crowding conditions.

Summary of All Age Groups

Summarizing the analysis of the acute cases it should be noted that variable x_{106} is positive for the 50-64 and for the over 65 age groups. Hence, older acute cases occur in homes with higher average value. Variable x_{110} has a negative relationship for all but the 14 and under age group, which indicates that fewer rooms in the owner-occupied unit are associated with increased leukemia rates. Variable x_{111}, is positive for the 15 to 49 age group and negative for under 14 and the over 65 age groups. Therefore, depending of the age group, the relationship varies. Variable x_{109}, is negative only for the 14 and under age group and does not indicate crowding. Finally, variable x_{112} is negative for age group 15-49 and positive for the over 65 age group. Again, one analysis suggests crowding to be a main factor while the other analysis suggests lack of crowding. In general, significant correlations are obtained but the relationships are not consistent. This suggests that the age-specific analysis is related to the classification of acute and chronic cases.

Summarizing the analysis of the chronic cases, it is worth noting that for age groups under 14 and 15-49 there are no significant relationships. Variable x_{111} has a negative regression coefficient for the 50-64 and over 65 age groups. This coupled with the positive relationship of variable x_{112} is very suggestive of crowding in the older chronic cases. However, variable x_{109} is negatively related in the 65 and over age group, which contradicts the results from variable x_{112}. The observed relationships for the age groups are noted in Table 33.

Leukemia and Income

Socio-economic status, as signified by income, is another variable which may be related to leukemia incidence. The median income as specified by the U. S. Census was used to determine if a relationship exists between leukemia and income.[10] Four income groups were established: median income below 4000 dollars, income between 4001 and 6000 dollars, income between 6001 and 8000 dollars, and income over 8000 dollars. The analysis was carried out on the 75 census tracts. The number of cases and the population was determined for each income group, after which an incidence rate was computed.

Reviewing Table 34 a definite pattern emerges. Thus, the highest incidence rate, 17.3 per 100,000 is in the highest income group. And, the second highest incidence rate, 11.7 per 100,000 represents the lowest income group. However, this latter rate is not in great variance with the 6001-8000 dollar income group. This means of course that high median income was found to be positivley related to leukemia incidence.

---

[10] U.S. Bureau of the Census, U.S. Census of Population: 1960, Series HC. Census Tracts. Buffalo, New York. Table P-1. pp. 16-21.

TABLE 33

OBSERVED RELATIONSHIP OF THE TESTED VARIABLES
LEUKEMIA AND DEMOGRAPHY

Exact definition of the variable	Hypothesized Relationship with Y	Observed Relationship with Y
Age Groups of the Deceased 14 and under 15-44 45-64 65 and over	Direct for Chronic cases; Inverse for Acute cases.	Direct for Chronic cases; Inverse for Acute cases.
Deceased Male/Female Ratio	Direct	Direct
Yearly Median Income Groups per Census Tract 4000 4001-6000 6001-8000 8000	Inverse	Direct
Occupation of the Deceased	Direct with the amount of "people contact"	Direct with the amount of "people contact"

TABLE 34

LEUKEMIA AND MEDIAN INCOME

Median Income	8-year Cases	Population for 8-years	Mean Annual Incidence Rate per 100,000
4000	31	265,920	11.7
4001-6000	220	2,601,352	8.5
6001-8000	139	1,233,568	11.3
8000	27	161,232	17.3

This relationship may possibly be the result of better diagnosis in the higher income levels. However, that the lower income groups had the next highest rate may be related to poverty, crowding, poor plumbing, etc. Table 33 indicates the hypothesized and the observed results of this variable.

Several papers have investigated income as related to leukemia incidence but with no conclusive results. For instance, it has been found that a higher incidence rate of leukemia are the lower economic groups (Goldenberg,1967), while others claim that the incidence is higher in economically more favored groups (Githens,1965; Pinkel and Nefzger, 1959; Sacks and Seeman,1947). Finally, there are studies in which no relationships whatsoever between income and incidence of leukemia has been reported (Mustacchi,1965; Wood,1960).

Leukemia and Occupation

Certain occupational hazards have been suggested as being re-

TABLE 35

LEUKEMIA AND OCCUPATION

Occupation	Number in Each Occupation
Hospital - Physicians	2
Practical Nurses	2
Kitchen Worker	1
Laundry Worker	1
Attendant (Orderly)	1
Sales - General	2
Discount Stores	4
Lumber	1
General Laborers	6
Secretarys	3
Bank officials	3
Electricians	3
Mechanics	2
Teachers	2
Truck Drivers	2
Taxi Drivers	2
Carpenters	2
Manager	2
Mail Carrier	2
Cement Finisher	2
Engineer	2
Meat Packer	1
Manager Food Store	1
Kitchen Worker - Restaurant	1
Railroad Laborer	1
Mail Clerk - Railroad	1
Ticket Agent	1
Serviceman - Electric	1
Accountant	1
Construction Worker	1
Insurance Agent	1
Hotel Maintenance	1
Postmaster	1
Member of Armed Forces	1
Plant Foreman	1
Clergyman	1
Watch Repair	1
Plumber	1
Lawyer	1
Telephone Lineman	1
Machine Operator	1
Librarian	1
Bank Custodian	1
Fireman	1
State Official	1
Tool Maker	1
Bookkeeper	1
Welder	1

lated to the incidence of leukemia. Physicians generally have a higher incidence rate than that of the lay population (Gilliam and Walter, 1958). Also, individuals occupationally exposed to X-radiation or benzene have increased rates. Graham found that a mother who had experienced radiation prior to conception had a risk of 1.6 of having children with leukemia (Graham,1966). The evidence of actual relationship of various occupations to leukemia incidence is relatively small.

In the present study, occupations were determined for 75 of the diseased individuals. This represents approximately 60 percent of those individuals who had been classified as being employed.

Table 35 gives the occupations for the individuals with leukemia. It can be seen that hospital workers represent the largest group. Two physicians, two practical nurses, a laundry man, an attendant and a kitchen worker represent the seven workers in a hospital. If leukemia were a transmittable virus it is probable that the contact would occur in an institution where the victims of leukemia are concentrated.

The second group of occupations, is sales. An occupation such as this exposes the employee to many different people in an average day. Again, if leukemia is a transmittable virus it does not seem unlikely that someone engaged in this profession might be exposed. The last group, food-related occupations, had three individuals which contracted the disease. As with the other victims, it is possible that these individuals were exposed to a larger number of people. Table 33 points out the hypothesized and observed relationships between leukemia incidence and income.

The final chapter, environment, is analyzed in the subsequent chapter.

CHAPTER VI

LEUKEMIA - SEASONAL ONSET AND AIR POLLUTION

Introduction

There are certain qualties of the physical environment which may support the viral etiology theory. Again, as in chapter four, the intuitive hypothesis is adopted to suggest a link between leukemia and the environmental variables in relation to this viral theory. Two variables will be investigated, namely the seasonal incidence of the disease and the relationship between leukemia and suspended particulate matter.

Seasonal Incidence of Leukemia

The seasonal incidence of leukemia is represented in Table 36. In the monthly variation, leukemia incidence is most prevalent in the month of March and the second most prevalent month is January. The monthly case distribution does not suggest any periodicity of incidence. However, by grouping the months into quarters a more interpretable pattern evolves. That is, that the winter quarter has the highest number of cases, while the summer quarter has the lowest number of cases. This type of pattern with a high in the winter season and a low in the summer season is typical of an infectious disease.

There are considerable variations in the interpretation of the seasonal variation in the incidence of leukemia. If leukemia is an

TABLE 36

SEASONAL VARIATION OF LEUKEMIA
BY MONTH OF ONSET*
1959-1966

Month	Cases by Month	Cases by Quarter
January	49	
February	40	138
March	49	
April	42	
May	35	121
June	44	
July	25	
August	41	93
September	27	
October	35	
November	37	112
December	40	

* Month of onset is the date of 1st symptoms by the patient or the date of 1st visit to the physician.

infectious disease one would expect the highest incidence to be in the winter season. A report by Hayes substantiating this, indicated that the seasonal variation in leukemia is more characteristic of an infectious disease than of a neoplasis (Hayes,1961). Dowsett also reported that the seasonal variation for acute leukemia was at a peak in winter (Dowsett,1965), which however was in contrast to a summer peak noted by Lee (Lee,1962;1963). Other investigations have data which suggest seasonal incidence but peaks of onset vary (Fraumeni,1963; Swan,1963). Furthermore, there are other reports which state that no seasonal incidence exists whatsoever. Both a paper from Norway (Bjelke,1964) and

one from Oregon (Meighan and Knox,1965) support this latter conclusion. Thus, the literature presents several contrasting views.

Leukemia and Air Pollution

The other environmental variable to be treated in this study is the concentration of suspended particulate matter. Data on suspended particulate matter could only be obtained for the years 1963 through 1966. The obtained values represent three monitoring stations in Buffalo. The locations of these stations are shown on Figure 6 along with the distribution of leukemia for 1963-1966.

Stations 1 and 3 are located in areas 9 and 12 - spatial low risk areas; while station 2 is in area 4 - a spatial high risk area. To determine the relationship of the suspended particulate data to the leukemia occurrences, concentric zones were drawn around the stations at a one and two mile radius. Table 37 gives the results of this analysis and suggests that there is no relationship between leukemia incidence and the volume of particulate matter. Notably, station 3 has the highest average value of particulates but the lowest number of leukemia cases within a one mile radius. The two other stations have more leukemia cases but a lower average value of particulate matter.

During most of the year Buffalo experiences a south-west wind from Lake Erie. This southwest wind coupled with the high amount of industrialization in the southern section of Buffalo is the primary reason for the high pollution values to the north of the industrial area. In fact, of the 11 stations in the Greater Buffalo area, station 3 had the highest pollution values for the 1963 to 1966 period. If a positive relationship existed between the volume of suspended part-

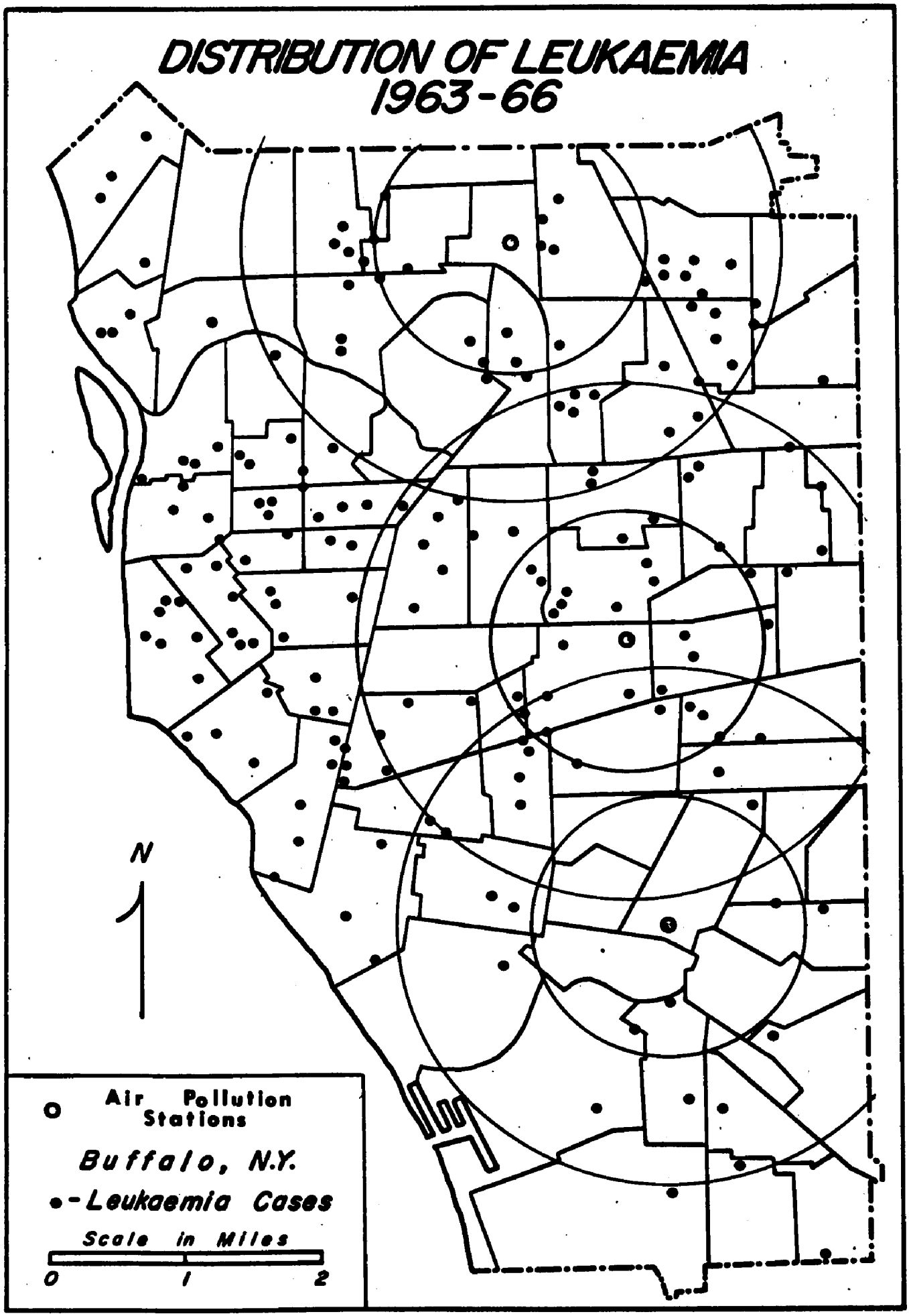

DISTRIBUTION OF LEUKAEMIA
1963-66
N
Air Pollution Stations
Buffalo, N.Y.
•-Leukaemia Cases
Scale in Miles
0
1
2

Figure 6

TABLE 37

LEUKEMIA AND AIR POLLUTION
1963-1966

Station	Concentric Zone in Miles	Suspended Particulate in Micrograms per Cubic Metre	Cases
1	1	89	13
	2		48
2	1	108	24
	2		61
3	1	249	3
	2		31

iculate matter and leukemia occurrences, one would expect a greater number of cases in this area. As noted, however, the opposite was true.

TABLE 38

OBSERVED RELATIONSHIP OF THE TESTED VARIABLES
LEUKEMIA AND ENVIRONMENT

Exact Definition of the Variable	Hypothesized Relationship with Y	Observed Relationship with Y
Average temperature per season	Inverse	Inverse
Air Pollution Micro-grams of particulate per cubic metre	Direct	Inverse

Summary

Seasonal variation was in accordance with some reports but contradicted others. The analysis of suspended particulates in relation

to leukemia showed no significance. The observed relationships between leukemia incidence and the environmental variables are noted in Table 38.

CHAPTER VII

SUMMARY AND CONCLUSIONS

As the preceding analyses of various factors related to leukemia have indicated, the findings of geographical epidemiological investigation are considerably complex. Not infrequently, their results are inconsistent (Mechanic,1968). Therefore it is important to note the several and frequently contradictory findings that are representative of the disease leukemia. These contradictory findings are not peculiar to leukemia but are evident in most geographical epidemiological research projects (Mechanic,1968). Nevertheless, an attempt to clarify the most important findings of this dissertation is provided in Table 39. This table illustrates those variables which have been found to be most relevant, at various scales, to the spatial intra-urban variation of leukemia. Even in view of the contradictory remarks the results do require more attention and elaboration.

Several thoughts and views on the etiology of leukemia have been presented in this study. Three particular factors were investigated to determine their effect on the intra-urban variation of the disease. As stated, housing, demography and environmental aspects were used for further analysis.

With this intuitive idea in mind, the study focused on four specific objectives. These objectives were: (1) to determine the

TABLE 39

SUMMARY OF SIGNIFICANT FINDINGS

FACTOR	SCALE	INCREASED INCIDENCE OF LEUKEMIA ASSOCIATED WITH
HOUSING	census tract	high average rent.
	block	fewer average number of rooms in the renter and owner-occupied units.
		more persons per house.
	tax district	fewer average number of rooms in the renter and owner-occupied units.
		low average rent.
		more persons per house.
	housing unit	apartments. fireplaces. decreased number of rooms in the home.
DEMOGRAPHY	block	crowding in the under 14; and the 65 and over age groups acute and chronic. crowding in the 50-64 age group: chronic.
	census tract	high income
	individual	hospital employees and sales employees.
ENVIRONMENT	city	seasonal onset typical of an infectious disease.

relationship of various housing characteristics to the spatial incidence of leukemia; (2) to determine the relationship of demographic aspects to the spatial pattern of leukemia; (3) to determine the relationship of environmental aspects to the spatial pattern of leukemia; and (4) to evaluate these preceding objectives under changing spatial scales.

Focusing on various scales within the intra-urban structure of Buffalo, house and house-related characteristics were investigated for associations with the pattern of leukemia occurrences. The results obtained were significant and the correlations were high. Hence, "spatial high and low risk areas" were identified for the spatial intra-urban pattern of leukemia but only for the tax district scale. In contrast, the use of the census tract and the block as a spatial unit appeared inadequate.

The identification of the spatial high and low risk areas was based on several variables. These variables were in turn tested at various other scales. Certain variables, for example the average number of rooms in the owner and renter-occupied units for the census tract and block analysis oscillated in their association to leukemia incidence. For instance, they were positively related for the census tract scale but negatively related for the block scale.

The standardized form of the more important variables provided no significant results for the census tract scale. However, the same variables for the block scale did suggest that increased leukemia is associated with crowding. Moreover, the variable, persons per house, is further suggestive of this crowding effect.

The crowding effect noted for the block scale was also evident for the identified spatial high risk areas. The correlations were

high and significant. The average number of rooms for owner and renter-occupied units and persons per house indicated a negative and positive relationship respectively. This, indicates a crowded condition.

Data on individual homes in whidh leukemia residents lived showed some interesting and curious relationships. The variables tested were those which described the house from a construction point of view, such as plumbing facilities, number of rooms and condition of the home. It may be noted that most of the variables showed no significant differences. But, the type of dwelling in which leukemia individuals lived did differ significantly from a control sample. The significance was attributed to the greater number of patients who had lived in apartments rather than in single or double dwellings. Because of their nature, large apartment houses are necessarily more crowded than single dwellings.

It was also noted that there was a higher frequency of fireplaces in the case homes as compared to the control homes. It is probable that this may be attributed to the variation in the type of dwelling in which a diseased individual had lived. For instance, the older, larger homes in the city had been converted to small one or two room apartments. Fireplaces were extremely prevalent in these homes to the extent of several per house. Thus, the significant difference in type of dwelling may indeed account for the prevalence of fireplaces in the diseased individuals' homes. Certainly, there are other aspects to consider. Frequency of fireplaces may be related to a high socio-economic status. Too, fireplaces provide and opening through which penetration may be made by various vectors.

Finally, the number of rooms per home were lower for the

diseased individuals as compared to the control sample. This is evidence of crowding in the home.

Certain aspects of demography and leukemia were found to be consistent with existing trends. Specifically the sex ratio, and the age-specific incidence were in agreement. Other aspects of demography suggested new hypotheses that might be seriously considered for research in the future. Age-groups were selected and divided into two classes based on duration. The variables tested yielded significant results, although the relationships obtained were generally inconsistent. It was concluded that the age group analyzed is significant to the interpretation of the results. For example, it was found that older acute cases (50-64 and over 65) occurred in homes of high average value and under crowded conditons. For the age group 15-49 there was no definite pattern. The relationships were contradictory; for instance, the average value of the home was higher while the average rent was lower. Also, the variables for indicating crowding or non-crowding did not agree. Likewise, the relationships did not agree for the under 14 age group. However, the correlation coefficient for this age group was higher than any other — acute or chronic. The chronic cases were definitely related to crowding, particularly in the older age groups (50-64 and over 65).

Leukemia cases occurred most frequently in the highest income groups. The second highest rate occurred in the lowest income group. Occupations of the diseased individuals appeared to be a pertinent factor. Hospital employees and sales personnel accounted for almost 20 percent of the diseased individuals; this included two physicians. Both of these occupations are conducive to the transmission of

infectious diseases.

A seasonal pattern was noted for the leukemia occurrences for the time period studied. Thus, the leukemia cases were more prevalent in the winter quarter.

Air pollution as measured by particulates, related to leukemia occurrences proved to be negative. In fact, low leukemia occurrences existed in the vicinity of high pollution.

It has been shown that leukemia exhibits many characteristics of an infectious disease. Certainly, the evidence presented is not conclusive but it is substantial enough to suspect that leukemia may be caused by a virus. Further, the evidence found in this paper gives support to the viral etiology theory although it has not been conclusive in this regard. To incriminate a virus as the cause of leukemia one must therefore pursue further research. Probably, there are several events of a chain of events that predispose individuals to leukemia. For instance, a chromosomal abnormality which, when one is exposed to a specific agent or a specific orderly set of events, "triggers" the disease. If this were the case, as it might well be, the obvious direction of research is to identify those individuals with a chromosomal abnormality and then consider their habits and life style over time and in space.

APPENDIX I

INTERNATIONAL CLASSIFICATION OF DISEASES
EIGHTH REVISION
SITE CODES 200-207

NEOPLASMS OF LYMPHATIC AND HEMATOPOIETIC TISSUE (200-207)

Excludes: secondary and unspecified neoplasm of lymph nodes (196)

200 Lymphosarcoma and reticulum-cell sarcoma

200.0 Reticulum-cell sarcoma (M964)
Reticulosarcoma
Reticulum-cell lymphosarcoma
Reticulosendothelial sarcoma (M972)

200.1 Lymphosarcoma (M961, 962, 963)
Lymphosarcoma (lymphoblastic)
(Lymphocytic)

201 Hodgkin's disease (M965,966,967,968)
Hodgkin's:
granuloma
lymphoma
Lymphogranuloma (malignant)
lymphogranulomatosis

202 Other neoplasms of lymphoid tissue

202.0 Giant follieular lymphoma (M969)
Brill-Symmers disease
Lymphoid follicular reticulosis
Macrofollicular lymphoma

202.1 Mucosis fungoides (M970)

202.2 Other primary malignant neoplasms of lymphoid tissue (M959,993,985)

Chloroma (M993)
Compound lymphoma
Leucosarcoma
Lymphoma, malignant
Lymphoma, NOS

Malignant neoplasm of:
lymph nodes or glands, specified as primary not classifiable under 200-202.1
spleen
Malignant Reticulosis (M989)

202.9 Other forms of lymphoma

Benign lymphoma
Benign or unspecified neoplasm of bone marrow
Lymphoid polyp (benign)
Ocular lymphoma

203 Multiple myeloma (M973)

Kahler's disease
Myeloma:
NOS
hemic
plasma cell
Myelomatosis
Plasmocytoma

204 Lymphatic leukemia (M982,983)

Includes: aleukemia, leukemia (aleukemia)
lymphatic
lymphoblastic
lymphogenous
lymphoid

204.0 Acute

204.1 Chronic

204.9 Unspecified

205 Myeloid Leukemia (m986)

Includes: aleukemai, leukemia (Aleukemia)
granulocytic
myeloblastic
myelocytic
myelogenous
myelosclerotic
myelosis,aleukemic

205.0 Acute

205.1 Chronic

205.9 Unspecified

206 Monocytic Leukemia (m989)

Includes: aleukemia, leukemia (aleukemia):
histiocytic
monoblastic
monocytic

206.0 Acute

206.1 Chronic

206.9 Unspecified

207 Other and unspecified leukemia (M980,981,983,987,988,990,991,992)

Includes: aleukemia, leukemia (aleukemia):
NOS megakaryocytic
giant cell stem cell

207.0 Leukemia, acute

207.1 Leukemia, chronic

207.2 acute erythremia (M984)
Di Guglielmo's disease Erthremic myelosis (acute)

207.9 Leukemia, unspecified

APPENDIX II

INTERNATIONAL CLASSIFICATION OF DISEASES
SEVENTH REVISION
SITE CODES 200-205

NEOPLASMS OF LYMPHATIC AND HEMATOPOIETIC TISSUES (200-205)

200 Lymphosarcoma and reticulosarcoma

This title includes primary malignant neoplasm of lymph nodes or glands. It excludes secondary and unspecified neoplasm of lymph nodes (198) and secondary neoplasm of spleen (199)

200.0 Reticulum-cell sarcoma
Reticulosarcoma
Reticulum-cell sarcoma

200.1 Lymphosarcoma
Lymphosarcoma, and site
This title excludes reticulum-cell sarcoma (200.) and Hodgkin's lymphosarcoma (201).

200.2 Other primary malignant neoplasms of lymphoid tissue
Chloroma
Chlorosarcoma
Leukosarcoma
Lymphochloroma
Malignant neoplasm NOS of spleen
Malignant neoplasm of lymph glands of any site, or of unspecified site, specified as primary except as classified under 200.0-200.1

201 Hodgkin's disease

Hodgkin's:
disease
granuloma
lymphoblastoma
lymphoma
lymphosarcoma
lymphadenoma
lymphadenosis NOS
lymphogranulomatosis NOS
multiple lymphadenomata
pseudoleukaemia, Hodgkin's type

202 Other forms of lymphoma (reticulosis)

202.0 Giant follicular lymphoma (Brill-Symmers' disease)
Brill-Symmer's disease
Giant follicular lymphoma
Lymphoid follicular reticulosis

202.1 Other
Benign or unspecified neoplasm of bone marrow
Malignant reticulosis
Reticulo-endothelioma
Reticulo-endotheliosis
Reticulosis (lymphoma) not classifiable under 200-202.0

203 Multiple myeloma (plasmocytoma)

Erythroblastoma
Haemic myeloma
Kahler's disease
Multiple myeloma
Multiple myelomata
Myelocytoma
Plasma cell myeloma
Plasmocytoma

204 Leukemia and aleukemia

This title excludes agranulocytosis (297)

204.0 Myeloid leukemia
leukemia, aleukemia leukocythaemia, leukosis:
lymphatic
lymphoblastic
lymphoid
lymphocythaemia

This title excludes the listed conditons if specified as acute (204.3).

204.1 Myeloid leukemia
leukemia, aleukemia, leukocythaemia, leukosis:
esoinophilic
myeloblastic
myelogenous
myeloid
myelaemai
myelocythaemia
myelosis, leukopenic

This title excludes the listed conditions if specified as acute (204.3).

204.2 Monocytic leukemia
Leukemia, aleukemia, leukocythaemia, leukosis
monoblastic (acute)
monocytic (acute)

204.3 Acute leukemia
leukemai, aleukaemia, leukocythaemia, leukosis: acute, unspecified as to type
Any condition in 204.3 or 204.1 specified as acute

204.4 Other and unspecified leukemia
Acute erythraemia (Di Guglielmo's disease)
Leukemia, aleukaemia, leukocythaemai, leukosis: not acute, unspecified as to type
Megakaryocytic leukemia

205 Mycosis fungoides
Granuloma fungoides
Mycosis fungoides

APPENDIX III

This is the field sheet which contained the data on the individual housing characteristics - diseased and non-diseased.

THE CITY of BUFFALO, N.Y., REAL PROPERTY ASSESSMENT

Assessors Maps and Survey / Section-Plate	ASSESSED TO	FRONT FEET	DEPTH	Street No.	ASSESSED ON	SIDE	DISTANCE FROM Feet	Course	Street	TYPE

Address

Unit Full Value of Land F.F. S.F.

Constructed or Moved on Lot in

Remodeled or Repaired in

WIDTH	DEPTH	DPTH. FACTOR	EFF. F.F.	AREA	UNIT VALUE	FULL VALUE
					FULL LAND VALUE	
					ASSESSED LAND VALUE	

SALES RECORD

YEAR	AMOUNT	REMARKS

RENTS:

BUILDING DESCRIPTION

BUILDING CLASS: Cheap, Average, Good, Expensive

TYPE
Ranch Dwelling
Single Dwelling
Double Dwelling
Duplex Dwelling
Bungalow
Gas Station
Apartment House
Store Building
Office Building
Hotel Building
Shop Building
Factory Building
Church Building
Garage Building
Warehouse

FOUNDATION
Stone
Concrete
Brick
Piers

BASEMENT
Full
Half
Quarter
Cement Floor
Laundry
Water Heater

CONSTRUCTION
Frame
Brick
Tile
Blocks
Re-Conc.
Mill
Steel Frame

EXTERIOR
Siding
Shingles
Stucco
Com. Brick
Face Brick
Compos.

ROOF
Wood Shingle
Compos. Shingle
Slate
Tile
Gravel
Prepared
Metal
Asbestos
Flat
Hip
Gable
Gambrel
Dormers

O. S. TRIM
Wood
Metal
Stone
Terra Cotta
Concrete

HEATING
Unit
Hot Air
Hot Water
Steam
Cen. Sta.
Fireplace

MISCELLANEOUS
Incinerator
Kitchen Cabinets
Pass. Elevator
Frt. Elevator
Fire Escape
Sprinklers
Air Conditioners
Vault

CONDITION
Good
Normal
Poor

EXTRAS

ROOMS	STORIES B	1	2	3	INTERIOR FINISH Floors	Trim	Walls
Living							
Dining							
Kitchen							
Bedroom							
Bathroom							
Lavatory							
Family							
Utility							
Recreation							
Office							
Storage							
Stores							

Complaint No. _____ Filed: _____ , 19____

FRONT

102

SELECTED BIBLIOGRAPHY

Anderson, U. M. et al. "High Risk Groups-Definition and Identification" New England Journal of Medicine. Vol. 273, 1965, pp. 308-313.

Arkin, H. and Colton, F. R. "Statistical Methods". Barnes and Noble, Inc. New York. 1969.

Bailar, J. C. and Dunham, L. J. "World Maps of Cancer Mortality Rates and Frequency Ratios". Journal of the National Cancer Institute. Vol. 41, No. 1, 1968, pp. 155-203.

Beard, J. W. "Viral Tumors of Chickens with particular reference to the Leukosis Complex". Epizootiology of Cancer in Animals, Annals New York Academy of Science. Vol. 108, 1963, pp. 1057-1085.

Bernstein, S. A. "Observations on the Effects of Housing on the Incidence and Spread of Common Respiratory Diseases among Air Force Recruits". American Journal of Hygiene. Vol. 65, 1957, pp.162-171.

Blalock, H. M. "Social Statistics". McGraw-Hill, New York, 1960.

Blalock, H. M. "Causal Inferences in Non-Experimental Research". The University of North Carolina Press, Chapel Hill, 1964.

Bjelke, E. "Leukemia in children and young adults in Norway" Cancer, Vol. 17, 1964, pp. 248-255.

Brown, R. "Explanation in Social Science." Aldine Publsihing Company, Chicago, 1968.

Dean, G. "The multiple sclerosis Problem." Scientific American, Vol. 223, No. 1, July 1970, pp. 40-46.

Dever, A. G. "A temporal-spatial analysis of leukemia occurrences in Buffalo, N. Y. 1959-1966." Unpublished Master's Thesis. State University of New York at Buffalo, 1969. (Will be published in "Readings in Medical Geography" Methuen, 1971.

Dowsett, E. G. "Leukemia in Kingston, Surrey. 1958-1964: An epidemiological study." British Journal of Cancer, Vol. 20, 1965, pp. 16-31.

Ederer, F. et al. "Temporal-spatial distribution of Leukemia and Lymphoma in Connecticut." Journal National Cancer Institute, Vol. 35, 1965, pp. 625-629.

Editorial. "The Virus etiology of Human Leukemia." JAMA, Vol. 186, 1963, p. 146.

Ezekiel, M. "Methods of Correlation and Regression Analysis." John Wiley and Son Inc., New York, 1941.

Fraumeni, J. F. "Seasonal Variation in Leukemia Incidence." British Medical Journal, Vol. 2, 1963, pp. 1408-1409.

Fraumeni, J. F. et al. "Temporal-spatial distribution of Childhood Leukemia in New York State - Special Reference to Case clustering by Year of Birth." Cancer, Vol. 19, 1966, pp. 996-1000.

Gilbertson, W. E. et al. "Housing, the residential environment, and health - a re-evaluation." American Journal of Public Health, Vol. 54, No. 12, 1964, pp. 2009-2013.

Gilliam, A. G. Unpublished data National Cancer Institute, U.S. Public Health Service, 1960.

Gilliam, A. G. "Age, Sex, and Race Selection at Death from Leukemia and Lymphomas." Blood, Vol. 8, 1953, pp. 693-702.

Gilliam, A. G. and Walter, W. A. "Trends of Mortality from Leukemia in the United States, 1921-55." Public Health Reports, No. 73, 1958, pp. 775-784.

Gilmore, H. R. and Zelesnick, G. "Environmental Hodgkin's Disease and leukemia." Pennsylvania Medical Journal, Vol. 65, 1962, pp. 1047-1049.

Githens, J. H. et al. "The relation of socio-economic factors to Incidence of Childhood Leukemia." Public Health Reports, No. 80, 1965, pp. 573-578.

Goldenberg, G. J. and Zarowski, V. S. "Leukemia in Manitoba. Case "clustering", Socio-economic Factors and Seasonal Incidence." Cancer, Vol. 20, 1967, pp. 2200-2212.

Grace, J. T. Jr. "Viruses as Etiologic Agents in Acute Leukemia." Presented at the Symposium on Clinical Aspects of Acute Leukemia, sponsored by the American Cancer Society and the National Cancer Institute, New York City, May 22, 1964.

Graham, S. "Social Factors in Relation to the Chronic Illnesses." Handbook of Medical Sociology, Prentice-Hall Inc., Englewood Cliffs, N. J., 1963, pp. 65-98.

Graham, S. et al. "The socio-economic Distribution of Cancer at Various Sites in Buffalo, New York., 1948-1952." Cancer, Vol. 13, 1960. p. 180.

Graham, S. et al. "Preconception, Intrauterine, and Postnatal Irradiation as related to Leukemia." National Cancer Institute monograph, No. 19, 1966, pp. 347-371.

Hagerstrand, T. "Innovation Diffusion as a Spatial Process." University of Chicago Press, Chicago, 1967.

Haggett, P. "Locational Analysis in Human Geography." New York, St. Martins Press, 1966.

Hayes, D. M. "The Seasonal Incidence of Acute Leukemia." Cancer, Vol. 14, 1961, pp. 1301-1305.

Heath, C. W. and Hasterlik, R. J. "Leukemia Among Children in a Suburban Community." American Journal of Medicine, Vol. 34, 1963, pp. 796-812.

Heath, C. W. et al. "Case clusters in the occurrence of Leukemia and Congential Malformations." Lancet, Vol. 2, 1964, pp. 136-137.

Hochstim, J. R. et al "Poverty Area Under the Microscope." American Journal of Public Health, Vol. 58, No. 10, 1968, pp. 1815-1827.

King, L. J. "Statistical Analysis in Geography." Prentice-Hall Inc., Englewood Cliffs, N.J., 1969.

Knox, G. "Epidemiology of Childhood Leukemia in Northumberland and Durham." British Journal Of Preventive Medicine, Vol. 1, 1962, pp. 17-24.

Lee, J. A. H. "Seasonal Variation in leukemia Incidence." British Medical Journal, Vol. 2, 1963, p.623.

Lee, J. A. H. "Seasonal Incidence in Clinical Onset of Leukemia in Young People." British Journal of Preventive Medicine, Vol. 1, 1962, pp. 1737-1738.

Lewis, E. B. "Leukemia and Ionizing Radiation." Science, Vol. 125, 1957, pp. 965-972.

MacMahon, B. "Epidemiology of Leukemia. Background for Future Studies." Bulletin, WHO, Vol. 26, 1962, pp. 579-584.

McPhedran, P. and Heath, C. W. "Multiple Cases of Leukemia Associated with one House." JAMA, Vol. 209, No. 13, 1969, pp. 2021-2025.

Maile, J. B. "Laboratory Medicine - Hematology." The C. V. Mosby Company, 1962.

Manning, M. D. and Carroll, B. E. "Some epidemiological Aspects of Leukemia in Children." Journal National Cancer Institute, Vol. 19, 1957, p. 1087.

March, H. C. "Leukemia in radiologists in a 20 year period." American Journal of Medical Science, Vol. 220, 1950, pp. 282-286.

Meadors, G. F. "Epidemiology of Leukemia." Public Health Report, Vol. 71, 1956, pp. 103-108.

Mechanic, D. "Medical Sociology" The Free Press, N. Y. 1968.

Meighan, S. S. and Knox, G. "Leukemia in Childhood - Epidemiology in Oregon." Cancer, Vol. 18, 1965, pp. 811-814.

M'Gonigle, G. M. "Poverty, Nutrition and Public Health." Proceedings Royal Society of Medicine, Vol. 26, 1933, pp. 677-687.

Miller, R. W. "The role of Epidemiology in the etiology of Leukemia." Journal National Cancer Institute, Vol. 10, 1964, pp.130-134.

Mustacchi, P. "Some intra-city variations of Leukemia Incidence in San Francisco." Cancer, Vol. 18, 1965, pp. 362-368.

Nelson, H. "Housing and Health." British Medical Journal, Vol. 2, 1945, pp. 395-397.

Pinkel, D. and Nefzger, D. "Some Epidemiological Features of Childhood Leukemia in the Buffalo, New York Area." Cancer, Vol. 6, No. 1, 1960, pp. 102-105.

Pinkel, D. and Nefzger, D. "Some Epidemiological Features of Adult Leukemia in the Buffalo, New York Area." Cancer, Vol. 12, 1959, pp. 351-358.

Polk, R. L.and Company Incorporated. Polk's Buffalo City Directory of Householders. Occupants of Office Buildings and other Business Places including a complete street and avenue guide. 1959-66.

Sacks, M. S. and Seeman, I. "A statistical study of mortality from leukemia." Blood, Vol. 2, 1947, pp. 1-14.

Schwartz, S. O. and Greenspan, I. "Leukemia Cluster in Niles, Illinois: Immunologic Data on Families of Leukemia patients and others." JAMA, Vol. 186, 1963, pp. 106-108.

Seigel, S. "Non-parametric Statistics" McGraw-Hill, New York, 1956.

Shimkin, M. B. "Epidemiology of Cancer: Spatial-Temporal Aggregation." Cancer Research, Vol. 25, 1965, pp. 1363-1373.

State of New York. "Cancer in New York State-exclusive of New York City." Bureau of Cancer Control, 1962.

Suchman, E. A. "The addictive diseases as Socio-economic Health Problems." Handbook of Medical Sociology, Prentice-Hall, Inc., Englewood Cliffs, N. J. 1963, pp. 123-143.

Swan, A. "Seasonal Variation in Leukemia Incidence." British Medical Journal, Vol. 2, 1963, pp. 1063-1064.

U. S. Bureau of the Census, U. S. Census of Housing: 1960, Block Statistics, Buffalo, New York.

U.S. Bureau of the Census, U. S. Census of Population: 1960, Census Tract Statistics, Buffalo, New York.

W. H. O. "Expert Committee on the Public Health Aspects of Housing." Technical Report Series, No. 225, Geneva, 1961.

Wald, N. "Leukemia in Hiroshima City Atomic Bomb Survivors." Science, Vol. 127, 1958, pp. 699-700.

Wilner, D. M. et al. "The Housing Environment and Family Life." The Johns Hopkins Press, Baltimore, 1962.

Wilner, D. M. and Walkley, R. P. "The Housing Environment and Mental Health." American Association for the Advancement of Science, 1959, pp. 143-174.

Wood, E. E. "A survey of Leukemia in Cornwall." British Medical Journal, Vol. 1, 1960, pp. 1760-1764.