CHILDHOOD CANCER MORTALITY IN ONTARIO 1976-1985

AN ECOLOGIC ANALYSIS OF CHILDHOOD CANCER MORTALITY IN ONTARIO 1976-1985

BY
CHRISTINE L. HAMPSON, B.A. (HONS)

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfillment of the Requirements

for the Degree

Master of Arts

McMaster University

MASTER OF ARTS (1991) **McMaster University**
(Geography) **Hamilton, Ontario**

TITLE: An Ecologic Analysis of Childhood Cancer Mortality in Ontario 1976 - 1985

AUTHOR: Christine L. Hampson, B.A. (Hons) (Brock University)

SUPERVISOR: DR. S.M. TAYLOR

NUMBER OF PAGES: ix, 150

ABSTRACT

Geographical variation in human cancer has been well documented among the populations of various areas. Past research on childhood cancer incidence and mortality, however, is inconclusive regarding the spatial variability and associated determinants of cancer in this population sub-group. This thesis examines childhood cancer mortality in Ontario from 1976 to 1985. An initial discussion of the differences between childhood and adult cancer, an overview of previous investigations of geographical variations in childhood cancer and a review of the broad range of correlates of pediatric cancer considered in the literature are provided. Then, using the ecologic study design, county level mortality rates were analysed for all sites combined over the ten year period for ages 0 to 19 by gender. Geographic analysis using maps, spatial autocorrelation, correlation and regression analyses, and an evaluation of regression residuals were used to examine the relationship between childhood cancer standardized mortality ratios and a set of 17 environmental and socio-demographic variables, including: ethnicity, urbanization, parental occupation and income.

The specific objectives of the research were: (a) to describe county level rates of childhood cancer mortality in Ontario for the period 1976 to 1985; (b) to analyse spatial variations in pediatric cancer mortality; (c) to determine to what extent those variations are related to variations in selected ecologic variables; and, (d) to comment on future directions for research in childhood cancer epidemiology.

The results of the spatial analysis of childhood cancer across Ontario revealed no consistent or distinctive patterns of geographic variation. This finding was confirmed by spatial autocorrelation analysis. Multiple regression analysis yielded four independent variables which demonstrated a significant association with sex-specific or total childhood cancer standardized mortality ratios in Ontario: native ethnicity, incidence of low income, number of manufacturing establishments per 1,000 urban population and employment in secondary industries.

Implications for future research in childhood cancer epidemiology include: small area analysis, the use of longer time periods and the incorporation of incidence data. As well, directions for further analysis of childhood cancer using the ecologic methodology and other research designs are discussed.

ACKNOWLEDGEMENTS

This thesis represents the conclusion of a 'substantial' part of my academic life and as such, has been touched by many individuals who have contributed directly or indirectly to its completion. I would especially like to express my sincere appreciation to my supervisor, Dr. Martin Taylor, for his guidance, encouragement and patience.

The data for this thesis was provided by the Ontario Cancer Research and Treatment Foundation. I am grateful to Dr. Loraine Marrett and Donna Reynolds for their contribution in this regard.

Special thanks are due to Susan Birnie for her invaluable assistance, sense of humour and moral support, which made even the most tedious tasks endurable.

I am also appreciative of the support of the Department of Geography at Brock University for giving me the opportunity to accomplish this goal. Special thanks to Colleen Catling for her assistance with the preparation of the manuscript and to Professor Josephine Meeker for her enthusiastic support, tenacious determination and constant belief in me.

Finally, my fondest thanks to Myra for absolutely everything. I did it - just as you always knew I would!

This thesis is dedicated to the loving memory of Phyllis Martin Hampson.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

1.1 Environment and Health

As infectious and parasitic diseases as a cause of death declined in the first half of the twentieth century in most of the developed world due to the impact of better medicine, prevention, nutrition and sanitation, chronic diseases have become the leading cause of death in our urban-industrial society (Greenberg, 1987). Individuals have become increasingly concerned about the impact of the environment and life-styles on their health, and medical researchers have focussed increasingly more attention on these factors as determinants of health status. Geographers have made significant contributions to our understanding of the complex relationships between environment and health, as questions concerning the relationship between people and the environment underlie geographic concerns.

Geographic variation in disease frequency has long been a focus in investigations of the health of human populations. In a recent introductory text in medical geography, Jones and Moon (1987; 38) articulate this approach as follows:

> Geographers in the disease ecology tradition have contributed to the understanding of disease aetiology (or causation) by employing the methods and techniques of epidemiology. Epidemiology has been defined by MacMahon and Pugh (1970) as 'the study of the distribution and determinants of disease frequency in man', and geographers have concentrated on the spatial variations of disease frequency in an attempt to provide clues to disease causation. Much epidemiological and geographical research accepts the biological definition of disease but extends the biomedical model by considering explanations that involve social and environmental causes.

This extension of the biomedical model of disease causation to include social and environmental factors has been most appropriate in the study of chronic disease - including cancer. Although cancer has been strongly resistant to the discovery of causes, it is generally accepted that "causation is extremely complex and multi-factorial, involving interaction among a multitude of environmental, genetic, and life-style factors over a considerable length of time" (Glick, 1982; 471). As a result of the recognition of the importance of social and environmental factors in carcinogenesis, studies of the geographical variation in human cancer have been used to attribute a high proportion of cancer to environmental factors. The extent to which this may be relevant to cancer in our youngest populations is the focus of this thesis.

Cancer in children is different from that found in adult populations with respect to age of occurrence, principal sites and latency. As a result of these differences, the role of socio-demographic and environmental factors versus genetic factors in the study of cancer in children is unclear. Few studies of the geographical variations of childhood cancer have been undertaken, due primarily to the relative rarity of malignancies in this segment of the population. However, despite the low frequency of pediatric cancer, it is among the most common natural causes of death for 0 to 19 year old populations in developed countries. For example, Canadian data, presented in Table 1.1, show that in 1985, after accidents, malignant neoplasms were the leading natural cause of death for 5-19 year olds. As little is known about the causes of malignant disease in children, a geographical analysis of variations of childhood cancer mortality is an appropriate initial step in developing clues to causation.

Table 1.1. Five Leading Causes of Death, by Age Group and Sex, 1985

Cause	Total		Male		Female	
	No.	Rate[1]	No.	Rate[1]	No.	Rate[1]
All ages						
Diseases of the heart	58,330	230.1	32,762	261.3	25,568	199.5
All malignant neoplasms	46,333	182.7	25,534	203.6	20,799	162.3
Respiratory disease	14,056	55.4	8,570	68.4	5,486	42.8
Cerebrovascular disease	13,874	54.7	5,930	47.3	7,944	62.0
All accidents	9,621	37.9	6,498	51.8	3,123	24.4
Under 1 year [2]						
Causes of perinatal mortality	1,212	322.6	697	360.7	515	282.2
Congenital anomalies	954	253.9	504	260.8	450	246.6
Sudden death	330	87.8	200	103.5	130	71.2
Respiratory disease	76	20.2	46	23.8	30	16.4
All accidents	66	17.6	42	21.7	24	13.2
1 - 4 years						
All accidents	240	16.2	150	19.8	90	12.5
Congenital anomalies	111	7.5	59	7.8	52	7.2
All malignant neoplasms	60	4.1	37	4.9	23	3.2
Respiratory disease	37	2.5	20	2.6	17	2.4
Homicide	17	1.1	10	1.3	7	1.0
5 - 19 years						
All accidents	1,223	22.0	865	30.3	358	13.2
All malignant neoplasms	253	4.5	143	5.0	110	4.1
Suicide	239	4.3	199	7.0	40	1.5
Congenital anomalies	87	1.6	41	1.4	46	1.7
Homicide	66	1.2	41	1.4	25	0.9
20 - 44 years						
All accidents	3,513	33.7	2,814	53.8	699	13.4
All malignant neoplasms	2,175	20.9	993	19.0	1,182	22.7
Suicide	1,756	16.8	1,417	27.1	339	6.5
Diseases of the heart	1,107	10.6	887	17.0	220	4.2
Homicide	324	3.1	212	4.1	112	2.2
45 - 64 years						
All malignant neoplasms	14,112	290.4	7,772	325.3	6,340	256.7
Diseases of the heart	10,712	220.5	8,119	339.9	2,593	105.0
All accidents	1,711	35.2	1,206	50.5	505	20.4
Respiratory disease	1,464	30.1	922	38.6	542	21.9
Cerebrovascular disease	1,451	29.9	802	33.6	649	26.3
65 years and over						
Diseases of the heart	46,400	1,758.3	23,716	2,129.7	22,724	1,487.7
All malignant neoplasms	29,722	1,125.3	16,582	1,489.0	13,140	860.2
Respiratory disease	12,176	461.0	7,406	665.1	4,770	312.3
Cerebrovascular disease	12,102	458.2	4,967	446.0	7,135	467.1
All accidents	2,868	108.6	1,421	127.6	1,447	94.7

[1] per 100,000 population

[2] per 100,000 live births

Source: **Canada Year Book**, 1988.

1.2 Context and Research Objectives

This investigation of spatial variations in childhood cancer mortality was initiated through preliminary discussions with researchers at the Ontario Cancer Treatment and Research Foundation. Their suggestion that an examination of childhood cancer in Ontario had yet to be undertaken and thus was needed, provided the rationale for the selection of this topic. Cancer mortality data for Ontario 1976 to 1985, provided by the Ontario Cancer Registry to a multi-disciplinary team of investigators at McMaster University and the University of Toronto and made available for this thesis, provided the opportunity to investigate relationships between social and environmental factors and health in this particular context.

One method for investigating the associations between socio-demographic and environmental variables and disease is through the use of ecologic analysis. In an ecologic study, patterns of disease rates for populations in defined geographic units are related to patterns of postulated risk factors in order to generate etiologic hypotheses. In the epidemiologic literature, ecologic studies have been criticized from the point of view that they are unable to demonstrate cause and effect reliably. On the other hand, in the early stages of research, especially where determinants are suspected to be of an environmental nature and distributed geographically, ecologic analyses are a very appropriate methodology.

The purpose of this thesis is to examine, through ecologic analysis, spatial variations in childhood cancer mortality in Ontario for the period 1976 to 1985 and associated environmental and socio-demographic correlates. Four research objectives will guide this examination: (1) to describe county level rates of childhood cancer mortality in Ontario for the period 1976 to 1985; (2) to analyse spatial variations in pediatric cancer mortality in Ontario; (3) to determine to what extent spatial variations in childhood cancer

mortality are related to variations in selected ecologic variables; and, (4) to comment on future directions for research in childhood cancer epidemiology and to assess the utility of the ecologic methodology in such investigations.

1.3 Chapter Outline

This thesis is organized into six chapters. A review of the literature, contained in Chapter Two, provides the theoretical background for the Ontario study. The chapter begins with a definition of childhood cancer and differentiates it from cancer in adults. An overview of previous investigations of geographical variations in childhood cancer follows. The remainder of the chapter addresses the broad range of correlates of pediatric cancer considered in previous examinations under three main headings: (1) demographic and genetic factors, which include race and ethnicity, age, sex and genetic considerations; (2) environmental factors, comprised of radiation, chemicals, viral infections, seasonality, parental occupation, socio-economic status, urbanization and other lifestyle variables; and, (3) mortality risk factors which can affect varying survival rates. The literature is somewhat inconclusive about the relative contributions of these different sets of factors to childhood cancer mortality and leaves open the possibility that each of these factors is important, thus deserving of further analysis through ecologic and other research designs.

Chapter Three focuses on two methodological issues of relevance to the Ontario study - ecologic analysis and mortality data. A discussion of the advantages of the ecologic methodology, which include its ability to facilitate the study of large populations using existing sources of data as well as its suitability for hypothesis generation, is presented. In addition, the disadvantages of the ecologic approach, such as ecologic bias or the 'ecological fallacy', choice of areal units and multicollinearity, are discussed. A review of

the literature with respect to ways of eliminating some of the bias inherent in ecologic studies is presented as it relates to the design of the Ontario study.

The second methodologic issue considered in Chapter Three concerns the use of mortality data as the outcome of interest in cancer studies. The appropriateness of the use of mortality data in epidemiologic research and its relationship to incidence is evaluated in the context of its availability for the Ontario study. As well, the limitations and disadvantages of mortality data, including diagnosis and recording errors in death certificates, the effects of population migration and the issue of long latency in carcinogenesis, are presented.

The research design of the Ontario study of childhood cancer mortality, 1976 - 1985 is the focus of Chapter Four. Following a presentation of the research objectives of the thesis, this chapter begins with a description of population-based cancer registries as sources of data. The Ontario Cancer Registry, the source of data for this investigation, is described and details concerning the collection, coding and completeness of data in this registry are provided. In addition, the method employed by the Ontario Cancer Registry to calculate population-at-risk data is outlined.

Following a rationale for the selection of the county as the spatial unit of analysis in the investigation of childhood cancer mortality in Ontario and the use of the 1981 **Census of Canada** as the source of data for the independent variables, this chapter details some of the issues involved in the use of these dependent and independent variables. The use of Standardized Mortality Ratios as the dependent variable in the analysis is discussed, and the data for childhood cancer mortality in Ontario for the ten year study period are provided in tabular form. The process whereby independent variables suitable for inclusion in this study were chosen is explained and the resulting 17 independent variables under the headings; urbanization, socio-economic status, parental occupation and ethnicity are

presented. A description of the process by which these variables are displayed visually in maps as well as a brief discussion of the spatial pattern of the independent variables is also provided.

The final section of Chapter Four provides a brief overview of the statistical methods used in the analysis of childhood cancer mortality in Ontario. The main components of spatial autocorrelation, regression and analysis of residuals are discussed.

Chapter Five reports on the findings of the ecologic analysis of childhood cancer in Ontario, 1976 - 1985 as they relate to the established research objectives. Specifically, these results include: (a) a discussion of spatial variations in childhood cancer mortality across Ontario; (b) a presentation of the results of spatial autocorrelation analysis of the Ontario childhood data set; (c) an examination of correlation and multiple regression analyses of the dependent and independent variables; and, (d) a description of the pattern of residuals from the regression analysis.

The concluding chapter contains a summary of the study, a comparison of its results with the findings of other studies of childhood cancer as well as a consideration of the broader implications and contributions of this research. Directions for further research into spatial variations in childhood cancer mortality and morbidity are suggested.

CHAPTER TWO

LITERATURE REVIEW

This chapter will review the literature on childhood cancer to provide background for the Ontario study. Three main aspects of the literature will be considered: first, a definition of childhood cancer will be provided. Secondly, an overview of literature dealing with spatial variations in childhood cancer rates will be presented. The bulk of this chapter will address the third aspect - the associations in the previous literature between childhood cancer and genetic, environmental and socio-demographic variables.

2.1 Childhood Cancer

In developed countries, cancer is the most common natural cause of death for 0 to 19 year old populations, exceeded only by accidents (Birch, 1983). Marked improvements in health care and social conditions since the Second World War resulted in a fall in mortality from infectious diseases. As a result, the impact of childhood cancer in the Western world has increased. Despite its prominence among causes of death, cancer in children is relatively uncommon, exhibiting low incidence and mortality rates (Miller, 1988). As indicated previously in Table 1.1, the cancer mortality rate for 5-19 year olds in Canada in 1985 was 4.6 per 100,000 for both sexes, 5.0 for males and 4.1 for females. This compares with cancer mortality rates for 45-64 year olds in the same year, which exhibited rates of 290.4 per 100,000 for both sexes, 325.3 for males and 256.7 for females.

The last forty years have witnessed a dramatic decline in childhood cancer mortality rates. This fall in mortality has been attributed primarily to early diagnosis and improvements in therapy (Miller and McKay, 1984). West (1984) reports a highly significant trend of decreasing mortality (p. <.001) for most of Western Europe, the United States and Canada. However, despite the gains made in reducing pediatric cancer mortality, the impact on incidence has not been significant (Neglia and Robison, 1988). Although there may have been a recent decrease in the incidence of certain neoplasms such as leukaemia, overall childhood cancer incidence rates have remained relatively stable (Young et al., 1986).

Childhood cancers are a diverse group of malignant diseases which are grouped and considered together strictly based on age of occurrence (Savitz and Zuckerman, 1987). Childhood cancers are generally classified by primary organ site or cell type of the tumour (Li, 1982). More than 50 forms of cancer in children have been identified; however the principal lethal cancer sites for persons under 20 years of age are: leukaemia (acute lymphocytic leukaemia (ALL) and acute myelocytic leukaemia (AML)), brain and central nervous system, lymphomas including Hodgkin's disease and non-Hodgkin's lymphoma, Kidney (Wilms') tumours, bone cancer and retinoblastoma (eye tumours). Leukaemia accounts for an average of 30 percent of all childhood cancer in North America; brain and other central nervous system for approximately 20 percent, lymphomas for 10 to 13 percent and the other sites for smaller proportions.

The justification for analysing childhood cancers as a separate entity lies primarily in the marked difference between the common cancers of childhood and those of adults (lung, colon, breast and prostate) (Savitz and Zuckerman, 1987). Hammond (1986) discusses a variety of important differences between cancer in adults and children. First, that the principal cancers of adult men and women rarely occur among children and, except

for the hematologic cancers and brain tumours, the principal types of cancers that affect children are rare in adults. In addition, epithelial tumours (carcinomas), which are frequently seen in adults are uncommon in children. Pediatric malignancies are primarily sarcomas, deep-seated cancers for which screening is impractical.

The differences between childhood and adult cancers are particularly significant when considering possible risk factors and etiology. Although a high proportion of human cancers can be attributed to environmental factors, their relevance to childhood cancer in particular has not been established (Birch, 1983). In adults, cancers which develop as a consequence of environmental exposures, including those associated with lifestyle, the workplace or with ambient air and water, often occur after a latent period of some 20 to 40 years. Cancers in children, conversely, do not have the opportunity for chronic exposure and long latency because they occur in very young children. However, as will be discussed later, it is possible that environmental carcinogens do play a role in the etiology of childhood cancer (Birch, 1983).

2.2 Spatial Variations in Childhood Cancer Rates

Having established what comprises childhood cancer, the next consideration is to determine what is known regarding spatial variations in childhood cancer mortality rates. Spatial variations in mortality have been generally well documented among the general populations of various areas (Mahoney et al., 1989). This interest in spatial variations in cancer rates has lead to the production of numerous cancer atlases. In a recent inventory and review, Walter and Birnie (1991) found that despite the emergence of these atlases, variations in childhood cancer specifically have not been documented. The strong geographic variation in cancer risk has long been a subject of intense interest to epidemiologists (Glick, 1982). However, differences in cancer mortality among

population sub-groups such as children cannot be discerned due to masking of these differences within the more homogeneous general population. West (1984) suggests that the small differences in childhood mortality when compared to malignant neoplasms in adults, could be explained by the consideration that malignancy in childhood is less influenced by environmental agents. This section will consider the usefulness of a spatial analytic approach in the analysis of cancer rates, and review the existing literature dealing specifically with the spatial distribution and clustering of childhood cancer.

Blot and Fraumeni (1982) point out that geographic variations in cancer mortality are useful for suggesting etiologic clues about risk factors and in generating hypotheses. "Geographical and temporal variations in incidence and mortality rates are among the most important sources of data used to formulate and test hypotheses regarding the causes of different types of cancers" (Brezlow and Langholz, 1983; 703). This method of analysis, then, is a valuable tool in the study of patterns of disease as it raises questions about the causes of cancer which leads to analytic studies aimed at determining causation.

Reports of cancer clusters tend to be well publicized in the media, often alleging a link between a cluster of cancers and some environmental hazard such as a nuclear plant or waste disposal site. Abnormally high occurrences of leukaemia in a small area for a limited time period have been reported for many years and demand attention as possible clues to an environmental determinant (Hanson and Mulvihill, 1982). However, as Corbett and O'Neill (1988) warn, the analysis and interpretation of clustering phenomena present scientific, epidemiological, social and political problems due to the difficulty in determining the role that chance plays in producing the observations. Further, statistical methods indicate that if enough places are studied over a sufficient period of time, a cluster is sure to appear by chance (Walter, 1988). Despite these limitations, cancer clusters provide useful points of departure for more rigorous epidemiologic methods.

Very few studies of the spatial variation of childhood cancer mortality have been undertaken. Before providing an overview of existing studies of this nature from the literature, it is necessary to examine why such a paucity of childhood cancer research exists. According to Breslow and Langholz (1983) the relative rarity of childhood cancer means that long time periods and large populations are required to produce stable rates. Further, the authors state that meaningful analysis of many childhood cancers requires information on cell type as well as topographic site, and this data is not routinely available in cancer registries.

The existing literature dealing with spatial variation in childhood cancer mortality rates indicates that the various forms of childhood cancer do have distinctive patterns of occurrence (Greenberg and Shuster, 1985). Interest is focussed on the characterization of the distribution of childhood cancer in the population and subgroups at high risk. Kramer et al. (1983) point out that this information is crucial in the identification of risk factors, the formulation of testable hypotheses regarding etiology, and ultimately in the designation of preventative strategies. Two spatial patterns are of interest in the case of mortality rates: first spatial variations in potential risk factors for childhood malignancies and second; spatial variations in socio-demographic characteristics which may influence mortality.

Greenberg (1983) analyses urbanization and cancer mortality by age for 1950 - 1975 in the United States. When looking specifically at child and teenage cancer, methodological difficulties to overcome small numbers and produce stable rates had to be managed. Greenberg (1983) did find overall regional differences in the geographical distribution of child and teenage cancer mortality during 1950 - 69. He found that counties with the highest rates for those 0 to 19 years old tended to be the urbanized counties in the New York/New Jersey region and northern California. These areas exhibited relatively high leukaemia, Hodgkin's disease, and to a lesser extent, non-Hodgkin's lymphoma

rates. Ten strongly urban counties in New England exhibited relatively high rates of Hodgkin's disease and six northern California urban counties manifested high leukaemia rates. The last regional pattern noted was the relatively high rates of non-Hodgkin's lymphoma in the urbanized counties near the Great Lakes.

The majority of childhood cancer literature dealing with spatial analysis focuses on clusters of cancer rather than variations over space. Craft and Openshaw (1985) indicate that reports of clusters of leukaemias and lymphomas in the literature have been numerous but the significance of these isolated clusters cannot be evaluated. Further, the authors point out that with rare diseases such as leukaemias and lymphomas, some clustering will occur by chance. Miller (1988) comments on a review of over 100 investigations of cancer clusters, primarily leukaemia, in the United States and none could be linked to an environmental cause.

Greenberg and Shuster (1985) provide a useful overview of previous investigations into childhood cancer clusters. Their conclusion after reviewing the literature is that there is little valid evidence of clustering, possibly due to a number of methodological considerations. In cases where the temporal and spatial boundaries are drawn around the cases ex post facto, chance cannot be excluded as an explanation for clustering. Studies which examined all ages under 20 rather than specific age groups, employed different analytic methods and considered different events (birth, onset of symptoms, diagnosis or death) complicate an assessment of the literature on clustering (Greenberg and Shuster, 1985). It is now thought that most childhood cancer clusters have occurred by chance rather than as the result of environmental or familial influences (Pratt, 1985). Despite the general failure to detect evidence of clustering of childhood cancer, the primary value of such studies is to generate etiologic hypotheses.

The demonstration that certain malignancies do not have a uniform incidence (with reference to sex, age, geographic location) indicates that their occurrence may not be due simply to chance. More importantly there may be identifiable factors which predispose to their development. Although the causes of neoplasms in early life remain unknown, certain observations provide a better understanding of some of the pathogenic factors. It is important to identify individual factors present in the host as well as in the environment which increase tumour susceptibility, as well as the mechanism of interactions of these factors (Lenarsky, 1983). The spatial analysis of childhood cancer mortality rates in Ontario to follow will be aimed at a preliminary examination of variations in space to point at potential avenues for further analysis.

2.3 Association to Genetic, Environmental and Socio-demographic Variables

Having established what constitutes childhood cancer and what is known about spatial variations in childhood cancer mortality rates, it is now appropriate to consider to what extent those spatial variations in childhood cancer mortality can be related to patterns of other variables. This section will briefly discuss a model for the development of cancer with specific relevance to pediatric cancer and then provide an appraisal of the range of covariants of childhood cancer that have been considered in the literature.

Cancer has been strongly resistant to the discovery of causes or the mechanisms involved (Glick, 1982). Causation is typically regarded as extremely complex and multifactorial, involving interaction among a wide range of genetic, environmental, occupational and life-style risk factors, acting over a prolonged time period (Foster, 1986). A detailed examination of models of carcinogenesis is outside of the scope of this paper except for considerations which relate specifically to childhood carcinogenesis.

The best synopsis of the hypothesis for childhood cancer induction is provided by Pratt (1985). He discusses a "two-hit" model of cancer induction proposed by A.G. Knudson Jr. which is consistent with many of the features of childhood cancer. This hypothesis, in its simplified form, proposes that two discrete mutational events may be required to transform a normal cell into a cancer cell. Individuals who have inherited a mutated "cancer gene" may have this defect present in all cells, and require only the second event to develop malignancy. These second events may occur as the result of several factors such as irradiation, chemical carcinogenesis or a virus. These individuals, then, could develop a specific tumour quite quickly and at an age earlier than those without inherited "cancer genes" who require two separate events (Pratt, 1985). Thus in most cases of childhood cancer, both heredity and environmental factors operate, although one or the other might predominate (Newell et al., 1989).

The range of covariants of childhood cancer which have been addressed by the literature is quite broad. This section will provide an overview of the range of factors which have been associated with childhood cancer in the literature and provide an evaluation of the strength of the association, if any, that was found. This will provide the theoretical background for the Ontario study which is the focus of this thesis. Two points must be made clear at the outset. First, that this review will consider the set of literature specific to covariants of **childhood** cancer not cancer in the general population. Secondly, that many studies use incidence, not mortality, as the outcome of interest. But given the preceding discussion regarding the theory of carcinogens, it is not certain whether this is relevant but is important to consider in light of the use of mortality data in this thesis.

Relatively few epidemiologic investigations of suspected risk factors for cancer in children have been published when compared to studies of adults (Neglia and Robison, 1988). Most of these studies are descriptive rather than analytic investigations which lead

to the formation of hypotheses regarding etiology. The greatest single obstacle or impediment to conducting more rigorous studies is the rarity of childhood cancers (Greenberg and Shuster, 1985). As a result of epidemiological, clinical and laboratory studies, the etiology of some cancers (i.e. carcinoma of the bronchus) is extensive. For other cancers (i.e. carcinoma of the cervix), sufficient has been determined to allow the isolation of high risk groups and to formulate some hypotheses. For many cancers, however, very little is known. Childhood cancer falls mainly into this latter category (Birch, 1983).

A survey of the relevant literature has revealed a number of possible risk factors for the malignant diseases of childhood. This survey of the literature has revealed over 14 categories of hypothesized factors of elevated and reduced risk of childhood cancer. However for methodological considerations which will be discussed later, causation is rarely established. Based on a review of the literature, three broad classifications of covariants of childhood cancer mortality can be identified and will be used in this review. The first two deal primarily with risk or etiology of disease (incidence) and the third with risk of mortality from the disease. All three, however, are interrelated. These broad classifications may be termed demographic and genetic factors, environmental (physical and socio-cultural) factors and mortality risk factors.

Many genetic and environmental factors have been studied and discussed in connection with the epidemiology of childhood cancer. Risk factors suggested include factors associated with the child, with the environment, with the mother during pregnancy, or with the parents before conception (Hemminki et al., 1981). Any study of the etiology of pediatric cancer must take into account the interaction of genetic and environmental factors.

2.3.1 Demographic and Genetic Factors

The first broad category is demographic and genetic factors, or characteristics of the host (child). These intrinsic factors are probably of greater relative importance in the development of childhood cancer than adult cancer. According to Birch (1983), etiological factors can be thought of as intrinsic, or the result of an individual's genetic constitution. The occurrence of congenital tumours, known associations with malformations, increased risk in patients due to certain hereditary syndromes and the familial nature of some childhood cancers, all indicate a potentially major role for genetic factors in the etiology of malignant disease in children. These intrinsic factors can be considered under two main headings: demographic characteristics including race and ethnicity, age and sex and genetic factors, comprised of heredity, congenital malformations, and familial considerations.

Demographic Characteristics

Race and Ethnicity: Racial differences have been discussed with regard to childhood cancer etiology. Pratt (1985) reviews the results from the 'Surveillance Epidemiology and End Results' (SEER) Program in the United States by observing that while no great difference in geographic incidence results from the SEER population base, there may be differences between races. Further, Greenberg and Shuster (1985) when analysing the SEER program results, conclude that overall, white children had incidence rates for cancer that were about 30 percent higher than those for black children. Kramer et al. (1983) examined a decade of childhood cancer incidence data in the Greater Delaware Valley Pediatric Tumor Registry and found incidence rates for all types of childhood cancer combined were higher among whites than non-whites. These two studies agree that racial differences, in all sites combined and differences between specific sites, exist. Bunin (1987) examined racial patterns of brain cancer and observed racial differences in the age-

incidence curves of various types of brain cancer. This, it was concluded, could imply genetic and/or environmental factors (Bunin, 1987). Miller (1978) concluded that differences between white and black children in mortality rates indicated racial differences in exposure or susceptibility to some agent which induces certain types of cancer (ALL) but not others (AML). Thus racial variations may reflect important etiological influences. Unfortunately, the Ontario study could not obtain data by race.

Ethnicity was dealt with separately in several studies. Breslow and Langholz (1983) studied Hispanics in New Mexico and found little difference between childhood cancer incidence rates between New Mexico Hispanics and New Mexico Whites. Greenberg et al. (1980) found a excess in childhood leukaemia and young adult Hodgkin's Lymphoma in areas with an excess of population of USSR foreign stock. Newell et al. (1989) noted higher childhood Hodgkin's disease rates in ethnic minorities in the U.S.. Birch (1983) found varying extremes in incidence of certain forms of leukaemia among Japanese and Chinese populations. The Ontario study will look at ethnicity and the importance of this covariant will be discussed later in this chapter when discussing mortality risk specifically.

Age: Several studies in the literature examined childhood cancer frequency by age (sub groups within 0-19 group). Greenberg and Shuster (1985), upon examination of data from the SEER program (based upon almost 26 million child-years of observation) noted incidence rates for cancer among children in the United States that differ by age and anatomic site. No explanations are offered. Newell et al., (1989) when looking at childhood peaks in the SEER program incidence data from 1973 - 1981, comment that clarification of the age distribution pattern of childhood malignancies helps in the identification and timing of exposures. Early peak incidence discovered in childhood or adolescence data for the U.S. suggest brief exposure to a carcinogenic agent in a period of

heightened mitotic activity. They conclude that different risk factors may be operative at different ages. Pratt (1985) suggests that incidence and mortality rates of specific pediatric malignancies vary greatly with age, with prenatal factors thought to affect the incidence of tumours in children under 5 years and postnatal effects related to the incidence in older children. Further, these postnatal factors may include environmental influences. Miller (1967) observes that mortality rates in children with specific forms of cancer exhibit dynamic changes with every year of age. Thus, age variations may reflect important etiological influences. Again, 0 -19 were grouped together in the Ontario study because of small frequencies so this cannot be analyzed.

Sex: A consistent feature of childhood cancer in the literature is an excess of male children. In both children and adults, males are affected more often by cancer than are females; the usual ratio is 1.2:1 with the variations being more evident in younger populations (Pratt, 1985). Breslow and Langholz (1983) studied international variations in sex ratios in **Cancer Incidence In Five Continents** and found a male excess in five of six major sites with a 30 percent excess incidence for males for leukaemia. Savitz and Zuckerman (1987) examined incidence rates for childhood cancer for the Denver Standard Metropolitan Statistical Area from 1976 - 1983, and as expected, males had higher rates than females. Similarly, a study of the Manchester Children's Tumour Registry 1954 - 1980 noted a preponderance of boys among patients (Birch, 1983). Kramer et al. (1983) found incidence rates for all types of cancers combined in the Greater Delaware Valley Pediatric Tumour Registry higher among males than females, regardless of race. In the analysis by Greenberg and Shuster (1985) of SEER Program data in the U.S., male incidence rates were found to be 20 percent greater than those for females. Thus, this pattern is consistent across racial and geographic boundaries. Neutel (1970), upon review

of the literature concludes that males have a higher risk of childhood cancer than females of the same age and suggested that this is not surprising since males are known to have an increased risk of many other diseases as well.

Regarding mortality, not much is written. Neglia and Robison (1988) suggest that sex is an important prognostic variable, with females having better survival and disease - free survival. The Ontario study will be able to look at differences in the mortality data by sex.

Genetic Factors

The ability to identify patterns of cancer incidence and mortality in children by demographic or personal characteristics, thus, may be helpful in providing clues regarding risk factors and may therefore be valuable in generating etiologic hypotheses (Kramer et al., 1983). The second type of intrinsic factor to be considered is genetic, which includes congenital malformations and familial considerations. In addition to demographic characteristics, genetic factors play a role in the etiology of malignant disease in childhood. Susceptibility to childhood cancer may be determined by the inheritance of a gene (or genes) which results in the development of a childhood cancer such as retinoblastoma. It may be influenced by the inheritance of a syndrome which predisposes to the development of malignant disease or by inheritance of susceptibility to environmental carcinogens (Birch, 1983). This section will briefly summarize the literature in this regard.

Familial aggregation of certain types of childhood cancers is reported frequently in the literature (Neglia and Robison, 1988) (Lenarsky, 1983). Pratt (1985) reported on the SEER program data which found specific types of cancer in excess, suggesting that such cancers in families were not chance or random occurrences. Further, Pratt (1985) reports that when cancer has occurred in one sibling, the risk of cancer in the remaining siblings is three times the expected risk based on the normal population rate. The actual risk,

however, is still low because of the rarity of cancer in childhood. Familial clusters of cancer, if not due to chance, may be hereditary, environmentally induced, or a combination of the two (Miller, 1975). Often, however, no etiology is evident and such families represent opportunities for etiological research (Mulvihill, 1982).

The second potential influence of genetics is the inheritance of a syndrome which predisposes the development of cancer. Cancer in children has been found to develop with unusually high frequency in persons with certain genetic diseases (Greenberg and Shuster, 1985). Miller (1967) analyses peaks in cancer mortality for very young children and suggests that the tumours are associated with congenital defects such as Down's syndrome and aplastic anemia. West (1984) also reports that genetic abnormalities such as Down's syndrome have been shown to be associated with childhood cancer. It is unclear, however, if the occurrence of congenital malformations and cancer together is a reflection of gene mutation or environmental influence (Pratt, 1985). Further research is needed to confirm or refute these associations and may reveal other associations (West, 1984).

The third consideration is that a susceptibility to environmental carcinogens can be inherited. Li (1982) points out that humans differ in susceptibility to cancer. Some are prone to cancer and others may be less susceptible and are unaffected despite intense exposure to a carcinogen. In general, the fraction of persons who develop cancer after exposure to a carcinogen is small, e.g. approximately 0.1 percent of Japanese children exposed to atomic bombing have developed radiation-associated leukaemia annually. Li (1982) concludes that those affected may be genetically susceptible.

Thus, there may be identifiable factors which predispose the development of malignancies in children (Lenarsky, 1983). These host factors in childhood cancer may not be particularly relevant in a spatial analysis of childhood cancer mortality in Ontario, but nevertheless warranted discussion here.

2.3.2 Environmental Factors

There is sufficient existing evidence to conclude that the risk of neoplastic disease during childhood is influenced by genetic as well as environmental factors (Greenberg and Shuster, 1985). Although a high proportion of human cancers can be attributed to environmental factors, their relevance to childhood cancer in particular has not been established (Birch, 1983). As discussed previously, cancers developing as a consequence of exposure to an environmental carcinogen often occur after a 20 - 40 year latent period. This latency period does not exist in cancers in children. However, it is possible that environmental carcinogens do have a role in the etiology of childhood cancer (Birch, 1983). Miller (1978) suggests that more than twenty environmental agents are known to cause human cancer and several have short enough latent periods so that the environmental exposure and diagnosis of neoplasia occur within the pediatric age - span.

Before discussing the competing roles of genetics versus environment as potential causes of childhood cancer, it is necessary to review the literature with regard to the role of the environment. A number of environmental agents have been investigated as possible risk factors for cancer among children. The development of cancer in young persons may be influenced by prenatal exposures, as well as by postnatal events (Greenberg and Shuster, 1985).

This section will review the range of physical and socio-cultural environmental covariants of childhood cancer addressed in the literature. This section will be divided into radiation, chemicals and drugs in pregnancy, viral infections, parental occupation, seasonality, socio-economic status, urbanization, and other lifestyle covariants.

Radiation: One of the most comprehensively studied topics is exposure to ionizing radiation and the subsequent occurrence of ALL in childhood (Neglia and Robison, 1988). Epidemiological studies have helped to identify a variety of

environmental agents which are oncogenic in man, but only one of these — ionizing radiation — thus far has been proven to induce leukaemia before 14 years of age (Miller, 1975). The literature on ionizing (natural background) radiation's role in childhood cancer is inconsistent, but there are strong suggestions that there may be a causal association (Savitz and Zuckerman, 1987). Miller (1978) provides a useful summary of previous findings regarding ionizing radiation and leukaemia. He notes that exposure to radiation from the atomic bombs in Japan caused an increase in the frequency of childhood leukaemia. Also, either as a result or coincidentally after a curtailing or elimination of radiation used for shoe fitting and routine flyoroscopic examinations in pediatric practises in 1956, leukaemia mortality rates for children under 5 years of age dropped substantially. Miller (1978) also reviewed a series of papers on increased frequency of leukaemia and all other forms of childhood cancer after exposure of the mother's abdomen during pregnancy to diagnostic irradiation and concludes that much of the evidence supports a causal relationship. Mulvihill (1982) supports this conclusion, and asserts that besides clinical observations, large epidemiological studies document the carcinogenicity of radiation in children. This author cites data from follow-up studies of (1) survivors of the atomic bombs in Japan, (2) survivors of other nuclear explosions and, (3) patients given radiation therapy, and notes that such studies rarely address the possibility that the effect seen in a large population could be accounted for by unusual host sensitivity of a segment of the group. Radiation is not associated with an excess of all types of cancer. Details of radiation carcinogenesis in children vary with the type of exposure, age at exposure, and the type of exposure. Prenatal exposure has been associated with up to a 60 percent excess of all types of malignancy during the first decade of life. Studies, however, were not sufficiently identical in design or results to support a causal relationship. Regarding postnatal exposure, Mulvihill (1982) points out that because of the long latent period for

most radiogenic cancers, few tumours appearing in childhood can be attributed to radiation. The exception is leukaemia.

Birch (1983), when assessing the role of radiation, states that there is no doubt that ionizing radiation is a potent carcinogen and children exposed at Hiroshima and Nagasaki developed leukaemias and solid tumours at an increased rate. Children who have received radiotherapy are also seen to be at an increased risk of developing further neoplastic disease. However, Savitz and Zuckerman (1987) found that rates for total childhood cancer were not excessive in the Denver, Colorado area in spite of the greater exposure to background ionizing radiation.

Numerous studies regarding prenatal exposure have demonstrated increased risks, but whether or not the association is causative has remained a controversial issue. At present, ionizing radiation cannot be considered an important etiological factor in the majority of cases of childhood cancer.

Chemicals: At least 54 chemical agents have been identified as potential human carcinogens (Pratt, 1985). Exposure of children to these agents may occur by transplacental passage which has been demonstrated with the association between diethystilbestrol (DES) in pregnancy and subsequent vaginal and cervical cancer among daughters. Birch (1983) provides two examples of human transplacental carcinogens: (1) the increased incidence of cancer in daughters of women treated with DES and, (2) diphenylhydantion (DPH) is suspected in children exposed to it in utero. Other drugs may well act as transplacental carcinogens. Miller (1978) concludes that any known human carcinogen that can cross the placenta may in theory induce cancer in the offspring, perhaps after a latent period of several decades. The role of industrial chemicals, which will be addressed later, is also relevant as Hemminki (1981) notes that prenatal and postnatal

exposure to pesticides has also been related to childhood malignancies. In addition, Neglia and Robison (1988) discuss speculation regarding pesticide and herbicide exposure and an increased incidence of ALL.

Greenberg and Shuster (1985) report that several investigators have concluded that chemical exposures, either before or after birth, may lead to carcinogenesis during childhood. They point out that a major limitation of published studies concerning exposure to chemicals is the difficulty in collecting information on specific chemical agents, misclassification and small number of cases. Since true relationships may appear spuriously weak in these studies, negative findings must also be interpreted with caution.

Viral Infections: In the 1960s there was great anticipation on the basis of laboratory research that viruses would soon be proven as a cause of human cancer (Miller, 1978). Viruses have received considerable attention in studies concerned with infections as a cause of human cancer (Lenarsky, 1983). Despite the fact that viruses have been shown to play a role in leukaemia in some animals, the lack of consistent associations in human studies provides no evidence of childhood leukaemia with either prenatal or postnatal infectious exposures (Neglia and Robison, 1988). Prenatal exposure to viral infections such as influenza have been reported and studied however the weight of the evidence suggests that influenza is not a strong risk factor for leukaemia in children (Greenberg and Shuster, 1985). Lenarsky (1983) reports other neoplasms occurring in children with epidemiologic patterns suggestive of a viral etiology such as Burkitt's lymphoma in Africa and its relation to Epstein-Barr virus. A more recent review by Newell et al., (1989) states that there is almost universal agreement that early onset of Hodgkin's disease is a consequence of a prevalent infection of low pathology.

Seasonality: Linked to the consideration of infections as a risk factor, is seasonality. The viral etiology of cancer, discussed above, has been studied in terms of hypotheses concerning possible viral modes of spread from one person to another (Miller, 1975). The rates for Burkitt's lymphoma, a childhood cancer which exhibits characteristics most suggestive of an infectious etiology, change quickly over time in a given locale, time-space clustering has been observed, and the distribution of the disease is confined to a climate that favours the proliferation of certain microbial agents (Miller, 1975).

Seasonal variations in the incidence of leukaemia have been the subject of several investigations to assess possible associations between infectious outbreaks and the incidence of ALL, and the observed seasonal variations in incidence in some adult cancers (Neglia and Robison, 1988). While seasonal variations have been observed in some studies, others have failed to identify differences. This inconsistency and lack of an apparent biologic justification for a seasonal variation makes the importance of these findings unclear (Neglia and Robison, 1988).

Harris et al. (1987) in their study of seasonal risk of ALL in the United States from 1973 to 1980, found important seasonal variation in the cumulative risk of ALL among youthful populations. The observed peaks in ALL risk coincide with seasonal variations in the rates of allergenic and infectious diseases, elements of which are capable of promoting lymphocytic proliferation and transformations. Harris et al. (1987) conclude that interactions between seasonality and potential leukomogenic latent viruses could possibly play some role in the development of ALL, but they caution that their interpretations are speculative.

Greenberg and Shuster (1985) provide a useful review the relevant literature. They state that the suspicion that some cancers of children may result from infectious processes has prompted evaluations of seasonal patterns of occurrence. The literature on this topic is

complicated by differences in the methods used to assess seasonality and a lack of uniformity in the event that was studied (e.g. birth, onset of symptoms, diagnosis, or death). Different configurations of study populations and use of a variety of statistical tests further complicate the results. When considered together, these findings provide limited support for the notion that leukaemia of childhood follows a seasonal pattern of onset, with cases tending to occur during the first six months of the calendar year (Greenberg and Shuster, 1985). These authors also conclude that in investigations of seasonal trends for Burkitt's lymphoma in African populations, the weight of the evidence favours a seasonal trend with symptoms typically arising in the spring and an excess of diagnoses during the summer. Other cancers of childhood studied for seasonal trends reviewed by these authors are: Hodgkin's' disease in the U.S. with a peak of birthdates in July and August and a trend for onset and diagnosis during the winter months; and a peak of deaths from neuroblastoma during the summer.

Parental Occupation: Since 1974, numerous studies have been published on the relationship between parental occupation and risk of childhood cancer. Table 2.1 presents, in chronological order, a critical appraisal of 17 case-control studies which have investigated this relationship.

Arundel and Kinnier-Wilson (1986) provide an extensive overview of parental occupation as a risk factor in the occurrence of childhood cancer. They point out that the theory of a correlation between parental occupation and childhood cancer is perfectly plausible and present several possible mechanisms by which this relation could arise: firstly, by effect of germ cell mutation in either parent exposed to an occupational noxious agent in the period leading to conception. Secondly, exposure of the mother before and

Table 2.1

Authors/Year	Case Group	Control Group	Source of Information	Conclusions
Fabia & Thuy (1974)	Fathers of 386 Quebec children who died from cancer before age 5 during 1965-70.	Fathers of 772 children whose birth registration immediately preceded and followed that of each case (2 per case).	Birth certificates.	Significant excess of fathers in hydrocarbon- related occupations (machinists, miners, and painters). Odds ratio = 2.1
Hakulinen et al. (1976)	Fathers of 852 incident cancer cases reported to the Finnish Cancer Registries 1959-68.	Parents of 852 children matched for date of birth and domicile.	Records of antenatal clinics.	No significant associations between childhood cancer and hydrocarbon-related occupations. Risk ratios = 1
Kantor et al. (1979)	Fathers of 149 incident Wilms' tumour cases (age 0-19) reported to the Connecticut Tumour Registry 1935-73.	Fathers of 149 children matched for sex, race and year of birth.	Birth certificates.	A positive statistically significant association between fathers' occupations involving lead and offspring with Wilms' tumour. Odds ratio = 3.7
Kwa & Fine (1980)	Fathers of 692 children born during 1947-57 and 1963-67 who died of cancer before age 15 in Massachusetts.	Fathers of 1384 children whose birth registration immediately preceded and followed that of each case child.	Birth certificates.	Positive association between paternal occupation as paper or pulp mill workers and nervous system tumors (relative odds 2.8); positive association between paternal occupation as mechanic or machinist and tumours of the urinary tract (relative odds 2.5); no association with hydrocarbon-related occupations.
Zack et al. (1980)	Parents of 296 children with cancer who were followed at Texas Children's Hospital Research Hematology Clinic 1976-77.	Three groups: Parents of children without cancer attending clinic at same time; siblings of parents of case group; Neighbours of case parents matched for number and age of children.	Face-to-face or telephone interviews.	No association between hydrocarbon-related occupations and child-hood cancer for any case-control comparison.

Table 2.1 - continued

Authors/Year	Case Group	Control Group	Source of Information	Conclusions
Hemminki et al. (1981)	Parents of childhood cancer cases (age under 15 at diagnosis) reported to the Finnish Cancer Registry 1959-75 (approximately 1600 cases).	For the 1959-68 cases: parents of the child born next after the case. For the 1969-75 cases: parents of the children born next before and next after the case.	Records of welfare centres attended by mothers during pregnancy.	Material occupations found significantly more frequently among cases than controls included farmers' wives, pharmacists, sales women, bakers, and factory workers. Paternal risk occupations appeared to be farming, motor vehicle driving, machine repair, painting and work of men who gave academic degrees as their occupation.
Sanders et al. (1981)	Fathers of all children dying from neoplasms in England and Wales 1959-63 (4395) and 1970-72 (2525).	Fathers of all children dying from all other causes in England and Wales 1959 - 63 (112840) and 1970 - 72 (54806).	Death certificates.	No overall association between deaths from neoplasms and hydrocarbon-related occupations but possibly between kidney tumours and hydrocarbon-related occupations. An increased risk for children of parents in higher social classes.
Peters et al. (1981)	Parents of 92 children with brain tumours age under 10 at diagnosis identified from 1972-77 by the Los Angeles County Cancer Surveillance Program.	Parents of 92 children matched by sex, race, age and social class.	Telephone interviews.	Association between tumours and maternal exposure to chemicals, paternal exposure to solvents, and paternal employment in the aircraft industry.
McWhirten (1982)	Parents of 127 incident cases (aged under 15) of ALL which occurred during 1973-79 gathered by the Queensland Childhood Malignancy Registry.	Deomographic factors for the relevant area.	Hospital records.	Positive significant association between ALL and social class ($p = 0.001$).

Table 2.1 - continued

Authors/Year	Case Group	Control Group	Source of Information	Conclusions
Gold et al. (1982)	Parents of 43 children developing leukaemia from 1969-74 and children diagnosed with brain tumours from 1965-74 in the Baltimore Standard Metropolitan Statistical Area.	Two groups: Parents of children with no known malignant disease selected from birth certificates at the Maryland State Health Department and matched for the date of birth, sex and race. Parents of children with malignancies other than leukaemia or brain tumours and matched for sex, age at diagnosis, and race.	Interviews with mothers.	No relationship between parental occupation and cancer in the offspring. The only significant association was between paternal motor vehicle related occupations, for both leukaemia patients and and brain tumour patients and their cancer controls.
Hicks et al. (1984)	Parents of 298 cases of childhood cancer identified from the Texas Children's Hospital, treated for cancer during March 1976 - December 1977.	Three groups: Parents of children without cancer seen at the clinic during the same interval. Siblings of parents in case group. Neighbourhood parents of children who did not have cancer and who were the same ages as the case children.	Face-to-face or telephone interviews and clinic records.	Parents of children with cancer were no more likely to have worked in occupations with potential ionizing radiation exposure. Children with cancer more often had fathers who were aircraft mechanics ($p = 0.04$). Children whose fathers had military jobs with potential ionizing radiation had an increased cancer risk ($p = 0.01$). Specific cancer types occurred significantly more often among children of fathers in certain radiation-related occupations.
Shaw et al. (1984)	Fathers of 255 childhood leukaemia cases born after 1964 that were reported to the California Tumor Registry and diagnosed between 1975 - 1980.	Fathers of 510 children of the same sex as the index case whose birth certificates immediately preceded and followed the index case.	Birth certificates.	No relationship between paternal occupation and and childhood leukaemia.
Wilkins & Sinks (1984)	Fathers of 62 patients with Wilms' tumour identified through the Columbus (Ohio) Children's Hospital Tumor Registry.	Two groups: Fathers of children matched for sex, race and year of birth. Fathers of children matched for sex, race, year of birth and mother's country of residence when the child was born.	Birth certificates.	No statistically significant association was found between paternal occupations defined as hydrocarbon or lead related and Wilms' tumour in offspring.

Table 2.1 - continued

Authors/Year	Case Group	Control Group	Source of Information	Conclusions
van Steensel-Moll et al. (1985)	Parents of 519 children with ALL selected from nationwide register of childhood leukaemia in the Netherlands 1973-80.	Parents of 507 children matched for sex and birth date and same municipal area as case children.	Postal question-naires to parents.	Positive association between mothers' pregnancy occupations (hydrocarbon-related), RR-25 and 24. Mother pre-occupations and fathers pre-diagnosis occupations showed no relation.
Olshan et al. (1986)	Parents of 51 children with brain tumours diagnosed in western Washington State during 1978-81.	Parents of 142 children selected at random from western Washington State population.	Interviews.	Increased risk associated with fathers' employment in aero-space industry for children under 10. RR = 2.10
Wilkins et al. (1988)	Parents of 491 Ohio-born children who died from brain cancer during 1959-78.	Parents of 491 Ohio-born children selected randomly and matched for sex, race and year of birth.	Birth certificates.	Case fathers were found more likely than control fathers to have been employed in agriculture, in metal-related jobs, in the construction industry and in the machinery industry.
Johnson & Spitz (1989)	Parents of 499 children who died in Texas from from childhood nervous system tumours 1964 - 1980.	Parents of 988 controls randomly selected from Texas livebirths matched for age, sex and race.	Birth certificates.	The odds ratio for paternal employment in industries involving potential electromagnetic field exposure was 1.6 ($p<0.07$). A risk of 3.5 ($p<0.05$) was detected for fathers who were electricians.

Adapted from: Arundel and Kinnier-Wilson (1986)

during pregnancy and during lactation to soiled clothing brought home from the father's workplace may lead to chromosome defects. Thirdly, exposure of the child to the reported or continued presence of certain items in the household or prolonged contact with the father's contaminated clothing may cause direct carcinogenesis.

A review of the 17 studies contained in Table 2.1 produces conflicting evidence concerning the possible association between parental occupation and cancer in childhood. Of the 17 studies, 11 indicate an association and six found no significant correlation. Each was a case-control study conducted by similar method - childhood cancer cases were collected together and controls were selected according to sex, age, domicile or other properties, the paternal (and, in some reports, maternal) occupation was obtained and categorized and statistical analysis was carried out. There are, however, several minor differences within this basic methodology, for example, source of information, type of case, and categorization of parental occupation which must be considered when attempting to compare survey results (Arundel and Kinnier-Wilson, 1986).

It seems impossible to conclude whether or not there is a relation between parental occupation and childhood cancer, since these reports are not directly comparable and they vary widely in areas fundamental to the study. The source of information is probably one of the most important variables, and several of these reports used either birth or death certificates which are often inaccurate (Arundel and Kinnier-Wilson, 1986). This issue will be considered in Chapter Three in more detail. Peters et al., (1981), Hicks, et al., (1984), Zack et al., (1980), Gold et al. (1982) and Olstan et al., (1986) obtained their information by personal interview which is a more reliable source but not always possible. Greenberg and Shuster (1985) point out that reliance upon parental recall of exposure histories may lead to invalid associations because parents of children with cancer are highly motivated to remember any event which may help to explain the cause of their child's illness. If the

recall of parents of cases differs systematically from the recall of other parents, an invalid association might be observed.

Secondly, some studies had incomplete coverage with cases which were either lost or discarded (i.e. Hemminki et al., (1981), McWhirter (1982) and Hakulinen et al, (1976)) which could lead to bias in the results. Thirdly, some control groups were more carefully selected than others - some matched their controls only by birth registration whereas other matched for sex, age and social class. Fourthly, not all of these studies included the occupations of both parents. If there is an association between parental occupation and childhood cancer, then the occupation of the mother is just as important as that of the father (Arundel and Kinnier-Wilson, 1986).

There is a serious statistical problem in the way some studies have examined a wide range of parental occupations and cancer types, while others have concentrated on small numbers with specifically chosen cancer categories. There are also problems of differentiating occupational risks from social class and residential effects. Greenberg and Shuster (1985) conclude that it is difficult to draw firm conclusions from the research on parental occupations and cancer in childhood. Most of these studies, have involved broad categories of employment, with relatively little information on actual levels of exposure. Further they point out that occasional false positive associations may result when large numbers of separate occupations are tested. Arundel and Kinnier-Wilson (1986) conclude that considering the several potential cancer categories involved, all the many occupations and industries, and confounding factors such as drugs and smoking, it would appear that large numbers of cases and better information on both histological diagnosis and parental occupation exposure will be necessary if more definitive answers are to be obtained. Discussion in Chapter Three will deal further with issues relating to study biases.

The next subheadings of covariants discussed in the literature under the heading of environmental factors are those related to the socio-cultural environment: socio-economic status, urbanization, and other lifestyle factors.

Socio-Economic Status: The evidence from the literature regarding the relationship between socio-economic status and childhood cancer relates to both international comparisons and case-control studies. Neglia and Robison (1988) describe the evolution of a 3 to 5 year old peak of incidence of ALL which does not occur uniformly through the world, being conspicuously absent in Africa and many developing nations. The evolution of the 3 to 5 year old peak in Great Britain only developed during the 1920s and 1930s as it did in the United States, which has prompted speculation that this peak may reflect environmental exposures associated with modernization. Similarly, the authors note that the occurrence of this peak in whites predated its occurrence in blacks, possibly reflective of socio-economic factors.

Newell et al., (1989) describe higher rates of childhood Hodgkin's' disease occurring in developing countries and in ethnic minorities nationally. The authors claim that there is almost universal agreement that early onset Hodgkin's' disease is a consequence of a prevalent infection of low pathologenicity (as discussed previously). Exposure during childhood may be delayed in economically advanced societies with increased susceptibility in early childhood. The authors note that as living conditions improve, the total number of childhood cases tends to decrease.

Spitz et al. (1986) also report this finding and claim that Hodgkins' disease of childhood and early adolescence has often been overlooked because of its relative rarity in industrialized nations. The high rates in developing countries is explained as follows: exposure to an oncogenic infective agent during childhood may be delayed in economically

advanced societies, resulting in low childhood disease rates and concomitantly in increased disease susceptibility in early adult hood. Childhood rates appear to be higher among ethnic groups with a low socio-economic status.

Birch (1983) describing international variations in leukaemia states that it tends to be an upper social class disease and socio-economic differences may account for the lower incidence in the Manchester region of England. Socio-economic differences may also in part account for the low incidence of ALL in U.S. blacks, although it is possible that ethnicity plays a role here.

Li (1982) states that international studies show correlations between the Hodgkins' disease rate and level of socio-economic development in an area. Better socio-economic conditions are generally associated with lower Hodgkin's disease rates in children and correspondingly, a higher rate among young adults. One hypothesis is that factors associated with socio-economic status, such as sibship size, hygiene, and medical practise patterns, influence age at exposure to etiologic agents, which are as yet unidentified. Analogies have been drawn to epidemiologic patterns of poliomyelitis, e.g., improved living conditions which result in infection at a later age and pose a greater risk of clinical disease.

In addition to the evidence from international comparisons, several case-control studies have been published on the relationship between socio-economic status and risk of cancer among children. Greenberg and Shuster (1985) review seven such studies which have dealt exclusively with leukaemia. They note a remarkable consistency of results across the various study locations and periods. With only one exception, all investigators found a positive association between high socio-economic status and increased risk of leukaemia among children. The indicators of socio-economic status used in these studies on leukaemia and additional ones on other cancer sites included: total family income,

median monthly rent, paternal occupation, residential area and education level. With forms of cancer other than leukaemia, there is less information on the relationship. Weak associations were noted for brain and Wilm's, but competing causes of death and misdiagnosis of cancer in children of lower socio-economic status could not be excluded as possible explanations for the observed socio-economic association.

The only evidence to date of a cancer in childhood that is associated with low socio-economic status according to Greenberg and Shuster (1985) is from a case-control study of rhabdomyosarcoma by Grufferman et al. (1982). They found that both mothers and fathers of cases had less education than parents of controls. Also, total family income for cases was less than for controls.

Neglia and Robison (1988) in a study in Baltimore, describe an increase in the incidence of leukaemia in black children. The striking increase observed in the incidence in blacks was most pronounced in those individuals within the higher socio-economic status categories. This observation has led to speculation that the increase may be the result of environmental influences associated with an upward socio-economic mobility in the black population in Baltimore during recent years.

The studies reviewed previously regarding parental occupation also provide some evidence of the effect of socio-economic status on childhood cancer. There are problems of differentiating occupational risks from social class and residential effects. McWhirter (1982) found that there was a positive significant association between ALL and the upper social classes and a negative correlation with lower social classes. Since the same trend was seen in the whole state of Queensland and the small area of Brisbane City, the results must be attributed to population characteristics rather than geographical factors. Arundel and Kinnier-Wilson (1986) offer several possible explanations for this: first, factors associated with differences in lifestyle among the various social classes may increase or

decrease the risk of developing ALL; secondly, it has been suggested that children in a pre-leukaemia phase have an increased tendency to infections, and it is possible that the children in lower social class families with higher risk of infections die before their leukaemia is recognized; or thirdly, it may be an artefact, though this appears less likely when the findings of Sanders et al., (1981) are also considered. These researchers found the most strikingly high proportional mortality ratios were for cancer deaths in children of social classes 1 and 2 (high), whereas the lowest values were among children whose fathers were in the social class 5 (low).

Urbanization: Neutel (1970) states that in studies examining the urban-rural distribution of childhood leukemia, the consensus is an increase in the risk of leukemia in urban areas. Greenberg, Preuss and Anderson (1980) found that for childhood leukemia, an indicator measuring traffic density consistently demonstrated correlations at the same level or higher than the control variables and the remaining risk factors. This clue is considered plausible because benzene, a substance capable of causing leukemia, may be two percent or more of gasoline. Neglia and Robison (1988) note that the increased death rates from leukemia in the Northwestern agricultural states have provided some of the motivation for studies assessing farm practises and the occurrence of leukaemia and lymphoma, especially given the potential exposures of farmers and their families to herbicides, fertilizers and pesticides.

Greenberg (1983) found, when investigating urbanization and cancer mortality in the total U.S. population, that urban areas have been characterized by higher overall population cancer mortality rates then rural areas. This author then set out to see if this conclusion was also true for the youngest populations. Greenberg (1983) prefaces this analysis by commenting that there are some reasons to suspect that there may not be an

urban-rural difference in child and teenage cancer. First, the major types of cancer afflicting children and teenagers exhibited relatively small urban-rural differences for the entire population. Second, child and teenage cancer mortality rates have been decreasing, which suggests different etiologies as well as improved survival rates for the young. Third, initial epidemiological studies suggest that rural as well as urban occupations are associated with higher child and teenage cancer incidence. Greenberg's findings were that strongly urban countries as a whole have higher rates of child and teenage cancer mortality rates than the entire nation. He comments that it is particularly interesting to note that the ratios for strongly urban areas are sometimes higher for 0 to 19 year old populations than for the total population. Further, within the group of most urbanized areas, regional differences in child and teenage cancer mortality could be identified.

Other Lifestyle: Lifestyle, or behavioural variables such as diet and smoking, have also received attention in the literature as risk factors for childhood cancer. Lenarsky (1983) lists the following dietary factors in the etiology of childhood cancer, by agent and source: aflatozins and tannins - natural constituents; nitrosamines - reaction of salivary nitrates with secondary or tertiary amines; polycyclic aromatic hydrocarbons - smoking, or grilling meat or fish, also from air pollution; artificial sweeteners and food dyes - food additives; pesticides and fertilizers - food contaminants; and overnutrition. Miller (1978) states that it has been difficult to derive strong evidence that individual components of the diet are carcinogenic in man (let alone children) because of the long latent periods involved, the role of metabolic conversion and possible interactions, which may be potentiating or inhibiting.

Regarding chemical contaminants (pesticides, PCBs, dioxin), Miller (1978) states that some of these chemicals are deposited in adipose tissues and are removed from the

body primarily in the fat of breast milk. This circumstance raises the possibility that in rare instances, when the mother has been heavily exposed to one of these agents, breast feeding may transmit the chemicals to the infant at a time when susceptibility is high and life expectancy is long, allowing ample time for development of cancers with latent periods of 50 years or more. No evidence exists of the contribution this would have to malignancies that manifest themselves in childhood. Birch (1983) discusses dietary and smoking habits during pregnancy as potential sources of transplacental carcinogens. He cites studies in laboratory animals, where the most potent transplacental carcinogens are among the nitroso-compounds, and asserts that an individual carcinogen may be considerable more potent in the fetus than in the adult. A low dose which is not capable of producing tumours in the adult can be carcinogenic in the fetus. Nitroso-compounds occur in many foods, in urban air and in cigarette smoke. They may also be synthesised endogenously from nitrite and amines or amides present in food and drugs. Although dietary factors may be of importance, the complex nature of mechanisms of carcinogenesis and the likelihood that very low doses of a combination of compounds would be involved renders that task of identifying such factors extremely difficult.

It has been suggested that cancer could result from transplacental exposure to maternal smoking (Lenarsky, 1983). Studies in human subjects have shown that components of tobacco smoke can reach the fetus and that human fetal tissues are capable of activating carcinogens similar to those in tobacco smoke. Fetal exposure to maternal smoking, therefore, could provide all the elements necessary for transplacental carcinogenesis in humans (Lenarsky, 1983). Neglia and Robison (1988) state that a 50 percent increase in risk for children exposed to cigarette smoke in utero has been reported, however, several large studies have failed to document this association. Neutel (1970) states that cigarette smoke is capable of inducing a placental enzyme which will metabolize

a potent carcinogen, therefore it is reasonable to postulate that smoking in pregnancy might increase the risk of childhood cancer.

Ecogenetics: Although genetic factors (resulting from an individual's genetic constitution) and environmental factors (arising from outside the individual) were considered separately in the previous review, all cancers certainly derive from an interaction between the two (Birch, 1983). Thus, any study of the etiology of pediatric cancer must take into account the interaction of genetic and environmental factors (Birch, 1983).

Kramer (1983) states that although for some childhood tumours specific genetic and environmental factors have been identified as being associated with high risk, for most neoplasms the etiologies are unknown. Investigators must recognize that, in most cases, both heredity and environment contribute to tumour frequency. However, one or the other category of risk factors may clearly be more influential, such as occurs with Ewing's sarcoma (genetic factors predominating) and Burkitt's lymphoma (environmental influences).

The importance or contribution of the role of the environment is the subject of disagreement by authors. Mulvihill (1982) suggests that why a child has cancer is usually unknown. As in cancer therapy, progress in cancer etiology has been made in recent years. The cliché that 80 percent of cancers are due to the environment was developed for adult tumours and does not withstand critical examination when applied to the 1 to 2 percent of malignancies that occur in those younger than 20 years of age. Whatever this evidence is for adult tumours, it is nearly absent with regard to childhood cancer. On the other hand, for many childhood tumours, genetic, congenital or familial determinants are conspicuous. According to Mulvihill (1982), considering cancer as either genetic or environmental in origin is a polarizing simplification. Any one tumour is likely to have many determinants;

hence, the appropriate conceptualization may be "ecogenetics". Analogous to pharmacogenetics (the study of genetic variations in response to drugs), ecogenetics means the study of genetic variations in response to any environmental agent. Although examples of ecogenetics in childhood cancer are few, they provide the models for understanding all human cancer through intensive study of just a small number of patients (Mulvihill, 1982).

According to Greenberg and Shuster (1985), there is sufficient existing evidence to conclude that the risk of neoplastic disease during childhood is influenced by genetic as well as environmental factors. Inherited susceptibility to cancer in children has been demonstrated for a variety of single gene disorders. These syndromes, however, account for a relatively small proportion of all malignancies among young persons. For most cancers in childhood, the contributions of hereditary factors are not well understood.

Greenberg and Shuster (1985) further state that a number of environmental risk factors for malignancies in children have been proposed. Some of these factors, such as prenatal infection with influenza, are commonplace, and thus the proportion of cancers related to such an exposure may be substantial, even if the corresponding estimated risk ratio is modest. At the same time, it must be recognized that these weak associations may not involve causal relationships. The most convincing evidence of environmental causes of cancer in children exists for factors with strong estimated risk ratios, such as prenatal exposure to DES or postnatal radiotherapy. As indicated in this review, very few strong environmental determinants for cancer in childhood have been characterized.

In considering the risk factors identified to date, Greenberg and Shuster (1985) speculate that many cancers of children develop through unusual host responses to common stimuli. Further, the authors observe that the concept of host susceptibility to cancer in children has not received adequate attention. They call for future analytic research on the epidemiology of these diseases which includes careful evaluation of effect

modification by such cofactors as age at exposure, gender, ethnicity, and family history. Ultimately, characterization of these host-environment interactions will provide a better understanding of the causes of cancer in children.

Lenarsky (1982) also comments on the importance of host-environment interactions. The development of a cancer may result from a multi-stage process. Initial and progressive stages of carcinogenesis may be modified by host and environmental factors. Genetic factors may alter susceptibility to the effects of potential carcinogens due to variations in metabolism. Lenarsky (1982) suggests that continued study of the host factors which increase susceptibility to malignancy and the interactions with environmental agents will lead to a better understanding of carcinogenesis.

2.3.3 Mortality Risk Factors

The previous overview has dealt with the range of genetic and environmental risk factors that have been associated with childhood cancer in the literature. These covariants have been considered in terms of their potential role in childhood cancer incidence - as determinants in "causation", not necessary in "dying from" the disease. However, the contribution of previously discussed factors specifically to patterns of childhood cancer mortality must be considered. Spatial differences in mortality rates are the result of higher prevalence of the disease in some areas **and/or** differences in rates of mortality (varying survival rates). The spatial pattern of interest, then, when analysing mortality data specifically, is the geographic distribution of **deaths** from a cause, not just of the cause(s) itself. The range of socio-demographic variables previously considered under causation, then, may play a role in mortality rates. In addition, differences in reporting and diagnosis may create spatial variations in mortality rates.

Bunin (1987) discusses the impact of both differences in diagnosis and in access to medical care with regard to childhood brain cancer. The author notes that black children in

the U.S. are a group whose lesser access to tertiary medical care may lead to lower rates of microscopic confirmation of brain cancer - thus both lower reporting of this cause of death and more incidence resulting in mortality due to late treatment. Variations in cancer mortality, then, may vary with access to medical care.

Schneider et al. (1990) examine pediatric cancer mortality rates among U.S. states to determine whether the differences in cancer mortality rates for the total population among states, noted to be decreasing in the literature, apply to childhood rates. Reasons they found in the literature to explain this convergence of rates for all populations include urbanization of formerly rural areas, increasing homogeneity of culture, interregional migration, better reporting and accuracy of incidence and mortality data in formerly rural areas, and the declining importance of European-born Americans and their cultural habits. Schneider et al. (1990) wanted to see if this convergence of rates is also present for pediatric cancer. They found a failure of New Jersey pediatric rates to converge with those of the U.S., and present a number of arguments to explain this. They conclude that based on demographic data, this could not be explained by ethnicity or immigration. An economic argument was made, citing the fact that crowding, poverty and lack of education make people less cognizant of the early warning signs of cancer. Economically disadvantaged individuals seek medical attention for their children later than those of higher socio-economic status and they may be less compliant with therapies, resulting in higher mortality rates. Schneider et al. (1990) state that in general, they believe that high socio-economic status brings with it better education about the early warning signs of cancer. This information contributes to higher diagnosis and reporting rates, an explanation for the relatively higher rates among wealthier New Jersey residents.

An environmental argument is also made, that chemical exposure, air and ground water contamination, and/or parental occupation contributed to higher pediatric cancer

mortality rates. Yet, Schneider et al. (1990) state that many industries, including heavy polluters, have left New Jersey and other north-eastern states. According to these authors, the result of this redistribution of sources of contamination should lead to a convergence of rates - but it has not .

Another explanation offered by these authors is that urban places historically have more physicians than rural places. With this increase in physicians comes more and better diagnostic equipment. Urban dwellers are more likely to die in the hospital and more likely to have a diagnosis confirmed by a physician than are those with less access to medical care whose death may be certified by a local coroner who may not have had any medical training. This is used to explain why urbanized places like New Jersey have historically had better reporting of cases (1950-59), although urban/rural differences may not be as significant presently.

If differences in diagnosis do not account for the differences in mortality rates, the possibility must be considered that the disease is truly more prevalent and/or more deadly in New Jersey, or areas which exhibit higher rates (Schneider et al., 1990). Thus two sets of covariants of childhood cancer must be considered: those genetic and environmental factors which are risk factors for incidence **and** those environmental (specifically socio-economic) factors which may increase or decrease mortality rates, irrespective of incidence.

2.4 Conclusion

This chapter has provided an overview of the range of factors which have been considered as potential correlates of childhood cancer. Classified under three broad headings, these correlates are: demographic and genetic factors; environmental factors; and, mortality risk factors. The literature is, however, inconclusive regarding the relative contributions of these different sets of factors to childhood cancer mortality rates. It

appears, based on the strength of evidence provided in the literature, that genetic predisposition is fundamentally important as a determinant of childhood cancer. In addition, cancers which develop as a consequence of environmental exposures are seen to occur after a latent period of some 20 to 40 years, thus apparently ruling out environmental factors as a cause of cancer in children. However, given the acceptance of the "two-hit" model of cancer induction which postulates that children born with a genetic predisposition to certain forms of cancer require only a second event to develop malignancy, it seems plausible that the risk of developing certain forms of cancer within the pediatric age-span, given a predisposition, may be increased further by early exposure to an environmental carcinogen. An example to illustrate this can be found in studies of childhood leukaemia, which show that genetically predisposed children living in close proximity to nuclear power facilities may be at elevated risk of developing certain forms of leukaemia (Mulvihill, 1982). In addition, socio-demographic variables which may influence access to medical care and differences in reporting, have been seen to affect childhood cancer mortality rates, irrespective of incidence.

Thus, it seems clear that in addition to hereditary and genetic factors, environmental and socio-demographic factors may be important and thus deserve further analysis. As these environmental and socio-demographic factors are known to vary geographically, a study which examines spatial variations in childhood cancer mortality is an appropriate design for investigating cancer in this sub-population in Ontario. The ecologic study design is, therefore, the focus of attention in the next chapter.

CHAPTER THREE

METHODS

This chapter will focus on two methodological issues of relevance to the Ontario study of childhood cancer mortality: ecologic analysis and mortality data. An assessment of ecologic analysis — the methodology employed in this thesis — will reveal its inherent strengths and weaknesses and its applicability to the analysis to follow. There are also special considerations with regard to mortality data that must be addressed in light of their use in the Ontario study.

3.1 Ecologic Analysis

The unique distinguishing feature of ecologic studies is that they are empirical investigations involving a group of individuals as the unit of analysis (Walter, 1988). Typically, the group or population is defined by some geographic boundary such as a state, county, or census tract (Morgenstern, 1982). The data may be limited to boundaries that have been constructed for administrative purposes but for which the data can subsequently be used as a source of information for epidemiologic research (Reynolds, 1989). Ecologic analysis has proven to be a very productive method in chronic disease epidemiology. In an ecologic survey, disease rates (mortality or incidence) are tabulated for defined geographic units and an attempt is made to relate the pattern of those rates to possible etiologic factors, such as socio-economic or environmental characteristics (Savitz and Redmond, 1985).

The two main aims of the ecologic design in epidemiology are the generation/testing of etiologic hypotheses , i.e. to explain patterns of disease occurrence; and the evaluation of health interventions, i.e. to test the application of our knowledge for preventing disease and

promoting health (Walter, 1988 and Morgenstern, 1982). The value of the ecologic study design is limited. Although relationships discovered between aggregate variables are judged by epidemiologists to be inadequate for hypothesis testing, the approach does provide valuable etiologic clues to be pursued in studies of individuals (Savitz and Redmond, 1985). Cleek (1979) states further, since much of this work is designed to generate hypotheses rather than to rigorously test associations, and so long as no one jumps to conclusions about individual-level relationships, there is no severe problem .

Three main types of ecologic study design are identified in the literature: exploratory studies of spatial variation, multiple-group comparison studies of spatial variations involving exposure data and time trend studies (Morgenstern, 1982 and Walter, 1988). The exploratory study is one in which geographic differences in disease rates among several regions or groups are observed. The objective is to search for spatial patterns that might suggest an etiologic hypothesis. Walter (1988) states that this type of study usually examines the spatial variation in disease rates, but without any direct incorporation of exposure information. Typical examples cited are investigations based on cancer atlases where the incidence rates for cancer sites of interest are examined for evidence of spatial autocorrelation/clustering, which might be related to possible environmental exposure variables. The analyses may be informal 'eyeball' assessments of the maps, or could involve formal statistical tests for spatial autocorrelation. Because exposure information is not directly incorporated into the analysis, this type of study is usually hypothesis generating rather than hypothesis testing (Walter, 1988).

Examples of exploratory ecologic studies in the childhood cancer literature include: Greenberg (1983), which examines the pattern of childhood cancer mortality in the United States as it relates to patterns of urbanization and Craft and Openshaw (1985), which analyse reports of childhood leukaemia clusters.

The multiple-group comparison ecologic study design observes the association between the average exposure level and the disease rate on a group basis for several regions (Morgenstern, 1982). A typical example of such a design would involve measurements of the hardness of the water supply in a number of communities, and the corresponding mortality rates from ischemic heart disease. The objective of the statistical analysis is then to decide if any association between exposure and health outcome is statistically significant and substantively meaningful, allowing for possible bias or confounding. The preferred analysis for this kind of data is regression, rather than correlation. Regression allows the estimation of relative risk associated with changes in exposure; under ideal circumstances this relative risk will be the same as that which would have been estimated in individually linked data (Walter, 1988).

There are numerous examples of multi-group exposure studies in the literature, where the geographic unit of analysis ranges from census tract to parish, town, catchment area, city, SMSA's, county borough, county to country (Walter, 1988). A typical example from the childhood cancer literature is a study by Greenberg, Preuss and Anderson (1980), which examines 71 hypothesized factors of elevated and reduced risk of childhood cancer (i.e. traffic density and leukaemia) for counties in five U. S. states.

In the time trend study design, single or multiple populations are assessed with respect to changes in average exposure level and changes in the disease rate over time (Morgenstern, 1982 and Walter, 1988). An association between exposure and disease outcome would be indicated if upward and downward changes in exposure are paralleled by similar changes in outcome (Walter, 1988). The study of childhood cancer mortality in Ontario to follow employs aspects of the exploratory ecologic design of spatial variation and the multiple-group design involving exposure data - time trend analysis is not used.

3.1.1 Advantages

Although the value of the ecologic study design is limited, certain advantages serve to make it an attractive methodology. The main advantages of the ecologic approach are that it facilitates the study of very large populations, at relatively low cost, using existing sources of data. A considerable increase in cost-efficiency is realized, as compared to designs where individual data are required because exposure and outcome information is used on a group basis (Reynolds, 1989). Data on individuals is often difficult to obtain, and surveys are expensive involving many difficult questions of sampling procedures. Typically, then, social scientists have analysed data at the level of some geographical unit, in terms of what it tells us about individual behaviour (Langbein and Lichtman, 1978).

The ability to use existing databases makes the ecologic design very useful under certain circumstances. For example, if exposure data, such as water quality, is routinely available in a particular geographic area, and if disease outcomes, such as cases of cancer, are routinely recorded in a registry, then the two sources of data may be used directly, without the necessity of contact with individual population members (Walter, 1988). The use of aggregate-level data has a tremendous advantage over the use of individual-level data in that it is readily available at many levels of aggregation, from city to county, state and nation, for many diseases and possible independent variables (Cleek, 1979). Because they can often be done by combining existing data files on large populations, ecologic studies are generally less expensive and take less time than studies such as the case-control, which involve the individual as the unit of analysis.

Ecologic studies can also achieve certain objectives generally not met with nonecologic designs (Morgenstern, 1982). The large amounts and types of aggregate data that exist make it especially well-suited to the generation of hypotheses. Similarly, the extensive manipulation necessary to unravelling the more complex disease relationships is facilitated by the use of aggregate data. There are simply many more possible control

variables available in the aggregate data than would be gathered in an individually-measured prospective or retrospective study (Cleek, 1979). Because large populations can be studied using ecologic designs, one may investigate relatively small increases in risk (Walter, 1988). Another advantage of the ecologic approach is its usefulness in the investigation of suspicious clusters of disease in relatively small geographic areas (Walter, 1988).

3.1.2 Disadvantages

There are several significant disadvantages of the ecologic approach, which often lead researchers to dismiss potentially important results without further investigation (Reynolds, 1989). This section will discuss three problems identified in the literature with respect to ecologic analysis - ecologic bias, choice of areal units, and multicollinearity - as well as some ways to minimize bias.

Ecologic Bias: Ecologic bias refers to the failure of aggregate-level associations to properly reflect individual-level associations (Greenland and Morgenstern, 1987). The major limitation of ecologic analysis for testing etiologic hypotheses is the potential for substantial bias in effect estimation. The central problem, known as the "ecological fallacy", results from making causal inference about individual phenomena on the basis of observations of groups (Morgenstern, 1982). The ecological fallacy was first demonstrated mathematically in 1950 by William Robinson, who popularized the term "ecological fallacy" to describe any incorrect inference about individual behaviour from grouped data (Langbein and Lichtman, 1978).

Two forms of ecologic bias have been identified in the literature - aggregation bias which results from the grouping of individuals, and specification bias, which is due to the confounding effect of the group itself. Aggregation bias occurs because as people are grouped at some level, i.e. a county, the grouping process is not likely to be random

(Cleek, 1979). Because the joint distribution of exposure and health remains unknown, an insignificant association between exposure and disease at the individual level can be inflated at the aggregate level in part because the aggregation process operated on the basis of the independent variable (Cleek, 1979). It is also possible that two variables which are correlated at the individual level show no association when studied in aggregate data (Walter, 1988). This possible distortion of the association between exposure and disease outcome is one of the major disadvantages of the ecologic methodology.

Specification bias can result from the actual composition of the group chosen for study. Certain extraneous risk factors can be differentially distributed within the group or some property of the group itself such as social organization, can affect disease occurrence. The grouping of individuals may alter the relative variance of independent and dependent variables, thereby affecting the values of correlation coefficients (Langbein and Lichtman, 1978).

The sum of these two distinct causes of discrepancies is called cross-level bias which can make ecologic association appear stronger or weaker than it is at the individual level. Morgenstern (1982) points out that aggregation bias and specification bias may in some cases cancel each other out, resulting in no bias in the estimation of effect. However, the author further states that cross-level bias ordinarily exaggerates the magnitude of the true association. Grouping variables does not always result in less precise measurements of the effects of one variable on another. If the relationship is properly specified, there may be no bias at all and in some cases, there may be a gain in precision by aggregation procedures (Langbein and Lichtman, 1978). For a given statistical model at the individual level, there can only be specification bias (confounding) while for the same model at the ecologic level, there can be both specification and aggregation bias (Morgenstern, 1982).

However, studies which use aggregate data at any level, run the risk of misrepresenting the individual level relationships - the ecological fallacy problem (Cleek, 1979).

The problems of inference inherent in the use of aggregate data have led some researchers to be quite critical of the ecologic method, disavow its use and advocate the use of individual level data only. The advantage with individual data is that the researcher is closest to the process of interest, the presence or absence of disease. However, the advantages of aggregate data, discussed in the preceding section, would seem to outweigh this disadvantage (Cleek, 1979). Walter (1988) further states that most epidemiologists would agree that it is generally preferable to use a nonecologic design, if this is feasible. However, if an ecologic design is selected, it requires considerable attention to methodologic rigour, in order to minimize the potential ecologic fallacy problem.

3.1.3 Choice of Areal Units

The choice of areal units can also influence the likelihood of a positive result as correlations generally increase as the data are aggregated into larger units (Reynolds, 1989). Depending on the scale, individuals are more or less likely to be grouped into aggregations on the basis of the dependent or an independent variable of interest. Specification bias, discussed previously, can result from the use of areal units containing individuals with similar characteristics, i.e. counties are often aggregations of individuals with similar socio-economic status (Cleek, 1979).

Stavraky (1976) points out two major problems in studies using ecologic analysis relevant to the choice of areal units. The quality of areal data is questionable in cases of cancer research where a 20 year latency requires the use of 20 year old data. The author points out that this is more important if relevant factors have changed in frequency over time. Another issue raised is the problem of obtaining disease frequency data and

environmental variables for the same areal units or same population sub-groups within these units. Changing administrative boundaries over time can create complications. In addition, researchers are often forced into the choice of an areal unit for study by the availability of incidence or mortality data and then need to adapt other data to this areal unit.

The second problem discussed by Stavraky (1976) deals with the adequacy of statistical methods used in ecologic analysis. Correlation and regression results are likely to differ depending upon the areal unit chosen. In general, correlation increases with the aggregation of data into larger units. This doesn't imply that larger units are more appropriate - by virtue of their greater heterogeneity they may be less suitable. In addition, since contiguous areas are likely to be similar in many respects, the assumption made in using a single or multiple linear regression model that the error terms are independent and normally distributed, may not be satisfied using areal data thus researchers must exercise caution interpreting results. This problem has been termed autocorrelation, and as of yet, no satisfactory statistical analyses have been developed and standardized to circumvent this problem (Reynolds, 1989).

Stavraky (1976) further points out that in any study exploring many relationships, the risk of finding some correlation co-efficients significant purely by chance must be kept in mind, thus one should emphasis the magnitude of the coefficient rather than statistical significance. Reynolds (1989) concludes that in such a case where the size of the geographic unit is in question, consistency among varying sized units would lend support to the findings thus conclusions derived at one geographic level may not be applicable to other levels.

3.1.4 Multicollinearity

Another problem with ecologic analysis is that certain predictor variables (especially sociodemographic and environmental variables) tend to be more highly correlated with each other at the aggregate level than they are at the individual level, a phenomenon called multicollinearity (Morgenstern, 1982). This implies that it is more difficult statistically to estimate the contribution of each factor, adjusting for possible effects of the other correlated or collinear risk factors. Multicollinearity between exposure variables is usually stronger at the ecologic level than at the individual level. This is because in the ecologic analysis, each geographic subunit is assigned a single value for each exposure variable, whereas there is in fact variation in exposure within the ecologic subgroup, usually such that the correlation between exposures within a subgroup is not 100 per cent (Walter, 1988).

Other problems with ecologic analysis involve migration and latency. An additional problem using the ecologic approach if latency is to be taken into account is that individuals may migrate over time, and hence be exposed to a variety of exposures (Walter, 1988). A problem specific to cancer is the substantial lag time between a carcinogenic stimulus and the occurrence of clinical disease. Because of this, aggregate causes must predate the observed disease rates by a number of years. The passage of time would make the linkage between some population characteristics and disease rates more tenuous, since the actual composition of the population changes as people migrate into and out of the area (Savitz and Redmond, 1985).

3.1.5 Minimizing Bias

One of the first steps to minimize bias is the use of regression techniques over correlation in order to more accurately assess the magnitude of the associations under study (Reynolds, 1989). The use of ecologic regression - not correlation - to estimate the magnitude of the desired association is frequently indicated in the literature. In the situation

where groups tend to be homogeneous with respect to one of the independent variables, ecologic regression coefficients, but not ecologic correlation coefficients, will result in unbiased estimates of their corresponding individual measures (Morgenstern, 1982). With the use of ecologic regression, as many risk factors as possible should be included in the statistical model (Morgenstern, 1982). Careful application of ecological regression is likely to be the most promising strategy for social scientists faced with problems of inferring individual behaviour from aggregate data. Other strategies are either logically flawed or have limited application to practical problems (Langbein and Lichtman, 1978).

A second way to minimize bias is to use moderately sized aggregated groups in order to make the groups as homogeneous as possible. County-level populations are the preferred unit of study since they are generally small enough to ensure some degree of homogeneity but also allow for rates of cancer to be fairly stable (Reynolds, 1989). By using smaller groups, the investigator may also increase the distorting effects of migration between regions. Furthermore, the use of smaller units of analysis tends to make the estimates of disease frequency for groups less stable, which may become a serious limitation, since most diseases are 'rare' events (Morgenstern, 1982).

A third way to minimize the problems of ecologic inference is to assess how the groups were, in fact, formed and analyse the data accordingly. Investigators should include in the model all variables thought to be related to the grouping process, even if these variables are not risk factors for the disease (Morgenstern, 1982). The goal of avoiding aggregation bias frequently requires including more variables in the aggregate model than are incorporated in the individual model it represents. County level data represents means or proportions computed for those individuals residing in each county. Aggregate models may be comprised both of variables which are theoretically relevant at the individual level and of variables which are added in order to remove the bias caused by

grouping, i.e. variables that reflect the grouping process itself (Langbein and Lichtman, 1978).

The findings of an ecologic analysis can be compared with the findings of other observational studies designed to test the same etiologic hypothesis to seek consistency of results (Morgenstern, 1982). Thus, if these considerations are taken into account, the ecologic design can yield useful hypothesis-generating results.

3.1.6 Conclusions

The ecologic methodology, with its primary aim of hypothesis generation, is a valuable tool in geographic studies of disease. Its cost-efficiency, ability to use existing data sources and study relatively large populations make it an attractive choice under many circumstances. An understanding of the limitations and potential sources of bias inherent in the ecologic approach enable researchers to use this method effectively. In the context of this thesis, the ecologic methodology is the most appropriate choice given both the availability of data, and the fact that it represents the initial study of childhood cancer mortality patterns in Ontario, 1976-1985.

3.2 Mortality Data

Cancer mortality statistics are an integral part of many epidemiological investigations of cancer. Mortality data are commonly used to indicate trends in cancer occurrence, in some instances because they are the only data available and in others because they may be considered preferable to the available cancer incidence data for this purpose (Devesa et al., 1984). The number of recorded deaths can be defined as a measure of mortality, whereas incidence is the number of cases occurring during a specific time period (Greenberg, 1983). The practise of using cancer mortality rates across geographical areas

to draw etiological inferences assumes that spatial variations in cancer mortality rates reflect similar patterns in incidence. This section will examine the nature and applicability of cancer mortality data in epidemiological research, its relationship to incidence, the advantages and disadvantages of mortality data, including the use of death certificates, migration and latency issues, and designs for producing better mortality studies.

Deaths from cancer are used in geographical studies, studies of time trends, correlation studies, and therapy evaluation as well as to identify cases for retrospective evaluation of possible etiologic factors (Percy et al., 1981). Availability is the overriding advantage of mortality data - in the absence of sufficient incidence data, mortality counts are the only option (Greenberg, 1983). Because mortality data are often much easier to collect and are generally more reliable than incidence data, the former are sometimes used in place of the latter to generate and test etiologic hypotheses, especially with highly fatal rare diseases (Kleinbaum et al., 1982).

Cancer registries are the most commonly utilized source of mortality data - a detailed discussion of cancer registries is contained in Chapter Four. Because of the varied uses of mortality data, it is important to consider the reliability, accuracy and suitability of its use in such investigations. According to Greenberg (1983), mortality closely reflects incidence for quickly fatal tumours but treatment has improved for many types of cancer (i.e. leukaemia and lymphomas) and in these cases there is much less association between mortality and incidence and, in turn, less association with local etiological factors. Because mortality rates are a function of both disease incidence and the case-fatality (survivorship) rate, the assumption that spatial variations in cancer mortality rates and incidence rates are highly correlated may not hold true if there are systematic differences in survival rates (Horner and Chirikos, 1987). Survivorship differences could arise from dissimilar geographical patterns of patient characteristics which potentially influence the course of the

disease, or from the direct effects of residency on that course. Differences in survival may also be related to the effects of geography on the opportunity for competing causes of death to influence site-specific mortality rates. Horner and Chirikos (1987) further assert that survivorship differences across areas are potentially important confounders of geographical comparisons of cancer mortality and, accordingly, deserve more attention than they have thus far received in the literature.

The 1987 paper by Horner and Chirikos examines whether survival rates of cancer patients are differentially distributed across geographical boundaries thus confounding geographical comparisons of cancer mortality. Their results indicate that for most cancers, survivorship is not likely to play a confounding role in geographical comparisons of adjusted cancer mortality rates. This finding, according to the authors, provides support for the continuing use of mortality data as a relatively fast and inexpensive method of identifying possible etiologic factors in preliminary epidemiological investigations.

When mortality is chosen as the outcome measure of interest in cancer studies, an understanding of the limitations and disadvantages of this data is necessary. Greenberg (1983) identifies three disadvantages of mortality data: diagnosis and recording errors in death certificates, population migration and the long latency periods of most cancers. The subject of missclassification of the cause of death on death certificates has received considerable attention in the literature. Death certificates are the most widely used source of information for mortality statistics, often serving as the only source of data for estimates of the incidence and prevalence for many diseases. Yet, the inaccuracies of cancer death certificates are frequently mentioned. Percy et al. (1981), conducted a study to determine the accuracy of cancer mortality data by comparing death certificates with an underlying cause of death of cancer to the hospital diagnosis for resident cases of cancer. The authors found that the underlying cause of death as coded on the death certificate was accurate for

about 65 per cent of the cancer deaths in the study. Certain cancer sites such as colorectal and brain cancers were particularly subject to misclassification problems. Most of this was attributed to physicians tending to report a non-specific site of cancer on the death certificate rather than the specific site identified in hospital diagnosis.

In a study conducted by Engel (1980), accuracy of death certification within an autopsied population in the United States is examined. The author found that the confirmation rate from autopsies for original death certificate diagnoses was 89 per cent. However, an underreporting of 10 per cent for cancer, and missclassification as to primary site of cancer in 28 per cent of the cancers was found. Both of the studies previously discussed indicate missclassification of the specific site of cancer as the major problem with death certificate data. In studies of childhood cancer, however, it is not known how relevant this might be. Small numbers of cases invariably lead researchers to study aggregated sites which would minimize the problems of misclassification.

Regarding the validity of the use of mortality data to indicate trends in cancer incidence, factors other than misclassification have been discussed. Devesa et al. (1984), state that the use of mortality data may not adequately facilitate the analysis of trends in cancer occurrence to gain information concerning possible changes in the prevalence of risk factors for several reasons. First, for some cancer sites, survival rates are so high that mortality rates greatly underestimate incidence rates and may not consistently or quickly reflect incidence rates. Further, changes in survival will alter mortality rates regardless of changes in cancer incidence rates. In addition, inaccuracies in death classification is also mentioned. On the other hand, cancer incidence data, which provide a measure more directly related to etiologic factors, are subject to problems of interpretation over long time periods, comparability of the data and a number of other possible artefacts.

A study conducted by Devesa et al. (1984), examines cancer incidence data and corresponding mortality data in the United States in an attempt to evaluate consistency among age groups and geographical areas and determine which data is most reliable. The findings of the study were that mortality trends may be influenced by shifts in diagnostic specificity on the death certificates and by changes in survival rates, and both incidence and mortality rates may be influenced by increased casefinding, improvements in diagnostic procedures, expansion of the medical care delivery system, and real changes in the prevalence of risk factors. The authors conclude that both incidence and mortality data should be used when attempting to assess trends in cancer occurrence because sole reliance on one or the other may lead to erroneous conclusions. In the context of investigations of childhood cancer clusters, Mathews (1988) notes that spatial and temporal comparisons of incidence data and mortality data may be misleading if registration bias for either onset or death is present, if there are differences in treatment which affect mortality, or if migration from place to place occurs between the time of diagnosis and the time of death.

Population migration and latency between exposure, occurrence and death are other disadvantages of the use of mortality data. Greenberg (1983) suggests that migration and latency problems should be considered together because the longer the latency period the greater the effect of migration. A study by Polissar (1980) examines the effect of migration on comparison of disease rates in geographic studies. The author discusses a lack of sensitivity in studies which examine potential environmental causes of cancer using county cancer mortality rates. He attributes this inability to detect some exposures to the fact that the distinction between exposed and unexposed persons is lost to a great extent to the migration of persons between exposed and unexposed areas. In diseases such as cancer, Polissar further states that the latent period from exposure until the terminal event is longer, thus the migration effect might be increased.

Polissar (1980) found that the effect of migration on risk estimates is a function of cancer site, the type of geographic units for which rates are calculated, and latent period. The older the population, the less the probability for migration. Therefore cancers that affect older people have higher risk retention rates than cancers that affect the young, because young people are more likely to migrate in and out of an area of exposure. Secondly, the type of geographic area used can affect the amount of risk retained. People are less likely to cross state boundaries than county boundaries and accordingly less likely to leave a county than a municipal area. The final impact of migration noted is that the longer the latency period, the less retention of excess risk because more opportunity for mixing between exposed and unexposed populations exists. Again, considering the special characteristics of cancer in children discussed in Chapter Two, the relevance of migration and latency issues is questionable.

In comparison to mortality data, the three problems previously identified are reduced with incidence data. Migration is less of a problem because a case history can be taken and diagnosis is more certain because tumours are usually more confined when initially diagnosed (Greenberg, 1983). In addition, the impact of treatment on survivorship is eliminated. However, in the absence of reliable incidence data, mortality data can be used with attempts to minimize the problems of migration, accuracy of diagnosis, local variations in rates of cure, and latency period (Greenberg, 1983). Aggregation of cancer sites, time periods and places may minimize errors caused by these problems. Polissar (1980) outlines some steps investigators could take to guard against the effects of migration in a geographic study. He suggests that study areas be as large as possible to minimize migration rates, that diseases with the shortest latent periods be chosen for study and calls for the development of appropriate migration indices for use in evaluating the effect of migration and additional research into the length of the latent period as a function of

widespread but very low levels of exposure, such as those usually found in geographic studies. As will be discussed further in Chapter Four, the Ontario study of childhood cancer mortality undertaken in this thesis incorporates many of these precautions.

3.3 Overall Conclusions

This chapter has provided an overview of two methodological issues of particular relevance to the analysis of childhood cancer mortality to follow. Two decisions regarding methodology had to be made prior to proceeding with this investigation of childhood cancer mortality in Ontario. The first was to use ecologic analysis, because of the benefits in cost-efficiency and the use of existing and available data sources. Given the suitability of the ecologic method for hypothesis-generating, the possibility that the determinants of childhood cancer are of an environmental nature and thus vary geographically, and the fact that this is the first investigation of spatial patterns of childhood cancer mortality undertaken with this data set, the choice is justified. The second decision was whether to use cancer incidence or mortality data. Mortality data was chosen for the Ontario study primarily because it was the only data readily available, however given the preceding discussion, it appears it was the best suited for this particular undertaking.

CHAPTER FOUR
RESEARCH DESIGN

4.1 Introduction and Research Objectives

The purpose of this chapter is to introduce and discuss issues pertaining to the research design of the Ontario study of childhood cancer mortality, 1976 - 1985. This chapter will focus on the sources of data utilized in the analysis, the dependent and independent variables selected, and the analysis methods employed in this investigation.

This thesis has four research objectives. The first is to describe county level rates of childhood cancer mortality in Ontario for the period 1976 to 1985. This analysis is related to an on-going study of geographical variations in cancer incidence and mortality in Ontario being investigated by a multi-disciplinary team under a Ministry of Health grant at McMaster University and the University of Toronto. This thesis represents an initial ecologic assessment of a seldom studied sub-population in cancer research in Ontario - children aged 0-19. A literature review of existing childhood cancer research was undertaken as an important part of this first objective.

The second research objective is to analyse spatial variations in pediatric cancer mortality in Ontario. An assessment of the geographic distribution of childhood cancer mortality rates at the county level, including an analysis of spatial autocorrelation, provides an important first step in developing hypotheses regarding underlying reasons for variations in these rates over space.

The third objective is to determine to what extent spatial variations in childhood cancer mortality are related to variations in other variables. Based on an assessment of previous literature in the field, detailed in Chapter Two, a set of plausible ecologic correlates of childhood cancer are developed. Using these constructs, independent variables available in the **Census of Canada** are selected and using multiple regression, the ecologic relationship between pediatric cancer mortality and the set of independent variables can be examined. Residuals from the regression analysis are also examined to gain further insight into the ecologic relationships.

The final research objective is to comment on future directions for research in childhood cancer epidemiology and to assess the utility of the ecologic methodology in such investigations.

4.2 Methodology

4.2.1. Cancer Registries

Cancer registries provide the primary source of data in most studies of this disease outcome. A population-based cancer registry is a collection of every single cancer case diagnosed in a given jurisdiction (Spengler, 1981). In this registration, selected information about all cancer patients in a precisely defined population, such as date of birth, site of cancer and residence, is collected in a uniform fashion as soon as possible after diagnosis. The ultimate objective of a population-based registry is to estimate and monitor the incidence of cancer, both overall and by site, in the entire population covered and in subgroups of that population, such as by age, sex, race, and geographic area (Marrett et al., 1986). The data typically undergo rigorous quality control in terms of coding, checking to ensure that a case is not reported more than once, and that every possible case has been identified (Spengler, 1981). Once assembled, the data can be used in numerous

types of epidemiological investigations including monitoring cancer levels, assessment of therapies, development of etiologic hypotheses, risk assessment in population subgroups and monitoring the effects of preventative measures (Marrett et al., 1986).

Although cancer registries provide the most reliable health data available for population-based research, the accuracy and completeness may vary between registries and according to the purposes for which the data from the registry is used. Swerdlow (1986) discusses several considerations relevant to interpretation of cancer registry data. Based on an assessment of registry systems in England and Wales, Swerdlow indicates that there may be missing data on variables such as birthplace or occupation; changes in administrative boundaries over time may occur in the geographic area covered by the registry; individuals may appear in more than one registry; cancer registration may be determined either by place of residence or the location of the treatment facility; administrative delays can cause late registration and deletions or alterations can be made to the file of registrations after data have been extracted; and coding errors can occur (Swerdlow, 1986 and Walter, 1988). These factors must be considered when using cancer registry data, however, many registries employ techniques designed to minimize potential biases and discrepancies.

The source of childhood cancer mortality data in this thesis is the Ontario Cancer Registry (OCR). Mortality frequencies by age, sex, site and county for the ten year period 1976 to 1985 were provided by the Ontario Cancer Treatment and Research Foundation (OCF) who maintain the OCR. The OCR is a population-based registry which obtains data from the following sources: hospital discharges with a mention of cancer; pathology reports with a mention of cancer; reports on patients referred to one of eight Regional Cancer Centres operated by the OCF; and death certificates with cancer as the underlying cause of death. Registration is based on a unique system using sophisticated computer record

linkage which employs medical logic, and result in the generation of incidence data (Reynolds, 1989). Death certificates supplied by the Registrar General of Ontario are the source of the mortality data.

Many of the methodological issues raised previously in this chapter with regard to cancer registries, and in Chapter Three concerning the use of death certificate data, are addressed by the OCR. According to Reynolds (1989), the completeness and accuracy of mortality data from the OCR is good, as both the Registrar General and the OCF conduct routine checks of data quality. The series of steps taken to produce the mortality data set used in this thesis can be summarized as follows: registration of death in Ontario is mandatory, with medical information completed by the attending physician at the time of death and the non-medical section usually completed based on information from the next of kin. The death certificate is then forwarded to the Registrar General of Ontario where it is re-coded and missing or ambiguous information is actively pursued and clarified. The underlying cause of death is then coded using the rules of the International Classification of Diseases. Residence codes, which correspond to census divisions, are assigned with actual place of residence rather than mailing address used, including death certificates for residents of Ontario who die out of the province. The OCF receives computerized tapes of the mortality data from the Registrar General and, combined with the other sources of data previously mentioned, the OCR maintains the data. With the exception of Ontario residents who die outside of Canada, it is believed that cancer death registration in Ontario is virtually complete and due to the record linkage system employed by the OCR, only approximately two per cent of cases in the OCR have a death certificate as their sole source of cancer information (Reynolds, 1989).

With regard to accuracy of the data, the OCR conducted an assessment of the accuracy of mortality data in Ontario. The result of this assessment was an 80 per cent

agreement rate between OCR incident diagnosis and cause of death as coded on the death certificate for all cancers combined (Reynolds, 1989). The major source of discrepancy was cross-classification of cancer sites; as will be discussed later in this chapter, this problem is not relevant in this study of childhood cancer mortality because all sites are combined. In addition, for the period 1976 to 1985 for all ages, the proportion of cases with missing information is very low: 0.095 per cent for census division; 0.003 per cent for age and 0 per cent for sex (Reynolds, 1989).

In addition to providing the necessary numerator data, the OCF also made available the denominator data - population frequencies. A valid ecologic study requires the identification of population denominators, or individuals at risk of experiencing the health event in question, which correspond to the health outcome numerators (Walter, 1988). As a census of the Canadian population is taken every first and sixth year of the decade and the study period of concern in this analysis is 1976 to 1985, three census points were used to produce population data: 1976, 1981 and 1986 (Reynolds, 1989). The population-at-risk data supplied by the OCF is produced using linear interpolation between census years, which is considered to be the best available method of population estimation (Reynolds, 1989). In an analysis of childhood cancer mortality rates, extremely small numerators serve to make any minor inaccuracies in comparatively large denominators insignificant in the calculation of rates.

4.2.2 Spatial Units of Analysis

The analysis of childhood cancer mortality rates in Ontario is conducted at the county level. Appendix One contains a base map of Ontario showing the name and location of the 47 counties used. The decision to conduct this investigation at the county level was made based on two primary considerations: the availability of data for the dependent and

independent variables and the use of the ecologic study design. As previously indicated, mortality data from the OCR is produced at the census division level, which can easily be converted to the county level, and as will be discussed further in this chapter, the independent variables required for the analysis were obtained from the **Census of Canada** in corresponding spatial units. The discussion of ecologic analysis in Chapter Three clearly indicates a preference for county-level populations to minimize potential sources of bias since counties are small enough to ensure some degree of similarity among populations but also allow for cancer rates to be fairly stable. Hoover et al., (1975) state that counties provide a compromise to state-by-state (province-to-province) analyses of geographic variation in cancer mortality because they are units small enough to be homogeneous for demographic and environmental characteristics that might influence cancer risk, and yet large enough for stable estimates of site-specific cancer mortality.

By using county-level data in this analysis, the problem of changing boundaries over time can be eliminated. Although there were several minor changes in the administrative boundaries of census divisions and sub-divisions between 1976 and 1985, the aggregations of several census divisions into counties allows for continuity in spatial units over the study period. In addition, there are fewer problems with comparability of residence classifications over time at the county level than if municipality data were used. People are more likely to incorrectly provide their address by municipality than by county, a potential problem in areas such as Metropolitan Toronto (Walter et al., 1988).

4.2.3 Census of Canada and Ontario Statistics

The sources of data for the independent variables used in this thesis were statistical reports based on the 1981 **Census of Canada** published by Statistics Canada and the Ontario Ministry of Treasury and Economics. Data for 1981 was selected because this year

is the closest mid-point between 1976 and 1985 available. While census data are subject to coverage, response, processing and sampling error given that they are based on a 20 per cent sample of the total population, the overall impact is thought to be minimal (Reynolds, 1989). This thesis will continue with the practise of considering census data as a good estimation of population characteristics and using census data for large population-based studies of this type. A more detailed discussion of the precise census data employed in this analysis will follow in the Independent Variables section of this chapter.

4.3 Dependent and Independent Variables

This section will detail some of the technical issues involved in the use of dependent and independent variables in the analysis of childhood cancer mortality in Ontario to follow. A discussion of standardization as well as age, sex and site considerations of the mortality data will be provided. As well, the process used to select appropriate correlates for the analysis as well as a discussion of the spatial distribution of these independent variables will also be presented.

4.3.1 Dependent Variables

The outcome measure of interest in this analysis of childhood cancer in Ontario is death by cancer for residents of Ontario aged 0 to 19 from 1976 to 1985. As previously indicated, yearly mortality frequencies by age, sex and county were obtained for the ten year period from the OCR, and because of small numbers, all ages, sites and years were aggregated. To facilitate analysis and comparisons between counties, crude mortality rates were initially calculated to allow various sized units to be compared. Tables 4.1, 4.2 and 4.3 present the frequencies and age-specific crude mortality rates for all counties in Ontario in tabular form. Age-specific crude mortality rates were calculated by dividing the total

Table 4.1 **Summary Data For Males 0-19, 1976-1985, All Sites**

	County	Population (000's)	Cancer Deaths	Crude Death Rate (1,000)	Standardized Mortality Ratios
1	Algoma	239.5	9	0.038	0.68
2	Brant	173.5	11	0.063	1.15
3	Bruce	104.1	6	0.058	1.05
4	Cochrane	184.6	17	0.092	1.68
5	Dufferin	57.5	3	0.052	0.95
6	Durham	499.4	26	0.052	0.95
7	Elgin	118.5	4	0.034	0.61
8	Essex	528.8	32	0.061	1.10
9	Frontenac	171.6	8	0.047	0.85
10	Grey	119.9	8	0.067	1.22
11	Haldimand/Norfolk	153.1	10	0.065	1.19
12	Haliburton	16.9	1	0.059	1.08
13	Halton	437.3	21	0.048	0.87
14	Hamilton/Wentworth	639.2	40	0.063	1.14
15	Hastings	180.3	10	0.055	1.01
16	Huron	97.0	3	0.031	0.56
17	Kenora	116.0	7	0.060	1.09
18	Kent	185.5	9	0.049	0.88
19	Lambton	211.7	15	0.071	1.29
20	Lanark	72.6	2	0.028	0.50
21	Leeds/Grenville	130.8	11	0.084	1.53
22	Lennox/Addington	58.4	3	0.051	0.93
23	Manitoulin	19.9	2	0.100	1.83
24	Middlesex	495.8	28	0.056	1.03
25	Muskoka	57.8	3	0.052	0.95
26	Niagara	599.2	40	0.067	1.22
27	Nipissing	148.0	8	0.054	0.98
28	Northumberland	106.8	4	0.037	0.68
29	Ottawa/Carleton	840.2	35	0.042	0.76
30	Oxford	145.8	4	0.027	0.50
31	Parry Sound	54.4	7	0.129	2.34
32	Peel	865.1	46	0.053	0.97
33	Perth	114.0	4	0.035	0.64
34	Peterborough	162.5	10	0.062	1.12
35	Prescott/Russell	93.2	3	0.032	0.59
36	Prince Edward	34.8	1	0.029	0.52
37	Rainy River	43.2	4	0.093	1.69
38	Renfrew	153.0	7	0.046	0.83
39	Simcoe	383.8	21	0.055	0.99
40	Stormont/Dundas/Glengarry	169.7	13	0.077	1.39
41	Sudbury	354.7	20	0.056	1.03
42	Thunder Bay	258.1	14	0.054	0.99
43	Timiskaming	74.3	3	0.040	0.74
44	Victoria	74.6	6	0.080	1.46
45	Waterloo	518.9	24	0.046	0.84
46	Wellington	220.6	8	0.036	0.66
47	York/Metropolitan Toronto	3463.5	196	0.057	1.03
	Total Ontario	13948.0	767	0.055	
	Mean			0.563	1.03
	Standard Deviation			0.203	0.37

Table 4.2 Summary Data For Females 0-19, 1976-1985, All Sites

	County	Population (000's)	Cancer Deaths	Crude Death Rate (1,000)	Standardized Mortality Ratios
1	Algoma	231.0	8	0.035	0.82
2	Brant	164.3	1	0.006	0.14
3	Bruce	98.0	2	0.020	0.48
4	Cochrane	174.8	7	0.040	0.95
5	Dufferin	52.8	1	0.019	0.45
6	Durham	475.2	17	0.036	0.85
7	Elgin	114.3	7	0.061	1.45
8	Essex	506.8	26	0.051	1.22
9	Frontenac	159.0	5	0.031	0.75
10	Grey	113.6	7	0.062	1.46
11	Haldimand/Norfolk	143.6	4	0.028	0.66
12	Haliburton	15.1	0	— —	
13	Halton	414.9	17	0.041	0.97
14	Hamilton/Wentworth	611.1	17	0.028	0.66
15	Hastings	169.9	8	0.047	1.12
16	Huron	92.4	5	0.054	1.29
17	Kenora	111.1	4	0.036	0.85
18	Kent	174.8	8	0.046	1.08
19	Lambton	202.1	9	0.045	1.06
20	Lanark	69.1	3	0.043	1.03
21	Leeds/Grenville	123.1	5	0.041	0.96
22	Lennox/Addington	54.7	3	0.055	1.29
23	Manitoulin	19.8	1	0.050	1.19
24	Middlesex	478.1	29	0.061	1.44
25	Muskoka	55.9	4	0.072	1.69
26	Niagara	571.1	28	0.049	0.33
27	Nipissing	139.3	4	0.029	0.68
28	Northumberland	103.6	6	0.058	1.37
29	Ottawa/Carleton	798.7	37	0.046	1.10
30	Oxford	138.5	9	0.065	1.54
31	Parry Sound	50.0	1	0.020	0.47
32	Peel	825.4	39	0.047	1.12
33	Perth	107.5	2	0.019	0.44
34	Peterborough	155.8	4	0.026	0.61
35	Prescott/Russell	90.0	4	0.044	1.05
36	Prince Edward	34.1	4	0.117	2.78
37	Rainy River	40.7	3	0.074	1.74
38	Renfrew	143.5	7	0.049	1.16
39	Simcoe	356.9	10	0.028	0.66
40	Stormont/Dundas/Glengarry	161.8	10	0.062	1.46
41	Sudbury	338.9	9	0.027	0.63
42	Thunder Bay	247.9	8	0.032	0.76
43	Timiskaming	71.5	3	0.042	0.99
44	Victoria	71.0	3	0.042	1.00
45	Waterloo	497.7	27	0.054	1.29
46	Wellington	208.9	6	0.029	0.68
47	York/Metropolitan Toronto	3310.4	140	0.021	1.00
	Total Ontario	13288.6	562	0.042	
	Mean			0.042	0.997
	Standard Deviation			0.020	0.47

Table 4.3 **Summary Data For Total 0-19, 1976-1985, All Sites**

	County	Population (000's)	Cancer Deaths	Crude Death Rate (1,000)	Standardized Mortality Ratios
1	Algoma	470.5	17	0.036	0.74
2	Brant	337.8	12	0.036	0.73
3	Bruce	202.1	8	0.040	0.81
4	Cochrane	359.4	24	0.067	1.37
5	Dufferin	110.2	4	0.036	0.74
6	Durham	974.6	43	0.044	0.91
7	Elgin	232.8	11	0.047	0.97
8	Essex	1035.6	58	0.056	1.15
9	Frontenac	330.6	13	0.039	0.81
10	Grey	233.4	15	0.064	1.32
11	Haldimand/Norfolk	296.7	14	0.047	0.97
12	Haliburton	32.0	1	0.031	0.64
13	Halton	852.2	38	0.045	0.92
14	Hamilton/Wentworth	1250.4	57	0.046	0.94
15	Hastings	350.2	18	0.051	1.06
16	Huron	189.4	8	0.042	0.87
17	Kenora	227.0	11	0.048	0.99
18	Kent	360.3	17	0.047	0.97
19	Lambton	413.7	24	0.058	1.19
20	Lanark	141.7	5	0.035	0.72
21	Leeds/Grenville	253.8	16	0.063	1.29
22	Lennox/Addington	113.1	6	0.053	1.09
23	Manitoulin	39.8	3	0.075	1.55
24	Middlesex	973.9	57	0.059	1.20
25	Muskoka	113.7	7	0.062	1.27
26	Niagara	1170.2	68	0.058	1.19
27	Nipissing	287.3	12	0.042	0.86
28	Northumberland	210.4	10	0.048	0.98
29	Ottawa/Carleton	1639.0	72	0.039	0.90
30	Oxford	284.2	13	0.046	0.94
31	Parry Sound	104.4	8	0.077	1.58
32	Peel	1690.5	85	0.050	1.03
33	Perth	221.5	6	0.027	0.56
34	Peterborough	318.3	14	0.044	0.90
35	Prescott/Russell	183.2	7	0.038	0.78
36	Prince Edward	68.8	5	0.073	1.49
37	Rainy River	83.9	7	0.083	1.71
38	Renfrew	296.5	14	0.048	0.97
39	Simcoe	740.7	31	0.042	0.86
40	Stormont/Dundas/Glengarry	331.5	23	0.069	1.42
41	Sudbury	693.6	29	0.042	0.86
42	Thunder Bay	506.0	22	0.043	0.89
43	Timiskaming	145.8	6	0.041	0.85
44	Victoria	145.7	9	0.062	1.27
45	Waterloo	1016.6	51	0.050	1.03
46	Wellington	429.5	14	0.033	0.67
47	York/Metropolitan Toronto	6774.0	336	0.050	1.02
	Total Ontario	27236.6	1329	0.049	
	Mean			0.50	1.02
	Standard Deviation			0.0127	0.26

number of cancer deaths for residents aged 0-19 in each county over the ten year period by the total population aged 0-19 in the county for the ten year period and multiplying by 1,000. This crude death rate, although allowing for comparisons between different sized populations, does not take into account differences in the population characteristics, i.e. age, of each county. Further, these crude mortality rates are extremely variable because of the very low frequency of childhood cancer mortality over the time period.

To overcome this methodological problem of small numbers and to control for differences in the population characteristics i.e. age of each county, standardized mortality ratios are used. Two methods of adjusting or standardizing rates are identified in standard epidemiology texts; the direct and indirect methods. The indirect method is used to compare populations in which rates are excessively variable because of small numbers (Mausner and Bahn, 1974). In this method, the more stable rates of the larger population are applied to the smaller study population, thus in this case the mortality rates for each county are compared with the rates for the entire province. The Standardized Mortality Ratio (SMR) is the total observed deaths in a county divided by the total expected deaths in that county. The expected deaths are calculated by multiplying the Ontario rate by the population of each county. If the mortality ratio is greater than 1.0, more deaths are observed in the county than would be expected based on rates for the entire province. If the ratio is less than 1.0, fewer deaths are observed than expected (Mausner and Bahn, 1974). The SMRs for all counties are presented in Tables 4.1, 4.2 and 4.3.

Based upon the preceding discussion, it was felt that SMRs would allow for greater confidence in interpreting the results of any statistical or spatial analysis of childhood cancer mortality in Ontario, thus they were used as the dependent variable or primary outcome measure in this thesis.

Further issues regarding the dependent variables had to be addressed prior to the analyses. Since hypotheses relating to childhood cancer etiology may vary depending upon age (i.e. 0-4 versus 15-19) and cancer site (i.e. leukaemia versus Hodgkin's Disease) it would be preferable to analyse the data by various age groupings and by different sites. However, due to the small numbers of childhood cancer deaths in Ontario over the study period, all sites and all ages were aggregated for this analysis. This was necessary to yield sufficient cases for meaningful analysis. The only breakdown of the data which was possible in this study was by sex; males, females and total populations are presented and analysed separately.

4.3.2 Independent Variables

The selection of potential ecologic correlates of childhood cancer mortality for analysis was limited by the availability of environmental and socio-demographic data. Chapter Two presents a review of the literature and identifies three broad classifications of covariants of childhood cancer - demographic and genetic factors, environmental (physical and socio-cultural) factors and factors related specifically to mortality. Demographic factors include race, ethnicity, age and sex while genetic factors concern hereditary factors which may predispose the development of malignancies in children. Environmental covariants linked to childhood cancer in the literature include radiation, chemicals, viral infections, seasonality, parental occupation, socio-economic status, urbanization and other lifestyle factors such as diet. Mortality risk factors include those factors which may increase or decrease mortality rates irrespective of incidence such as access to medical care.

From this broad set of potential correlates of pediatric cancer identified in the literature, variables suitable for this particular investigation were selected. Suitability was determined by considerations of both etiologic plausibility and availability in existing data

sources. In addition, the advantages and limitations of the ecologic methodology discussed in Chapter Three - the ability to use existing data sources and to study groups not individuals - must be kept in mind. This consideration eliminates inclusion of genetics and heredity as well as the direct incorporation of physical environmental variables as they are not available. With those essentially eliminated and considering the previous age and sex discussion, four broad constructs emerge from the literature as potential correlates of variations in childhood cancer mortality to investigate: urbanization, socio-economic status, parental occupation and ethnicity. Table 4.4 presents a list of 17 independent variables organized under these four headings which were used in the analysis, indicating the source of each data set.

Urbanization: Two measures of urbanization were selected directly from the census data: percent of the population urban (over 1,000 people and 400 or more per km^2) and population density. One additional measure of urbanization was used because these first two measures are considered too broad to adequately differentiate between urban and rural areas. The percentage of the population living in places with over 5,000 inhabitants was calculated from the census data as a measure which may more clearly differentiate between rural and urban counties.

Socio-economic Status: Three measures of socio-economic status available in the census were used: median household income, residents with less than a Grade Nine education and incidence of low income. Previous studies have relied on median income as a measure of income. Reynolds (1989) suggests that incidence of low income may be a more suitable indicator, thus it is used here.

Parental Occupation: The measures available from calculations of the census data include manufacturing establishments per 1,000 population and per 1,000 urban population, and employment by primary, secondary and tertiary sectors.

Table 4.4 **Independent Variables**

Independent Variables	Definition	Source
Urbanization		
1. Urban (5,000)	Total population of the county minus all of the cities, towns and villages with populations of 5,000 or more. Townships,territories etc. not included as urban even if population is over 5,000.	Statistics Canada, **1981 Census of Canada, Population, occupied private dwellings, private households and census and economic families in private households. Selected social and economic characteristics.** Catalogue 95 - 942 Volume 3 - Profile Series B, Table 1.
2. Urban (Census)	Percent of total population living in an area having a population concentration of 1,000 or more **and** a population density of 400 or more per km ^{2}.	Ontario Ministry of Treasury and Economics. **Ontario Statistics, 1982.** Table 2.16.
3. Population Density	Population density by km ^{2}.	Statistics Canada, **1981 Census of Canada, Population, occupied private dwellings, private households and census and economic families in private households. Selected social and economic characteristics.** Catalogue 95 - 942 Volume 3 - Profile Series B, Table 1.
Socioeconomic Status		
4. Education < Grade 9	Percent of the population 15 years and over with less than grade 9 education.	Statistics Canada, **1981 Census Canada, Population, occupied private dwellings, private households and census and economic families in private households. Selected social and economic characteristics.** Catalogue 95 - 942 Volume 3 - Profile Series B, Table 1.
5. Incidence of Low Income	Incidence of Low Income in Percentage On the basis of the total income of a family unit adjusted for federal Child Tax Credit, size of the family unit and size of the area of residence, the position of eacheconomic family is determined in relationto low income cut-offs based on the 1987 Family Expenditure Survey and updated by changes in the Consumer Price Index. The incidence of low income is the percentage of units below the low income cut-offs.	Statistics Canada, **1981 Census of Canada, Population, occupied private dwellings, households and census and private economic families in private households. Selected social and economic characteristics.** Catalogue 95 - 942 Volume 3 - Profile Series B, Table 1.

Independent Variables	Definition	Source
6. Median Income	Median Household Income in dollars. Median income is the amount which divides the income size distribution of the group in two parts, one having incomes below the median and the other having incomes above the median.	Statistics Canada, **1981 Census of Canada, Population, occupied private dwellings, private households and census and economic families in private households. Selected social and economic characteristics.** Catalogue 95 - 942 Volume 3 - Profile Series B, Table 1.
Parental Occupation		
7. Manufacturing (1,000)	Number of Manufacturing Establishments per 1,000 population. Establishment is the smallest operating unit capable of reporting specified input and output data; usually a plant or a mill.	Statistics Canada, **Manufacturing Industries of Canada: sub-provincial areas, 1981, Annual Census of Manufacturers.** Catalogue 31 - 209, Table 6.
8. Manufacturing (1,000 Urban)	Number of Manufacturing Establishments per 1,000 urban population. See above.	Statistics Canada, **Manufacturing Industries of Canada: sub-provincial areas, 1981, Annual Census of Manufacturers.** Catalogue 31 - 209, Table 6.
9. Primary Employment	Percent of the Labour Force 15 Years and older employed in Primary Industries. Includes: farming, horticultural and animal husbandry occupations; fishing, trapping and related occupations; forestry and logging occupations; mining and uarrying including oil qand gas field occupations.	Statistics Canada, **1981 Census of Canada, Population, occupied private dwellings, private households and census and economic families in private households. Selected social and economic characteristics.** Catalogue 95 - 942 Volume 3 - Profile Series B, Table 1.
10. Secondary Employment	Percent of the Labour Force 15 Years and older employed in Secondary Industries. Includes: processing occupations; machining and related occupations; product fabricating, assembling and repairing occupations; construction trades occupations; transport equipment operating occupations; material handling and related ; occupations other crafts and equipment operating occupations.	Statistics Canada, **1981 Census of Canada, Population, occupied private dwellings, private households and census and economic families in private households. Selected social and economic characteristics.** Catalogue 95 - 942 Volume 3 - Profile Series B, Table 1.

Independent Variables	Definition	Source
11. Tertiary Employment	Percent of the Labour Force 15 Years and older employed in Tertiary Industries. Includes: managerial, administrative and related occupations; occupations in natural sciences, engineering, and mathematics; occupations in social sciences and related fields; occupations in religion; teachingand related occupations; occupations in medicine and health; artistic, literary, recreational and related occupations; clerical and related occupations; sales occupations; service occupations.	Statistics Canada, **1981 Census of Canada, Population, occupied private dwellings, private households and census and economic families in private households. Selected social and economic characteristics.** Catalogue 95 - 942 Volume 3 - Profile Series B, Table 1.
Ethnicity		
12. British	Proportion of the population of British origin, 1981 Census of Canada.	Reynolds, Donna, 1989. **The Geographic of Cancer Mortality in Ontario: An Ecologic Analysis on Ethnicity, Industry and Socio Economic Characteristics, 1976-1985.** M.Sc. Thesis, University of Toronto.
13. French	Proportion of the population of French origin, 1981 Census of Canada.	Reynolds, Donna, 1989. **The Geographic of Cancer Mortality in Ontario: An Ecologic Analysis on Ethnicity, Industry and Socio Economic Characteristics, 1976-1985.** M.Sc. Thesis, University of Toronto.
14. Dutch	Proportion of the population of Dutch origin, 1981 Census of Canada.	Reynolds, Donna, 1989. **The Geographic of Cancer Mortality in Ontario: An Ecologic Analysis on Ethnicity, Industry and Socio Economic Characteristics, 1976-1985.** M.Sc. Thesis, University of Toronto.
15. German	Proportion of the population of German origin, 1981 Census of Canada.	Reynolds, Donna, 1989. **The Geographic of Cancer Mortality in Ontario: An Ecologic Analysis on Ethnicity, Industry and Socio Economic Characteristics, 1976-1985.** M.Sc. Thesis, University of Toronto.

Independent Variables	Definition	Source
16. Italian	Proportion of the population of Italian origin, 1981 Census of Canada.	Reynolds, Donna, 1989. **The Geographic of Cancer Mortality in Ontario: An Ecologic Analysis on Ethnicity, Industry and Socio Economic Characteristics, 1976-1985.** M.Sc. Thesis, University of Toronto.
17. Native	Proportion of the population of Native origin, 1981 Census of Canada.	Reynolds, Donna, 1989. **The Geographic of Cancer Mortality in Ontario: An Ecologic Analysis on Ethnicity, Industry and Socio Economic Characteristics, 1976-1985.** M.Sc. Thesis, University of Toronto.

Ethnicity: The five largest ethnic groups in Ontario included in the 1981 Census were British, French, Dutch, German and Italian. In addition, Native population was selected because of the prominence of this group in several Ontario counties.

As the research objective relevant to the investigation of these independent variables is to determine to what extent spatial variations in childhood cancer mortality are related to variations in other variables, this set provides a useful first step in developing etiological hypotheses warranting further investigation. Appendix Two contains the percentages or other measures obtained from the census for each of these variables.

Representation of the independent (and dependent) variables for spatial analysis was conducted using the mapping package AtlasGraphics, a Geographic Information System program available at McMaster University. The maps produced by this program display each variable by county in quintiles, with the counties with the highest values shaded darkest and shading of decreasing intensity corresponding to lower values. Appendix Three contains a set of 17 maps, illustrating the spatial distribution of each variable by county. Although a detailed discussion of the spatial pattern of each independent variable is not relevant here, a brief comment on each construct is warranted.

Urbanization: A pattern of high urbanization in the counties comprising the Golden Horseshoe of Southern Ontario (from Durham through Metropolitan Toronto and the surrounding area to Hamilton and Niagara) as well as the Ottawa-Carleton area is seen.

Socio-economic Status: No obvious pattern of socio-economic status as measured by income and education level is apparent, with median household income highest in the Northwest (Sudbury, Thunder Bay), and Metropolitan Toronto areas. Residents of Northern Ontario are generally less likely to have attained over a Grade Nine education. The pattern of incidence of low income shows Northeastern Ontario (Nipissing, Temiskaming, Parry Sound) to have the highest percentage of residents living below the

low income cutoff while Northwestern Ontario (Kenora, Thunder Bay) and many Southern Ontario counties have the smallest proportions.

Parental Occupation: The highest number of manufacturing establishments per 1,000 population can be found in South Central Ontario (Durham, Metropolitan Toronto and area to Wellington and Oxford). When considered per 1,000 urban population, the counties with the highest number of manufacturing establishments appear in a scattered pattern throughout the province (Parry Sound, Victoria, Prescott and Russell). Primary industries employ the highest proportions in Southwestern Ontario (Bruce, Perth, Elgin); secondary industries in Central (Waterloo, Perth) and Eastern Ontario (Leeds and Grenville, Lanark, Hastings) and tertiary sector in several areas (Metropolitan Toronto area; Nipissing - Parry Sound; Middlesex and Ottawa-Carleton).

Ethnicity: The highest proportions of British population are found in counties in Central and Eastern Ontario; French in Northeastern Ontario; Dutch in Southwestern Ontario; German in Central and Northeastern Ontario; Italian in Northwestern Ontario and Southern Ontario; and Natives in the North.

4.4 Analysis Methods

This section is designed to provide a brief overview of the statistical methods used in the analysis of childhood cancer mortality in Ontario to follow. Spatial autocorrelation, regression and analysis of residuals were used as analytic tools in this study, thus the main components of each will be discussed.

4.4.1 Spatial Autocorrelation

The second research objective set out in this thesis is to analyse spatial variations in childhood cancer mortality across Ontario. The most obvious method for such an analysis

is to examine maps of the spatial distribution of SMRs and look for patterns. To add to this "eyeball" assessment, spatial autocorrelation can be used as a technique to assess the degree to which SMRs in one county are similar to SMRs in other counties located nearby. Goodchild (1986) describes the role of autocorrelation in spatial analysis as follows: spatial analysis deals with two distinct types of information; attributes of spatial features (SMRs) and locations (counties). Many types of analysis, such as multiple regression and factor analysis, look only at feature attributes without making explicit reference to location. Location is involved indirectly in the sense that it is used to determine whether a case falls inside or outside the study area, but the locations of cases within the study area in no way affects the outcome of the analysis. Spatial autocorrelation is a technique which deals simultaneously with both locational and attribute information. For example, a pair of spatial features such as counties, may be similar or dissimilar in attributes (SMRs) and their proximity will determine how similar they are in spatial location. Spatial autocorrelation compares the two sets of similarities. If counties which are similar in location also tend to be similar in SMRs, the pattern shows positive spatial autocorrelation. Conversely, negative spatial autocorrelation exists when counties which are close together in space tend to be more dissimilar in SMRs than counties which are further apart. The case of zero spatial autocorrelation occurs when SMRs are independent of county location.

Goodchild (1986) states further that the practical importance of spatial autocorrelation is that as an index, it provides a type of information about a spatially distributed phenomenon which is not available in any other form of statistical analysis, and which can be vital to correct interpretation. In interpreting maps of mortality, then, spatial autocorrelation measures give a precise and objective value to something which would otherwise have to be perceived subjectively from a map. In the context of this thesis, areas of the province which show higher than expected ratios of childhood cancer mortality can

be evaluated to determine whether these areas of high cancer risk tend to show spatial aggregation.

Walter et al., (1991) report on a literature review conducted to assess various statistical methods used on epidemiologic data to test for clustering. They conclude that the methods cited most consistently in the literature, and which fully utilize data similar to county level cancer data, are the Moran Index and the Geary Ratio. Walter et al., (1991), through the use of simulated data, further assess the effect of population size on these statistics, in light of the fact that within Ontario the variation in county populations is extremely wide, and conclude that the Moran Index shows the least effect. Further simulations were conducted to investigate the relative and absolute power of these methods, and the results show that the level of power afforded by the Moran Index is quite high, thus giving confidence in the results obtained from this statistic.

The Moran Index is used in this analysis since it seems to be preferred to Geary due to its efficiency and stability. The Moran Index is positive when nearby counties tend to be similar in SMRs, negative when they tend to be more dissimilar than expected, and approximately zero when SMRs are arranged randomly and independently in space (Goodchild, 1989). The results of the spatial autocorrelation analysis are provided in Chapter Five.

4.4.2 Regression

Regression is the method of choice for examining the effects of several variables simultaneously on the risk or rate of disease (Reynolds, 1989). A review of the literature presented in Chapter Three advocates the use of regression techniques over correlation to minimize bias in ecologic analyses, with as many risk factors included as possible. The use of regression allows for more accurate assessments of the magnitude of each

association under examination. Wellington et al. (1979), in examining state patterns of cancer mortality in the U.S., state that correlations between mortality patterns and factor variables can be misleading. Association may be exhibited that are artificial, due to the statistical association of variables. Because of this multicollinearity among the variables, multiple regression techniques that determine the net effect of each factor on a mortality pattern after correction for its relationship with all other variables considered in the model are preferred (Wellington et al., 1979).

Multiple regression is used in this thesis to satisfy the third research objective - to determine to what extent spatial variations in childhood cancer mortality are related to variations in other variables. Multiple regression is specifically used to examine the ecologic relationship between SMRs and the set of 17 socio-demographic and environmental variables. Using SPSS/PC+, a statistical package designed for the IBM PC, the SMRs for males, females and the total population aged 0-19 for each county in Ontario were run against the set of data for the independent variables for each county in a step-wise fashion. This step-wise multiple regression selects the independent variable(s) which best explains the variation in SMRs across the geographical study area. The results of the regression analysis are presented and discussed in Chapter Five.

4.4.3 Residuals

The examination of residuals from the regression analysis is the fourth analysis method used in this study. The mapping of residuals enables the researcher to investigate variance in the dependent variable which was not accounted for in the regression analysis. In other words, it allows for an analysis of the effects of variables that were not selected in the step-wise regression. If a pattern or clustering of residuals is apparent, further investigation of potential predictor variables can be considered (Gardner, 1973). In this

study, if several counties demonstrate significant variance (significant underprediction or overprediction not explained by the independent variable(s) selected in the regression analysis), hypotheses concerning what these counties have in common may suggest other variables which should be considered in the analysis of childhood cancer mortality. The maps of the residuals as well as a discussion of their pattern is included in Chapter Five.

4.5 Conclusion

This chapter has provided details concerning sources of data, choice of dependent and independent variables, and analysis methods used in this study of childhood cancer mortality in Ontario. The results of this analysis follow.

CHAPTER FIVE
ANALYSIS

Introduction

The purpose of this chapter is to report on the findings of the ecologic analysis of childhood cancer mortality in Ontario, 1976 - 1985. The results of four stages of data analysis, beginning with a geographic analysis using maps, followed by spatial autocorrelation, correlation and regression analysis and finally an analysis of residuals will be presented and findings will be interpreted as they relate to the research objectives of this thesis.

5.1 Spatial analysis

As detailed in Chapter Four, the second research objective of this investigation is to analyse spatial variations in childhood cancer mortality in Ontario, as an important first step in developing hypotheses regarding underlying reasons for variations over space. While it must be clearly stated that geographic description of pediatric mortality is not capable of generating evidence sufficient to demonstrate causation or identify potential etiologic agents, the spatial analytic approach presents a valuable starting point. A careful examination of the spatial distribution of childhood cancer mortality in Ontario has the potential to reveal several things about the pattern in question. First, an advantage to mapping mortality or other variables is that a large volume of data can be visually comprehended almost instantaneously (Blot et al., 1979).

Second, at the initial stage in the study of a particular cancer, the spatial approach contributes in a very useful way by providing information as to whether or not the cancer in

question is influenced by environmental factors (Glick, 1982). For example, if the distribution of childhood cancer mortality reveals distinct geographical variation over space, the influence of localized environmental or socio-demographic factors would be indicated. Clustering of high rates in certain areas of the province would, for example, serve as clues to environmental exposures that may be uncovered through further study (Blot et al., 1979). On the other hand, if the distribution of childhood cancer mortality across Ontario appears random, this lack of distinct geographical variation may point to a genetic etiology or suggest the influence of a non-localized stimulus (Glick, 1982). Mathews (1988) concludes that the absence of detectable clustering implies either that the causes of a disease are ubiquitous and are distributed uniformly in time and place or that the clustering is difficult to detect against a background of random variation.

One limitation to the ability of this geographical approach to cancer epidemiology to generate such conclusions regarding the role of environmental influences is scale. Clearly, the degree of spatial variation exhibited by cancer mortality rates depends on the scale at which the investigation is carried out. In the context of this thesis, a finding of significant spatial variation in childhood cancer mortality rates would point to a subset of environmental and socio-demographic influences which vary at the same scale as that on which the cancer rate is aggregated - the county. Conversely, discovering that childhood cancer mortality does not exhibit significant geographical variation at the county level need not suggest a lack of environmental dependence, rather it allows the researcher to rule out potential environmental and socio-demographic factors that operate at that county scale (Cleek, 1979).

The major advantage of examining the spatial distribution of pediatric cancer mortality across Ontario is the potential to generate etiologic hypotheses from the geographic patterns. The geographical variation in cancer has been used to postulate that

approximately 80 percent of human cancer is due to environmental factors (Glick, 1982). This study of childhood cancer mortality rates by county may reveal to what extent this is a relevant statement when applied to this specific sub-group of the population. Thus, an examination of the geographical patterns of childhood cancer mortality can lead to the formation of hypotheses regarding factors which seem important for further testing using other analytical techniques such as case-control studies. The usefulness of the spatial analytic approach as the first level in this investigation of cancer mortality, then, is to raise questions about factors which may be relevant rather than answering them.

Figures 5.1, 5.2 and 5.3 present a visual display of the dependent variables by county for males, females and the total population aged 0 to 19 respectively. The data for the SMRs by county are contained in Tables 4.1, 4.2 and 4.3. The maps were produced using AtlasGraphics, discussed previously, which ranks the SMRs and places them into approximate quintiles. The point must be made that placing the SMRs into quintiles irrespective of the range in magnitude of the values, essentially forces spatial variation to be observed on the maps. It is therefore important to pay attention to the range of values of the SMRs when examining the mapped pattern produced. A further assessment of the range of the SMR values can be facilitated by an evaluation of the mean and standard deviation for each data set. The mean and standard deviation values for each SMR are contained in Tables 4.1, 4.2 and 4.3. Due to the wide variation in the populations of Ontario counties and the rare occurrence of childhood cancer mortality, it is expected that some extreme SMR values will be generated.

Figure 5.1 presents the spatial distribution of cancer mortality for males aged 0 - 19 in Ontario over the ten year study period. The standardized mortality ratios for males range from a minimum of 0.5 to a maximum of 2.34, with a mean value of 1.03 and a standard deviation of 0.37. The range in male SMR values is skewed by one county (Parry Sound)

Figure 5.1

STANDARDIZED MORTALITY RATIOS

Males 0–19

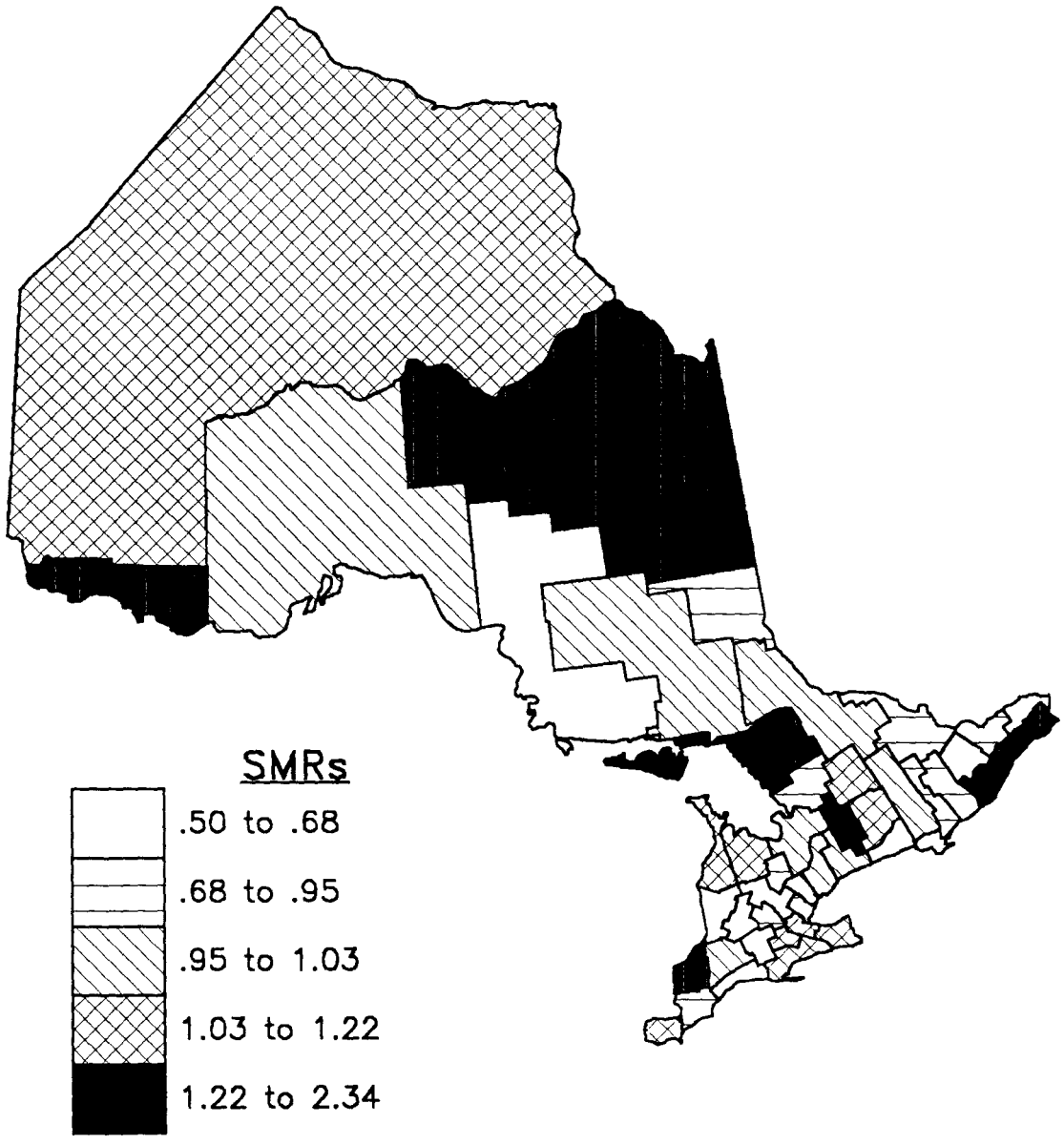


Figure 5.2

STANDARDIZED MORTALITY RATIOS

Females 0–19

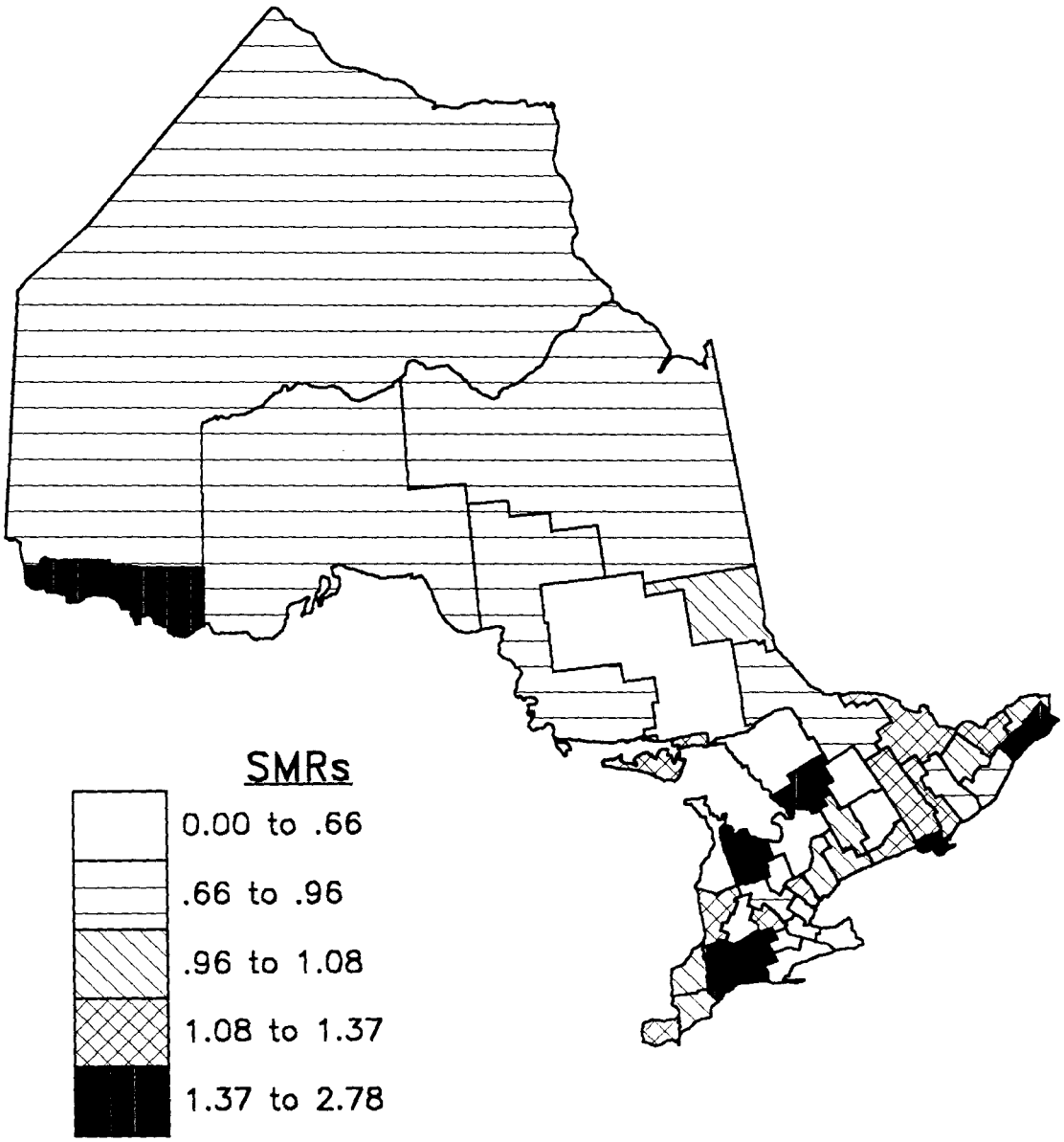
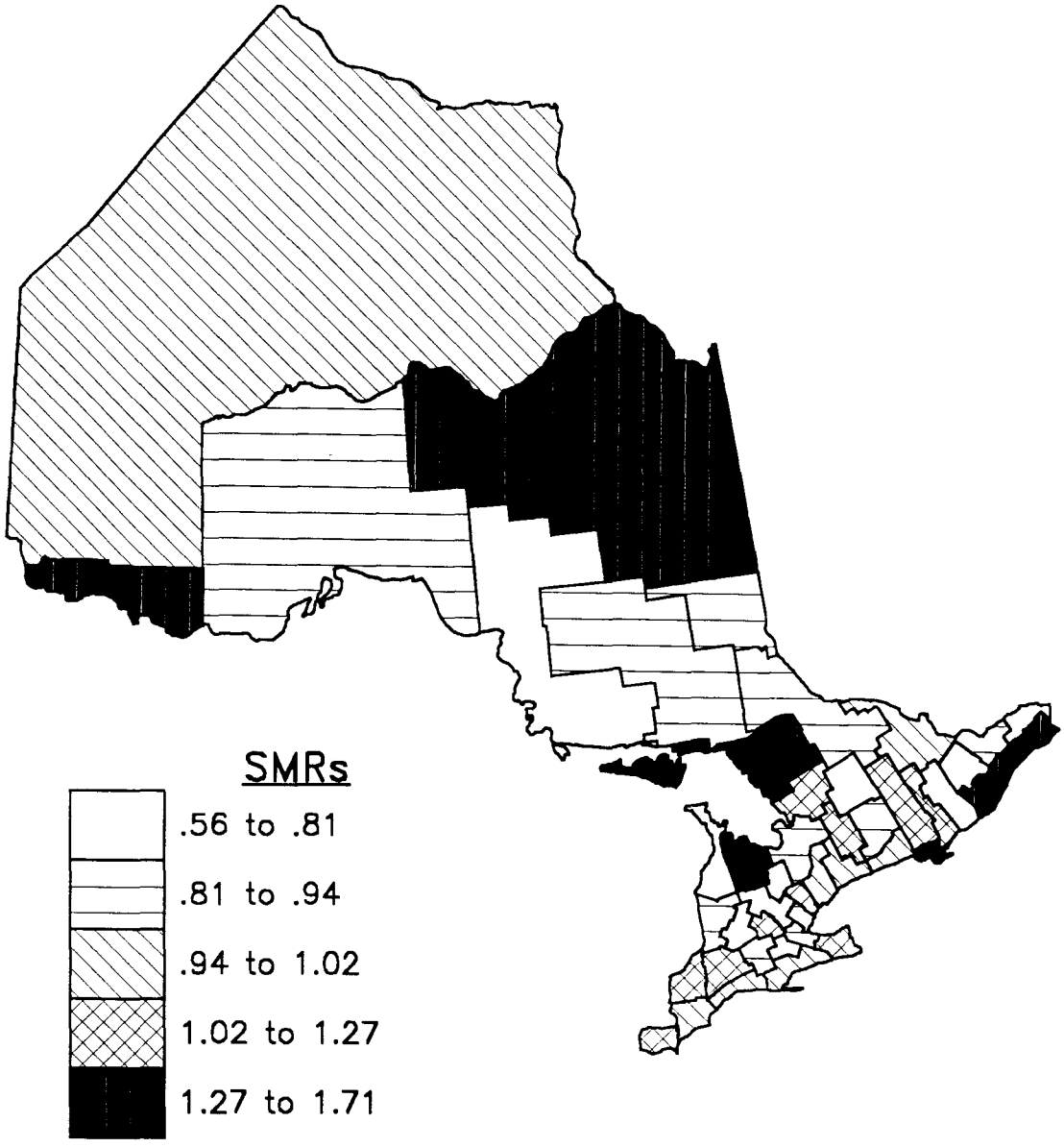


Figure 5.3

STANDARDIZED MORTALITY RATIOS

Total 0–19



which exhibits an extreme value of 2.34. The values for the remaining 46 counties cluster within two standard deviations from the mean, suggesting a limited amount of geographical variation. No clear pattern indicating areas with consistently higher or lower than expected mortality is evident, however higher ratios, represented by dark shading on the map, appear to be found in the more northern and rural counties, outside of the major population centres. It should be noted that the large northern counties have a disproportionate visual impact when viewed on a map of the entire province and may obscure the observed pattern. Counties with lower than expected ratios, indicated in white in the map, appear to be quite randomly distributed, with one cluster of counties apparent in Southwestern Ontario (Huron, Perth, Wellington). Lack of a consistent pattern is further evidenced by the presence of counties in the highest quintile immediately adjacent to counties in the lowest quintile.

Figure 5.2 displays cancer mortality by county for females aged 0 to 19, across Ontario. The standardized mortality ratios for females range from a minimum of 0 to a maximum of 2.78, with a mean value of 0.997 and a standard deviation of 0.47, indicating a slightly larger amount of variation than those for males. However, extreme values for two counties; Haliburton which had no female childhood cancer deaths over the ten year study period and Prince Edward, which due to its small population and slightly elevated number of female cancer deaths has a SMR of 2.78, extend the range in SMR values. The pattern illustrated on the map is quite different than the previous one for males with rates in Northern Ontario consistently at or lower than expected with the exception of Rainy River. Counties with higher than expected ratios demonstrate no obvious pattern, however a cluster of three Southwestern Ontario counties (Middlesex, Oxford, Elgin) is evident. Overall, no distinctive geographical variation of female cancer mortality over space is revealed.

Figure 5.3 presents the spatial distribution of childhood cancer mortality in Ontario for males and females aged 0 to 19 combined. The standardized mortality ratios range from a minimum of 0.56 to a maximum of 1.71, a very narrow range suggesting a limited amount of variation between counties. The mean for this data set is 1.02 with a standard deviation of 0.26. Three counties with low populations (Rainy River, Manitoulin and Parry Sound) fall slightly more than two standard deviations away from the mean. Again, patterns indicating areas of the province with clusters of counties consistently higher or lower than expected are not evident. Counties with SMRs higher than expected are found around Georgian Bay, in Southwestern Ontario and discretely distributed in Rainy River and Cochrane to the north. Counties with lower than expected SMRs appear to be randomly distributed in the Southwestern, Eastern and Central areas of the province. Lack of a consistent pattern of spatial aggregation of rates is further evidenced by the appearance of counties in the highest quintile immediately adjacent to those falling in the lowest quintile.

What information has been revealed by this initial stage of analysis? The most obvious conclusion is that the distribution of childhood cancer mortality in Ontario does not reveal evidence of significant spatial clustering or distinct and systematic geographical patterns. This is consistent with the existing body of literature on spatial variations in childhood cancer (see Chapter Two). The preceding discussion of the spatial analytic approach to cancer epidemiology suggested that if the distribution of mortality appears random across the province, a genetic etiology or factor which is distributed uniformly across the province may be indicated. Further, due to the fact that this investigation is carried out at the county level, the lack of systematic spatial variation or clustering at this scale does not indicate that environmental and socio-demographic factors are not relevant, simply that they cannot be detected visually at the county level. Thus, this geographical

analysis of spatial variations in childhood mortality in Ontario yields no hypotheses relating to underlying environmental or socio-demographic factors which may explain the variations. Further levels of analysis are therefore required to determine to what extent, if any, factors which cannot be detected by a mapped pattern may be relevant.

5.2 Spatial autocorrelation

The second analytic method used to investigate spatial variations in childhood cancer mortality in Ontario is spatial autocorrelation. As detailed in Chapter Four, spatial autocorrelation is a technique which assesses the degree to which mortality ratios in one county are similar to ratios in counties located nearby. In the context of this study, spatial autocorrelation analysis provides a precise and objective measure for patterns which were subjectively perceived from maps in the previous discussion. Spatial autocorrelation analysis is designed to evaluate the presence of clusters of counties with elevated or lower than expected mortality ratios. This could then lead to the identification of potential environmental and socio-demographic variables to be included in the regression analysis to follow.

Before presenting the results of the spatial autocorrelation analysis it is appropriate to consider what this analysis is capable of revealing about childhood cancer in Ontario. If strong spatial autocorrelation is found, counties with high or low childhood cancer mortality are spatially clustered. Hypotheses regarding what factors these contiguous counties have in common which may explain these elevated or depressed rates could then be generated. Conversely, if measures of negative autocorrelation are produced, contiguous counties are more dissimilar in SMRs than counties which are further apart, suggesting no regional variation in underlying factors which could influence rates. If, however, zero spatial autocorrelation results, two conclusions could be drawn about

variations in childhood cancer mortality in Ontario. First, that the dependent variable is distributed randomly and independently in space and therefore cannot easily be explained by the distribution of environmental or socio-demographic variables. Or, secondly, that the factors which might be associated with pediatric cancer mortality are not spatially distributed when aggregated at the county level. The lack of spatial autocorrelation, then, implies that random spatial variations in the dependent variable may be influenced by the distribution of some factor or factors which are also randomly distributed at the county level.

The results of the spatial autocorrelation analysis of the Ontario childhood cancer mortality data set are presented in Table 5.1. A Moran Index value of +1.0 would indicate spatial clustering of the standardized mortality ratios. As evidenced in Table 5.1, none of the Moran Index values approach +1.0 therefore no spatial autocorrelation is present. The significance level is determined by a Z-score of 1.96 at the 5 percent level and 2.58 for the 1 percent level. Again, Table 5.1 shows that none of the Z-scores approach this level, thus none of them are significant.

The findings of this spatial autocorrelation analysis confirm the conclusions drawn from the visual inspection of the maps - no spatial clustering of childhood cancer mortality rates in Ontario is present at the county level. Despite the lack of spatial autocorrelation, the possibility remains that there are socio-demographic and environmental effects at the county level. This could occur if counties which exhibit high childhood cancer mortality rates, though not spatially contiguous, are characterized by similar environmental and socio-demographic chatacteristics.

5.3 Correlation and Regression Analysis

The lack of clear spatial clustering of childhood cancer mortality rates across Ontario, as evidenced by the results of the previous evaluation of mapped patterns and the

Table 5.1 **Spatial Autocorrelation Statistics - Moran Index**

Dependent Variable	I - Value	Z - Score*
Standardized Mortality Ratio Males	-0.04	-0.26
Standardized Mortality Ratio Females	-0.06	-0.42
Standardized Mortality Ratio Total	+0.02	+0.44

* Level of Significance: 5% (±1.96)

(Critical values of Z) 1% (±2.58)

spatial autocorrelation analysis, does not mean that further examination of variations in these rates is meaningless. Rather it indicates that no hypothesized risk factors which might influence variations in childhood cancer are readily apparent thus facilitating the selection of a sub-set of factors warranting more detailed analysis. The third technique employed in this analysis of childhood cancer mortality in Ontario is multiple regression. This section begins with a preliminary assessment of the correlations between the mortality outcomes and the independent variables. Then, as discussed in Chapter Four, step-wise multiple regression is used to examine the ecologic relationship between SMRs and the set of 17 socio-demographic and environmental variables. The results of such an analysis will identify the independent variable or variables which best explain the variations in SMRs across Ontario. These 'predictor' variables which are selected in the regression analysis imply statistical associations between childhood cancer mortality and the particular environmental or socio-demographic variables, they are not in themselves evidence of causation.

Before presenting the results of the regression analysis, it is valuable to examine the correlations between the dependent and independent variables as well as among the independent variables themselves. Correlation analyses provide a quick way to estimate the magnitude and direction of associations between variables (Reynolds, 1989). Table 5.2 displays the correlation matrix for SMRs and the 17 socio-demographic and environmental variables. The full correlation matrix is contained in Appendix Four.

As indicated by the magnitude of the correlation coefficients contained in Table 5.2, no strong associations between SMRs and any one of the 17 independent variables are found. This indicates that as anticipated, no single 'predictor' variable capable of explaining variations in childhood cancer mortality across Ontario emerges. However,

Table 5.2 **Correlation Results**

	Dependent Variables		
Independent Variables	Standardized Mortality Ratio Males	Standardized Mortality Ratio Females	Standardized Mortality Ratio Total
1 Urban (5,000)	-0.120	-0.160	-0.183
2 Urban (Census)	-0.177	-0.123	-0.215
3 Population Density	-0.032	-0.019	-0.031
4 Education < Grade 9	*+0.287	+0.006	+0.237
5 Incidence of Low Income	+0.271	-0.054	+0.193
6 Median Income	-0.251	-0.066	-0.255
7 Manufacturing (1,000)	*-0.308	+0.101	-0.174
8 Manufacturing (1,000 Urban)	+0.010	*+0.457	*+0.342
9 Primary Employment	-0.204	+0.154	-0.055
10 Secondary Employment	*-0.311	+0.029	-0.241
11 Tertiary Employment	-0.054	-0.004	-0.050
12 British	-0.067	+0.084	+0.001
13 French	-0.100	-0.020	-0.034
14 Dutch	-0.279	+0.158	-0.084
15 German	-0.175	-0.051	-0.169
16 Italian	-0.027	-0.257	-0.155
17 Native	*+0.388	+0.012	*+0.311

*$p<0.05$

despite the low magnitude of the correlations, several interesting relationships between SMRs and the socio-demographic and environmental variables emerge.

As shown in Table 5.2, four independent variables correlate significantly ($p< 0.05$) with SMRs for males: education less than Grade 9 (+0.287); number of manufacturing establishments per 1,000 population (-0.308); parental employment in secondary industries (-0.311) and native ethnicity (+0.388). For female SMRs, only the number of manufacturing establishments per 1,000 urban population is significant (+0.457). Total SMRs correlate significantly with two independent variables: number of manufacturing establishments per 1,000 urban population (+0.342) and percent of the population of native ethnicity (+0.311). Of the 17 independent variables examined, four are significantly correlated with the outcome measures: education less than Grade 9 with male SMRs (+0.287); number of manufacturing establishments per 1,000 population with male SMRs (-0.308); number of manufacturing establishments per 1,000 urban population with SMRs for females (+0.457) and both sexes combined (+0.342); parental employment in secondary industries with male SMRs (-0.311) and native ethnicity with SMRs for both males (+0.388) and total (+0.311).

Although no strong associations between any single independent variable and the disease outcomes emerge from this regression analysis, ecologic relationships linking manufacturing, native populations, low income and education to variations in childhood cancer mortality are suggested. However, of particular interest in this study of childhood cancer mortality in Ontario is the ability to investigate ecologic relationships by controlling for many variables simultaneously. Thus multiple regression analysis is used. The results of the multiple regression analysis of SMRs for males, females and both sexes combined against the 17 independent variables by county are presented in Table 5.3. A discussion of the results follows here.

Table 5.3 **Regression Results**

Dependent Variable	Coefficients		T - Statistic	R^2
Standardized Mortality Ratio Males	Constant	0.97	17.80***	
	Native	0.02	2.82**	0.15**
Standardized Mortality Ratio Females	Constant	0.59	4.39***	
	Manufacturing (1000 Urban)	0.16	3.44**	0.21***
Standardized Mortality Ratio Total	Constant	1.02	10.46***	
	Manufacturing (1000 Urban)	0.09	3.20**	
	Secondary	0.05	-2.62*	0.24**

*** $p < 0.001$
** $p < 0.01$
* $p < 0.05$

For cancer mortality of male children aged 0-19, the percentage of the population of native ethnicity is the only independent variable selected. As indicated by the R^2 value, 15.1 percent of the variation in male SMRs is explained by the proportion of native population in a county. The equation generated by the multiple regression analysis is: SMR(male) = 0.96678 + 0.02464Native. Thus, as the proportion of a county's population of native ethnicity increases, so does male childhood cancer mortality. Partial regression for each of the remaining 16 independent variables after native is removed fails to enter additional variables, however incidence of low income is positively and significantly correlated with elevated SMRs.

For females, the number of manufacturing establishments per 1,000 urban population is the independent variable that best accounts for the variations in SMRs for this sex. The R^2 value indicates that 20.9 percent of the variance in female cancer mortality is explained by this independent variable. The equation generated is: SMR(female) = 0.58797 + 0.16333 Manufacturing 1,000 Urban. Thus, as the number of manufacturing establishments per 1,000 urban population in a county increases, so does the female childhood cancer rate. None of the remaining independent variables were selected in successive steps of the regression analysis.

The multiple regression analysis for total childhood cancer SMRs results in the selection of two independent variables: number of manufacturing establishments per 1,000 urban population and the percent of the labour force 15 years and older employed in secondary sector industries. The R^2 values indicate that 11.7 percent of the variation in SMRs can be explained by the measure of manufacturing establishments, 11.9 percent of the variation is accounted for by the measure of secondary employment once the previous variable removed, for a total of 23.6 percent for the two variables combined. The equation for this dependent variable is: SMR(total) = 1.02106 + 0.08682 Manufacturing 1,000

Urban - 0.04906 Secondary. Thus, as the number of manufacturing establishments per 1,000 urban population of a county increases, so does overall childhood mortality and, after controlling for this effect, as the percent of the labour force employed in secondary industries decreases, so does childhood cancer mortality.

What do the results of this multiple regression analysis reveal about childhood cancer mortality in Ontario overall? First, that the 17 socio-demographic and environmental variables under investigation fail to explain a high proportion of the variance in cancer mortality in children. One reason for this could be explained by the choice of independent variables. As detailed in Chapter Four, the selection of socio-demographic and environmental variables was constrained by the availability of existing data bases and etiologic plausibility as determined by co-variants identified in a literature review. The fact that none of the 17 variables entered into the regression model successfully accounted for a large proportion of the variation in childhood cancer mortality across Ontario could be because the relevant correlates were not selected for investigation. It is appropriate at this point in the discussion to consider this possibility based on the results of an analysis of residuals from the multiple regression analysis, before proceeding with further interpretation of the findings of the regression analysis.

5.4 Residuals

The final analysis method employed in this investigation of childhood cancer mortality in Ontario is the examination of residuals. As discussed in Chapter Four, residual analysis facilitates an investigation of variations in SMRs that are not accounted for in the regression analysis. The pattern of residuals could, then, point to potential predictor variables which were not previously considered in the regression analysis. A visual

assessment of the spatial distribution of residuals as displayed in maps and a statistical analysis of spatial autocorrelation of residuals follow.

Figures 5.4, 5.5 and 5.6 illustrate the spatial distribution of the residuals generated from the regression analysis for male, female and total SMRs respectively. If a pattern or clustering of residuals is evident on these maps, that is areas of the province where SMRs in contiguous counties were poorly explained by the regressor variables, hypotheses regarding other factors which these counties may have in common and thus may be related to childhood cancer mortality can be generated. If however, the resulting pattern appears random across the province, no strong evidence that the addition of other environmental and socio-demographic variables would substantially increase the level of explanation is provided.

Residuals of male SMRs, displayed in Figure 5.4, range from -0.52 to +0.24. No clear pattern of residuals is evident, however Central and Northeastern Ontario demonstrate grouping of counties with high positive residuals whereas high negative results are found in the Southeastern and Southwestern regions of the province. Residuals for female SMRs, shown in Figure 5.5, range from -0.86 to +0.31 and demonstrate no clear pattern on the map. Contiguous counties with high positive and negative residuals are found in several regions of the province. Figure 5.6 displays total SMR residuals, which range from -0.39 to +0.19. No geographical pattern of residuals is evident, as counties with high positive and negative residuals are evenly distributed across Ontario. As indicated by all three maps, no obvious spatial pattern or clustering of residuals, suggestive of other factors which should have been entered into the regression analysis, is present .

The spatial autocorrelation statistics for the residual analysis are presented in Table 5.4. As discussed in section 5.2 of this chapter, spatial autocorrelation analysis is designed to evaluate the presence of clusters of counties with high positive or negative regression

Figure 5.4

REGRESSION RESIDUALS

Male, Standardized Mortality Ratio

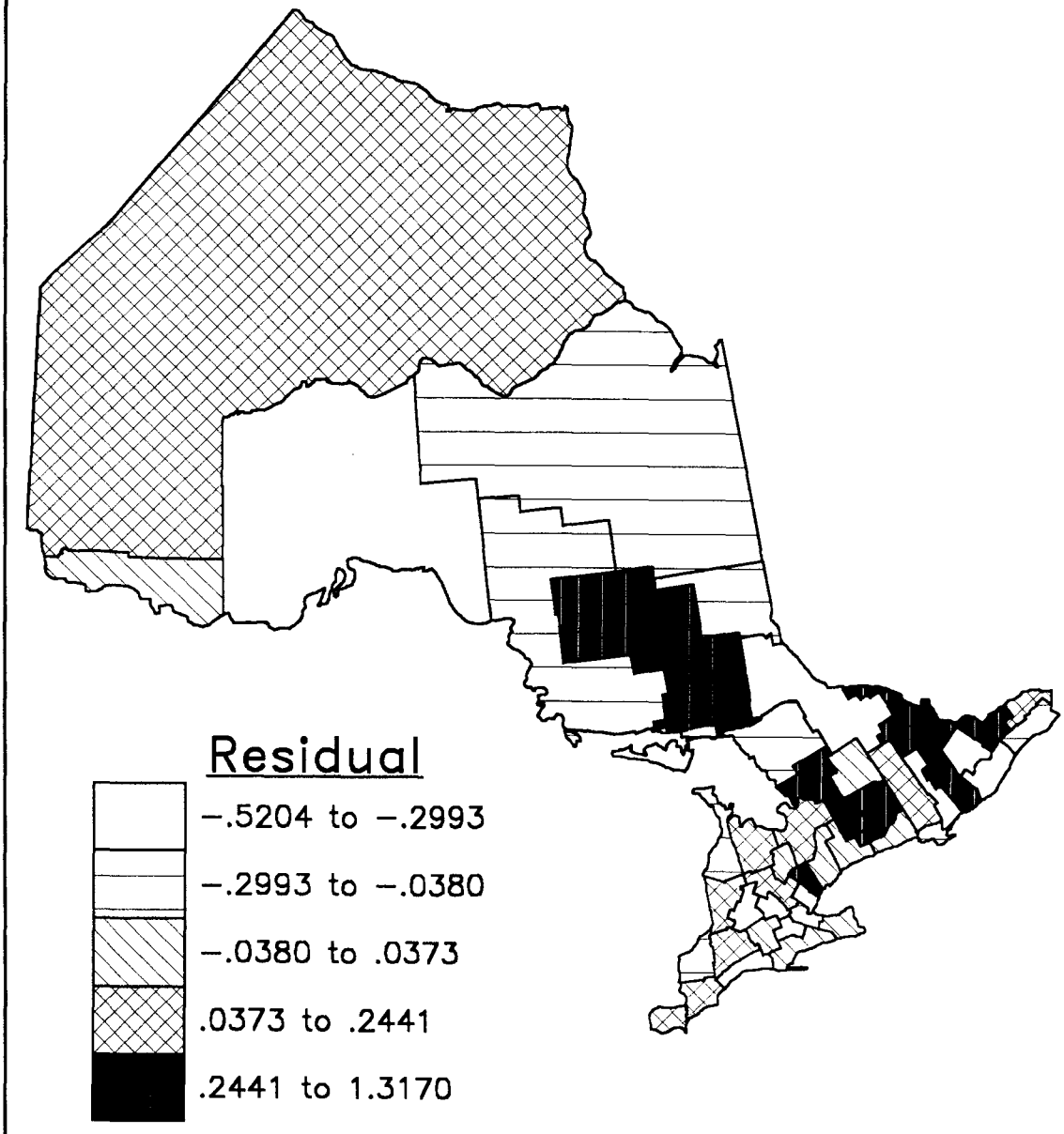


Figure 5.5

REGRESSION RESIDUALS

Female, Standardized Mortality Ratio

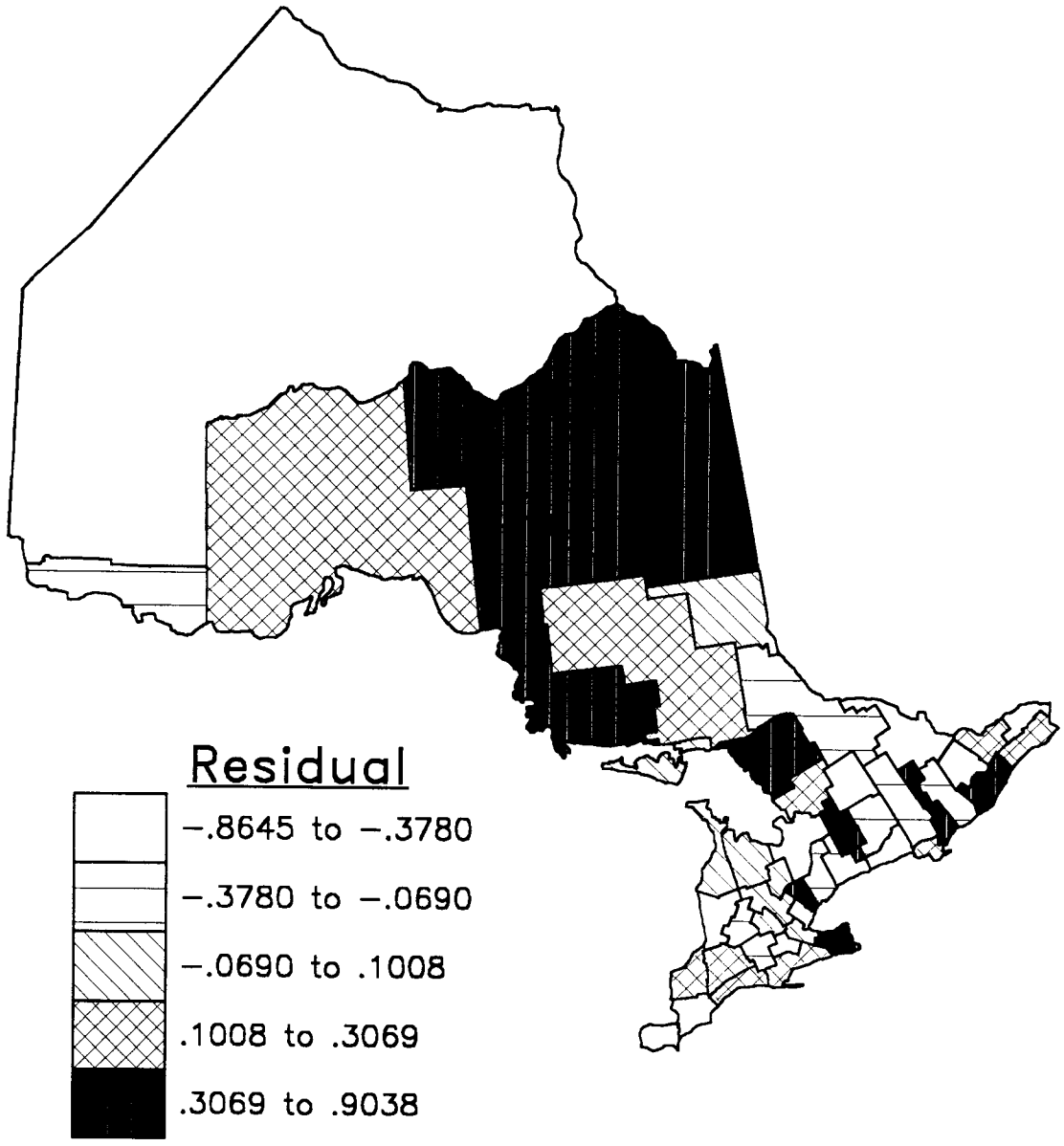


Figure 5.6

REGRESSION RESIDUALS

Total, Standardized Mortality Ratio

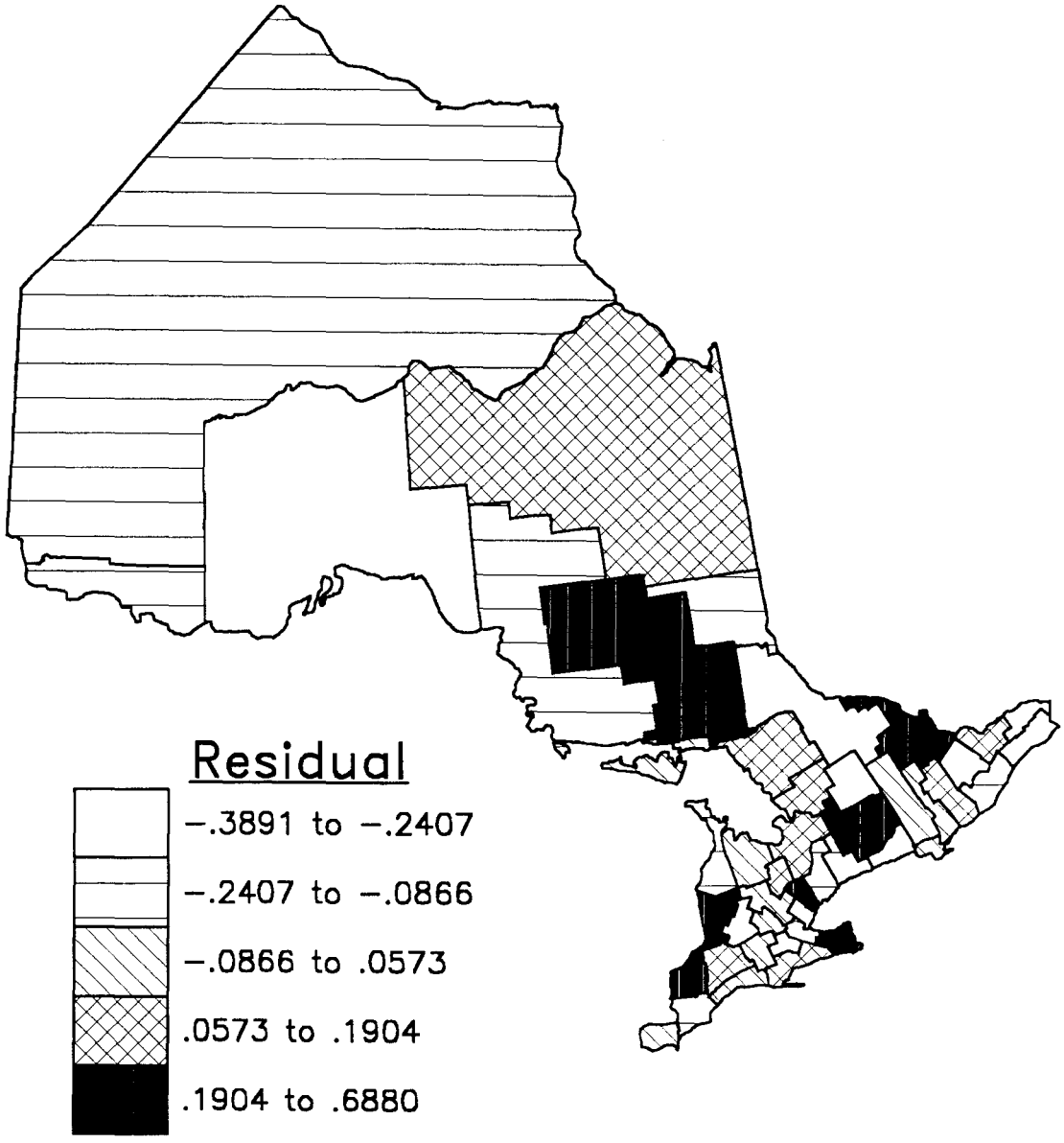


Table 5.4 Residuals - Spatial Autocorrelation Statistics Moran Index

Residuals Dependent Variables	I - Value	Z - Score*
Standardized Mortality Ratio Males	-0.05	-0.27
Standardized Mortality Ratio Females	+0.09	+1.24
Standardized Mortality Ratio Total	-0.12	-1.08

* Level of Significance: 5% ($\pm$1.96)

(Critical values of Z) 1% ($\pm$2.58)

residuals. A Moran Index of +1.0 is indicative of a spatial clustering of residuals. None of the Moran Index values in Table 5.4 approach +1.0 therefore no spatial autocorrelation is present in the residuals. The significance level is determined by a Z-score of 1.96 at the 5 percent level and 2.58. Again, none of the Z-scores in Table 5.4 approach this level, thus none are significant.

The results of this spatial autocorrelation analysis supports the conclusion reached in the previous discussion of the mapped pattern of residuals. No spatial pattern or clustering of residuals from the regression analysis is indicated which would point to additional independent variables warranting consideration in this investigation.

5.5 Discussion and Conclusions

The preceding discussion has considered one possible explanation for the inability of the regression analysis to yield factors which describe a large proportion of the variation in childhood cancer mortality rates across Ontario. The results of the residuals analysis provide no conclusive evidence that additional variables should have been included in the analysis. Several other lines of inquiry merit consideration as potential explanations of these findings.

As discussed throughout this thesis, this ecologic analysis of childhood cancer mortality in Ontario was conducted at the level of the county. Another possible explanation for the identification of no clear socio-demographic or environmental correlates of patterns of pediatric cancer mortality could be the scale of the investigation. Despite the justifications for proceeding at the level of the county, which include the availability of data in corresponding spatial units and the ability to generate stable rates of the rare outcome measure of interest, it could be that the county level is inappropriate for investigating cancer mortality in children. The county level is perhaps not capable of revealing patterns of

underlying factors relevant to malignancies in this population cohort, which does not necessarily rule out the influence of socio-demographic and environmental factors at other scales. Chapter Six will discuss suggestions for future ecologic studies of childhood cancer designed to overcome the inadequacies of large spatial units and small outcome measures, including analyses at the census tract level using cancer incidence data.

The review of previous literature relating to childhood cancer epidemiology contained in Chapter Two, suggests that genetic factors potentially play a major etiologic role. Further, the previous discussion of spatial autocorrelation analysis in this chapter indicates that the random and independent spatial distribution of childhood cancer across Ontario may point to underlying factors which are also randomly distributed and exhibit no clustering at the county level. If genetic factors are significant with respect to cancer in children and the assertion that approximately 80 percent of human cancer is attributable to environmental influences is not applicable to malignancies in this population sub-group, then the results of this investigation seem appropriate. Genetic characteristics, which are likely to be distributed randomly in individuals across the province, cannot adequately be investigated in a county level ecologic study.

Although none of the independent variables considered in this investigation describe a large proportion of the spatial variation of childhood cancer mortality in Ontario, four of the variables do emerge which demonstrate a reasonably high degree of association with the Ontario cancer data set. Native ethnicity, incidence of low income, number of manufacturing establishments per 1,000 urban population and employment in secondary industries accounted for an average of 15 to 20 percent of the variation in the sex-specific or total childhood cancer SMRs. Several hypotheses can be drawn from this.

A high proportion of persons of native origin in a county could indicate the presence of a genetic influence in childhood cancer mortality, although there is no evidence

in the literature to suggest such an association. Remoteness, low education and language barriers all act as deterents for studying native populations (Reynolds, 1989). However, one plausible explanation for the correlation of counties with substantial native communities and high incidence of low income with counties exhibiting high childhood cancer mortality could relate to factors which influence mortality rather than cancer incidence. Low education, low income, language barriers and geographic remoteness could be indicative of poor access to health care facilities, early cancer diagnosis and treatment for children in native populations. If this is the case, the results of this analysis would indicate that native children do not necessarily have a higher incidence of cancer but that more native children die as a result of cancer.

The number of manufacturing establishments relative to the urban population of a county is found to be correlated with patterns of childhood cancer mortality in females and both sexes combined. This variable serves as a potential proxy measure of a wide range of environmental conditions, such as air and water quality and industrial/chemical contaminants. The relationship of this variable with childhood cancer mortality, then, could be suggestive of an environmental etiology which could only be uncovered through the incorporation of more direct environmental exposure measures.

The final ecologic correlate which demonstrates a significant correlation with childhood cancer mortality in Ontario is the percentage of the labour force employed in secondary sector industries. The literature review of previous investigations of childhood cancer presented in Chapter Two, considered a relatively substantive set of published studies looking at the relationship between malignancies in children and parental occupation. Although few of these studies clearly demonstrated a causal association, methodological issues and inconsistencies between studies make the evidence inconclusive. The measure of employment in secondary industries used in the Ontario investigation

includes occupations such as material processing, machining, construction and transport equipment operation. The association between high childhood cancer mortality and high proportions of employment in secondary industries in Ontario counties could point to the appropriateness of further investigation of this hypothesis.

The findings of this investigation of socio-demographic and environmental correlates of childhood cancer mortality in Ontario 1976-1985 will be discussed further in Chapter Six, as they relate to the research objectives set out in this thesis and as they point to further investigations of childhood cancer.

CHAPTER SIX
SUMMARY AND CONCLUSION

6.1 Summary

This thesis has examined the ecologic correlates of childhood cancer mortality in Ontario for 1976 to 1985. In so doing, the thesis had four objectives: (1) to describe county level rates of childhood cancer mortality in Ontario for the period 1976 to 1985; (2) to analyse spatial variations in pediatric cancer mortality in Ontario; (3) to determine to what extent spatial variations in childhood cancer mortality are related to variations in selected ecologic variables; and, (4) to comment on future directions for research in childhood cancer epidemiology and to assess the utility of the ecologic methodology in such investigations.

An extensive review of previous literature on childhood cancer, undertaken to set the context for the Ontario study, revealed several important findings. Firstly, that significant differences exist between childhood and adult cancers and the role of environmental factors in pediatric cancer etiology has not been established. Secondly, that few studies of the spatial variation of childhood cancer mortality have been undertaken, and of these investigations, little valid evidence of clustering or systematic variations suggestive of etiologic clues about risk factors was found. Thirdly, that previous investigations of childhood cancer have considered a broad range of potential correlates, including: demographic and genetic factors such as age, ethnicity and sex; environmental factors such as radiation, chemicals, parental occupation, socio-economic status and urbanization; and that pediatric cancer may be influenced by the interaction of genetic and environmental factors.

In order to address these research objectives, an evaluation of the ecologic methodology and the use of mortality data was undertaken. Cancer mortality frequencies for children aged 0 to 19 were obtained from the Ontario Cancer Registry by age, sex and county for the ten year period 1976 to 1985. Standardized mortality ratios were calculated from the data for use as the independent variable. Based on the broad set of ecologic correlates of childhood cancer emerging from the literature review, 17 independent variables were selected from available data sources for inclusion in the Ontario study.

Four stages of data analysis designed to meet the research objectives were undertaken: geographic analysis using maps, spatial autocorrelation, regression analysis, and an evaluation of regression residuals. The spatial analysis of childhood cancer mortality across Ontario revealed no consistent or distinctive pattern of geographic variation. Spatial autocorrelation analysis confirmed this finding, as spatial clustering of childhood cancer mortality across Ontario was not present at the county level.

Correlation and multiple regression analysis failed to yield independent variables which are strongly associated with variations in childhood cancer mortality across Ontario, however, several socio-demographic and environmental correlates did emerge which demonstrate a significant and reasonably high degree of association with the Ontario data set. Native ethnicity, incidence of low income, number of manufacturing establishments per 1,000 urban population and employment in secondary industries were selected in the multiple regression analysis as the variables exhibiting the highest degree of association with variations in sex-specific or total childhood cancer SMRs in Ontario.

6.2 Conclusions

To examine how the findings of this investigation of childhood cancer mortality in Ontario compare with the findings of other studies of cancer in this population subgroup,

two issues must be considered. Firstly, with regard to spatial variations in childhood cancer, the results of the Ontario study are consistent with the rare examples in the previous literature of investigations of this phenomenon in other geographical locations. In general, childhood cancer rates demonstrate no systematic or recurring pattern over space. Those few individual studies which did identify geographical patterns of childhood cancer rates, achieved their results in investigations of regional patterns across a nation (i.e. Greenberg, 1983), or cancer clusters at a localized scale (i.e. Craft and Openshaw, 1985). No examples of studies examining variations in childhood cancer mortality at the county scale in a jurisdiction similar to Ontario were found in the literature to provide a basis for direct comparison.

The second consideration in this evaluation of the comparability of the Ontario study to other investigations of childhood cancer concerns the ecologic correlates identified which demonstrated an association with pediatric cancer. As discussed in the literature review presented in Chapter Two, the studies which analysed cancer in this age cohort considered a broad range of potential correlates and, as a group, fail to identify strong and consistent associations between childhood cancer and specific environmental or socio-demographic variables. Even among those studies investigating associations between malignancies in childhood and parental occupation, which constitute a large proportion of previous investigations, evidence supportive of a clear and strong association is inconclusive. Thus, the findings of the Ontario study are consistent with the literature, as the relationships which result are similar both in nature and in the strength of association to those previously demonstrated.

The research in this thesis makes several contributions to both childhood cancer epidemiology in general as well as to the expanding literature within medical geography concerning the environmental determinants of health. This study represents an initial

investigation of cancer in a population subgroup which has received very little attention in the literature. In addition, this thesis constitutes the first investigation conducted with childhood cancer data in the Ontario Cancer Registry. As cancer is the most common natural cause of death for 0 to 19 year old populations in most developed countries, the need for investigations such as this is clearly indicated. This thesis did not attempt to provide evidence sufficient to demonstrate causation between childhood cancer mortality and etiologic agents, but rather to develop hypotheses regarding the potential role of socio-economic and environmental factors as ecologic determinants of health. In this sense, this work has successfully met its objectives.

6.3 Directions for Further Research

The results of this analysis have several implications for future work in spatial variations in childhood cancer mortality and morbidity. The assertion that the county level may be inappropriate for investigating cancer mortality in children has been attributed primarily to the rarity of malignancies in this age group. Studies conducted at the county or census tract level for the entire Canadian population aged 0 to 19, may reveal regional differences in the geographical distribution of cancer mortality not detected when considering one province in isolation, and worthy of further research. In order to facilitate such an examination, a national cancer registry or provincial registries similar to the OCR would need to be developed. This would, however, not overcome the problem of the small number of cases. Conversely, investigations at a smaller scale, such as the census tract, may be necessary to reveal ecologic relationships between childhood cancer and socio-demographic and environmental variables which also vary at this scale. The use of longer time periods could provide sufficient mortality frequencies necessary to generate stable

rates, however difficulties with changing administrative boundaries over time and accuracy of the data would have to be overcome.

The second area of consideration for further research into geographical variations in childhood cancer is the use of incidence data in addition to mortality data. The improvements in early diagnosis and therapy introduced in the past forty years have reduced childhood cancer mortality significantly, but have had little apparent impact on morbidity (Young et al., 1986). An ecologic analysis of incidence rates, then, may yield numbers high enough to generate stable rates for smaller geographic areas. Another advantage of the incorporation of incidence data is the ability to investigate specific cancer sites (i.e. leukaemia and Hodgkin's disease) and to compare age cohorts (i.e. 0-4, 5-9, 10-14, 15-19), which is not possible with mortality data in Ontario. The existing literature dealing with spatial variation in childhood cancer rates indicates that the various forms of childhood malignancies do exhibit distinctive patterns of occurrence (Greenberg and Shuster, 1985). As well, the previous literature suggests that age variations may reflect important etiological influences (Miller, 1967). Thus, the ability to disaggregate childhood cancer data by site and age through the use of incidence data may reveal important spatial relationships.

The final set of suggestions for further investigations into childhood cancer are derived from the hypotheses generated by this analysis in Ontario. Studies on cancer among low income populations, particularly Native children, are needed. The development of comprehensive environmental data bases to allow for more direct incorporation of exposure measures in cancer studies is called for. In addition, the association between childhood cancer and parental occupation clearly warrants further investigation using study designs such as case-control.

Finally, some comment regarding the appropriateness of the ecologic methodology in investigations of environment and health is warranted. As indicated previously in this thesis and elsewhere, ecologic studies have been criticized for their inability to demonstrate causal associations between disease outcomes and environmental variables. However, as demonstrated by the results of this investigation of childhood cancer mortality in Ontario, the ecologic methodology successfully facilitated an examination of geographical variations in both disease rates and postulated determinants, thus pointing the direction for further analysis of cancer in children by other research designs capable of demonstrating cause and effect.

APPENDIX ONE

Ontario Counties 1981

Appendix One

ONTARIO COUNTIES 1981

01	Algoma	25	Muskoka
02	Brant	26	Niagara
03	Bruce	27	Nipissing
04	Cochrane	28	Northumberland
05	Dufferin	29	Ottawa/Carleton
06	Durham	30	Oxford
07	Elgin	31	Parry Sound
08	Essex	32	Peel
09	Frontenac	33	Perth
10	Grey	34	Peterborough
11	Haldimand/Norfolk	35	Prescott/Russell
12	Haliburton	36	Prince Edward
13	Halton	37	Rainy River
14	Hamilton/Wentworth	38	Renfrew
15	Hastings	39	Simcoe
16	Huron	40	Stormont/Dundas/Glengarry
17	Kenora	41	Sudbury
18	Kent	42	Thunder Bay
19	Lambton	43	Timiskaming
20	Lanark	44	Victoria
21	Leeds/Grenville	45	Waterloo
22	Lennox/Addington	46	Wellington
23	Manitoulin	47	York/Metropolitan Toronto
24	Middlesex		

APPENDIX TWO

Data for Independent Variables

Appendix Two — Data For Independent Variables

	County	01 Urban (5000) %	02 Urban (Census) %	03 Pop. Density	04 Education <Grade 9 %	05 Incidence of Low Income %	06 Median Income $	07 Mfrg. (1,000) #	08 Mfrg. (1,000 Urban) #	09 Primary Employment %
1	Algoma	74.44	74.52	2.6	19.43	10.2	23,957	0.57	0.79	0.44
2	Brant	78.33	79.49	95.6	19.62	11.5	20,907	2.03	2.55	2.29
3	Bruce	19.84	42.42	14.8	21.35	11.2	19,127	1.20	2.82	3.05
4	Cochrane	72.27	69.01	0.7	26.35	11.1	21,990	0.57	0.79	0.99
5	Dufferin	44.12	57.24	20.9	15.89	8.9	22,579	1.74	3.03	2.15
6	Durham	88.03	82.14	113.9	13.65	8.3	25,885	1.29	1.57	0.70
7	Elgin	47.95	54.06	37.1	20.45	10.9	19,717	1.45	2.68	5.57
8	Essex	72.99	80.10	167.8	19.77	13.9	20,934	1.70	2.12	1.34
9	Frontenac	48.66	72.08	28.3	14.43	12.5	20,032	0.87	1.21	0.48
10	Grey	35.49	50.12	16.4	22.82	13.9	16,642	1.73	3.46	1.92
11	Haldimand/Norfolk	69.71	42.18	30.7	23.13	10.4	20,316	1.44	3.42	9.40
12	Haliburton	—	—	2.7	23.66	15.6	14,044	1.05	—	0.66
13	Halton	100.0	91.48	264.8	9.33	6.3	29,496	1.80	1.97	0.64
14	Hamilton/Wentworth	91.68	91.22	369.7	20.16	14.0	21,766	1.40	1.53	0.57
15	Hastings	46.75	64.73	17.9	18.33	13.1	18,468	1.71	2.64	0.72
16	Huron	13.04	35.83	16.5	21.96	13.2	17,362	1.78	4.98	4.14
17	Kenora	27.70	44.81	0.1	26.13	7.6	21,129	0.75	1.52	1.02
18	Kent	49.02	64.48	42.9	21.61	12.6	19,488	1.68	2.61	2.86
19	Lambton	41.23	69.03	41.2	15.49	9.8	23,136	1.02	1.48	1.53
20	Lanark	44.05	54.73	14.9	17.62	11.0	19,466	1.79	3.28	1.14
21	Leeds/Grenville	24.59	44.22	23.9	15.82	10.1	19,479	1.32	2.99	1.59
22	Lennox/Addington	—	33.03	11.6	18.25	12.1	19,297	0.70	2.11	1.41
23	Manitoulin	—	13.70	3.0	30.69	16.4	13,020	0.82	6.00	1.61
24	Middlesex	82.67	86.56	94.7	13.96	11.6	20,794	1.35	1.55	1.13
25	Muskoka	75.74	42.51	9.5	18.82	12.0	16,940	1.38	3.25	0.71
26	Niagara	95.70	87.53	199.0	19.22	12.3	21,460	1.41	1.62	1.29
27	Nipissing	71.41	74.43	4.5	20.99	14.0	19,338	1.08	1.46	0.55
28	Northumberland	32.91	47.99	30.8	18.07	10.0	19,454	1.63	3.40	2.11
29	Ottawa/Carleton	89.77	89.72	198.3	10.56	10.9	24,860	0.63	0.70	0.25
30	Oxford	53.06	59.95	42.3	21.48	9.9	20,204	2.11	3.51	4.44
31	Parry Sound	18.26	28.51	3.3	24.81	15.5	14,348	1.28	4.48	0.64
32	Peel	100.0	88.91	400.4	11.48	7.5	29,189	2.25	2.53	0.30
33	Perth	47.34	61.14	30.2	21.51	10.4	19,280	1.77	2.90	3.88
34	Peterborough	59.17	66.68	25.9	16.20	12.1	19,804	1.22	1.83	0.86
35	Prescott/Russell	18.72	41.77	26.4	26.64	12.6	20,154	1.44	3.45	2.48
36	Prince Edward	—	24.37	21.3	21.17	11.5	18,195	1.61	6.67	3.82
37	Rainy River	39.06	62.75	1.4	22.23	9.4	20,176	0.75	1.52	1.94
38	Renfrew	44.29	57.40	11.4	22.52	12.7	18,319	1.35	2.35	1.12
39	Simcoe	44.10	61.82	46.5	17.26	10.2	20,261	1.35	2.18	1.22
40	Stormont/Dundas/Glengarry	31.33	56.95	30.6	22.95	14.8	17,787	1.31	2.29	1.94
41	Sudbury	37.63	82.36	4.1	22.14	13.3	21,066	0.56	0.68	0.51
42	Thunder Bay	73.04	82.97	1.4	19.73	8.9	24,265	0.80	0.96	0.66
43	Timiskaming	43.05	66.70	3.2	26.11	15.6	16,291	1.19	1.78	2.48
44	Victoria	28.42	35.36	15.6	20.31	10.6	17,465	1.42	4.02	1.43
45	Waterloo	87.18	91.23	224.7	19.20	10.6	22,022	2.44	2.92	0.56
46	Wellington	59.70	70.59	48.7	17.20	9.1	21,759	2.06	2.92	1.73
47	York/Metropolitan Toronto	98.52	98.06	1001.6	17.49	12.5	26,605	2.38	2.43	0.16

Appendix Two (Continued) Data For Independent Variables

	County	10 Secondary Employ-ment %	11 Tertiary Employ-ment %	12 British %	13 French %	14 Dutch %	15 German %	16 Italian %	17 Native %
1	Algoma	2.22	55.07	46.17	15.22	0.88	2.75	9.30	3.18
2	Brant	7.46	54.40	62.19	2.65	3.24	3.78	2.95	4.59
3	Bruce	3.94	50.39	64.72	3.22	2.67	14.21	0.36	1.87
4	Cochrane	1.90	53.65	24.69	49.25	0.53	1.43	2.49	6.08
5	Dufferin	4.80	55.40	74.09	2.02	4.30	3.39	1.12	0.36
6	Durham	5.11	60.55	67.43	3.48	3.44	3.13	2.17	0.32
7	Elgin	6.29	48.58	65.88	2.46	5.78	7.66	0.74	0.31
8	Essex	4.59	58.02	41.65	15.36	1.13	4.83	7.30	0.24
9	Frontenac	2.24	64.86	70.09	4.57	2.56	2.72	1.13	0.28
10	Grey	3.91	33.16	73.63	1.61	2.58	9.58	0.23	0.21
11	Haldimand/Norfolk	4.00	44.13	59.46	2.26	5.58	7.72	0.66	0.97
12	Haliburton	2.73	57.60	79.02	2.46	1.48	3.94	0.67	0.18
13	Halton	3.78	71.03	63.17	3.41	3.49	4.02	2.91	0.16
14	Hamilton/Wentworth	4.99	58.32	53.47	3.38	2.75	3.74	9.45	0.68
15	Hastings	6.99	59.85	71.94	5.04	2.71	3.16	0.75	1.75
16	Huron	3.96	47.46	69.56	2.52	7.58	8.42	0.14	0.14
17	Kenora	1.63	57.10	34.34	6.96	0.89	5.06	1.19	26.12
18	Kent	5.28	52.52	59.74	9.77	6.22	2.59	1.28	0.62
19	Lambton	2.69	57.49	64.59	5.65	5.33	3.02	1.88	2.44
20	Lanark	6.82	59.97	80.96	4.81	1.08	1.80	0.35	0.22
21	Leeds/Grenville	6.54	57.75	76.48	4.99	3.89	2.29	0.44	0.25
22	Lennox/Addington	4.64	57.24	74.51	3.75	3.58	3.27	0.33	0.37
23	Manitoulin	1.49	55.33	56.28	2.57	0.55	0.73	0.27	30.61
24	Middlesex	4.42	88.90	63.97	2.71	4.23	4.14	2.35	1.05
25	Muskoka	3.04	59.82	75.41	3.53	1.25	4.05	0.65	0.86
26	Niagara	3.80	58.10	50.16	6.48	4.18	6.63	8.18	0.37
27	Nipissing	2.75	64.30	40.49	33.03	0.65	2.86	2.09	1.78
28	Northumberland	6.39	53.09	78.33	2.68	3.80	2.18	0.57	0.51
29	Ottawa/Carleton	1.45	82.95	48.67	20.86	1.19	2.69	2.89	0.30
30	Oxford	5.80	50.75	65.02	2.21	7.31	7.21	1.04	0.24
31	Parry Sound	2.13	57.69	68.63	5.30	1.64	5.60	1.24	2.28
32	Peel	6.06	67.12	53.24	3.05	1.85	3.10	7.83	0.38
33	Perth	7.88	49.90	59.12	1.67	4.58	17.23	0.87	0.21
34	Peterborough	4.17	62.91	78.04	2.83	2.06	2.38	0.98	1.05
35	Prescott/Russell	4.72	54.13	16.73	75.83	0.87	0.84	0.36	0.11
36	Prince Edward	3.77	56.60	77.89	2.09	4.15	2.76	0.70	0.56
37	Rainy River	2.67	55.17	43.81	6.52	2.17	4.98	0.77	9.18
38	Renfrew	4.75	59.74	51.42	10.93	1.49	12.72	0.23	0.78
39	Simcoe	6.32	60.12	68.99	7.00	3.00	3.07	1.68	0.78
40	Stormont/Dundas/Glengarry	5.54	54.63	46.84	32.78	3.27	1.92	0.61	0.83
41	Sudbury	1.62	58.27	32.75	35.25	0.67	2.26	4.68	1.81
42	Thunder Bay	2.33	58.37	38.96	8.34	1.38	2.78	6.78	3.78
43	Timiskaming	1.61	55.20	46.58	30.34	0.76	2.70	0.87	1.19
44	Victoria	5.75	52.45	80.76	2.36	2.16	2.44	0.51	0.34
45	Waterloo	8.88	59.08	45.15	3.11	1.92	21.26	1.10	0.24
46	Wellington	6.73	58.54	62.82	1.84	4.72	6.80	5.29	0.20
47	York/Metropolitan Toronto	5.53	71.03	67.43	3.48	3.44	3.13	2.17	0.32

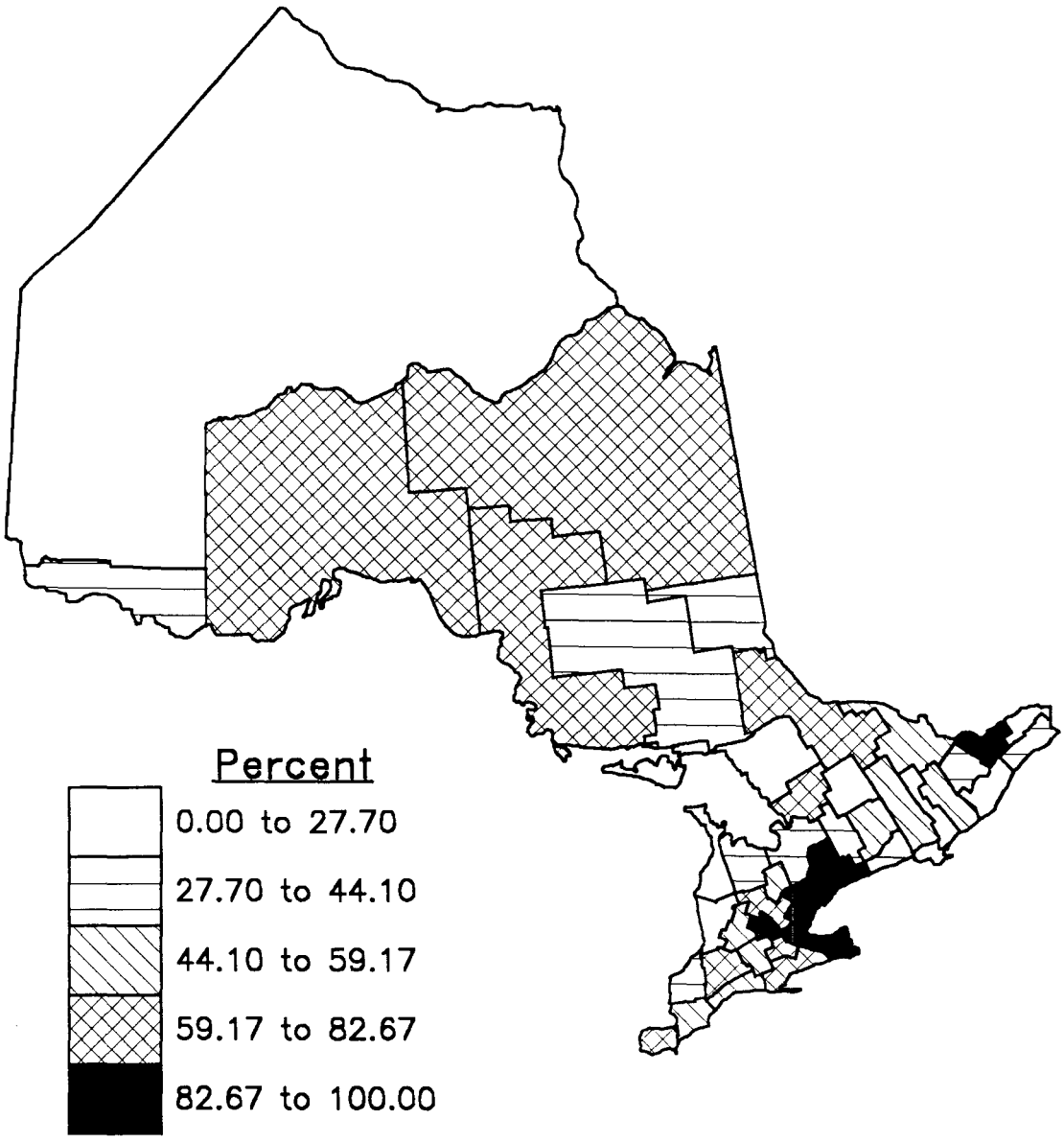
APPENDIX THREE

Maps of Independent Variables

Appendix 3.1

URBANIZATION 1981

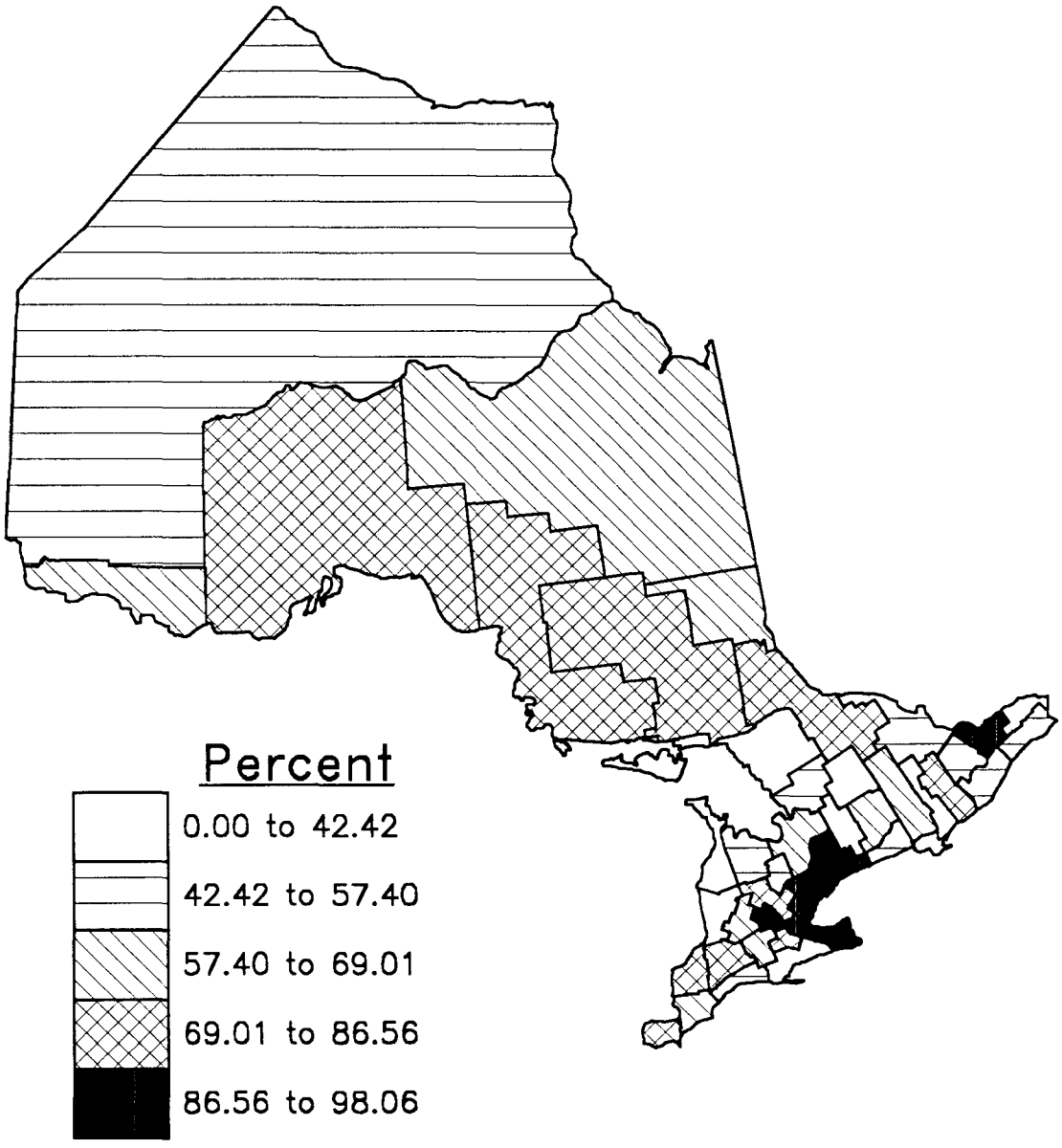
Percent of Population in Places Over 5,000



Appendix 3.2

URBANIZATION 1980

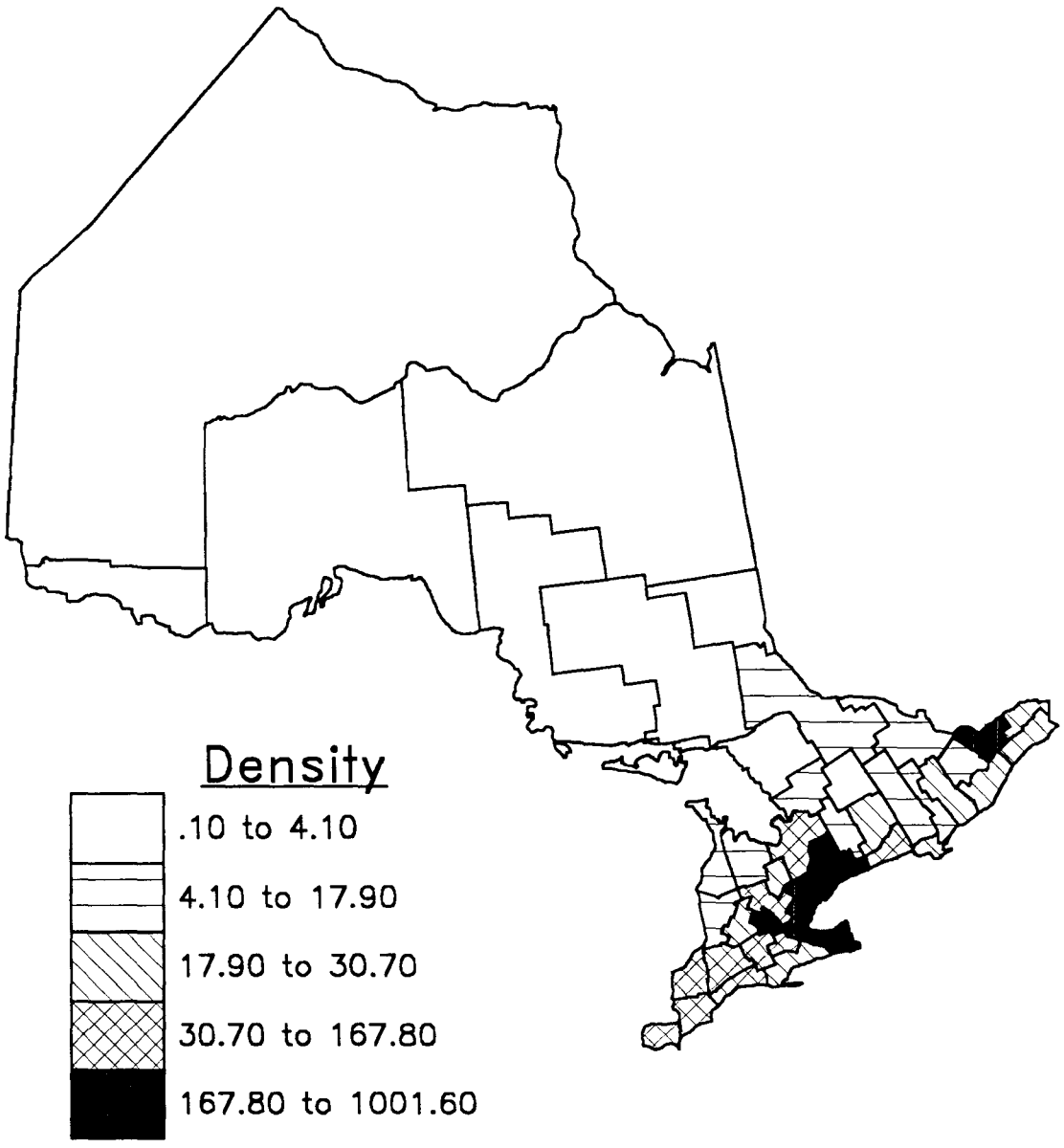
Percent of Population Urban



Appendix 3.3

POPULATION DENSITY 1981

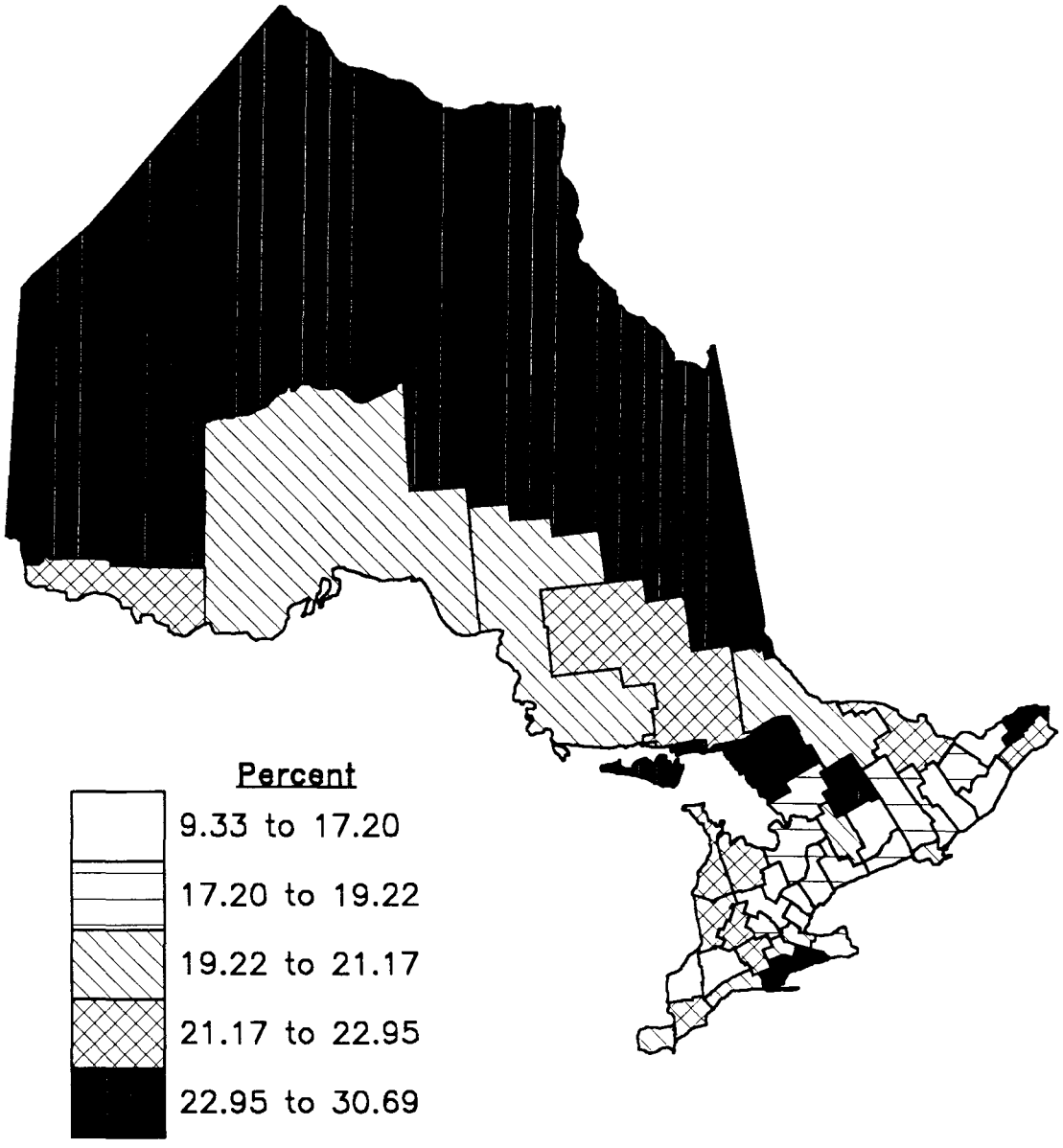
Density per Square Kilometre



Appendix 3.4

PARENTAL EDUCATION 1981

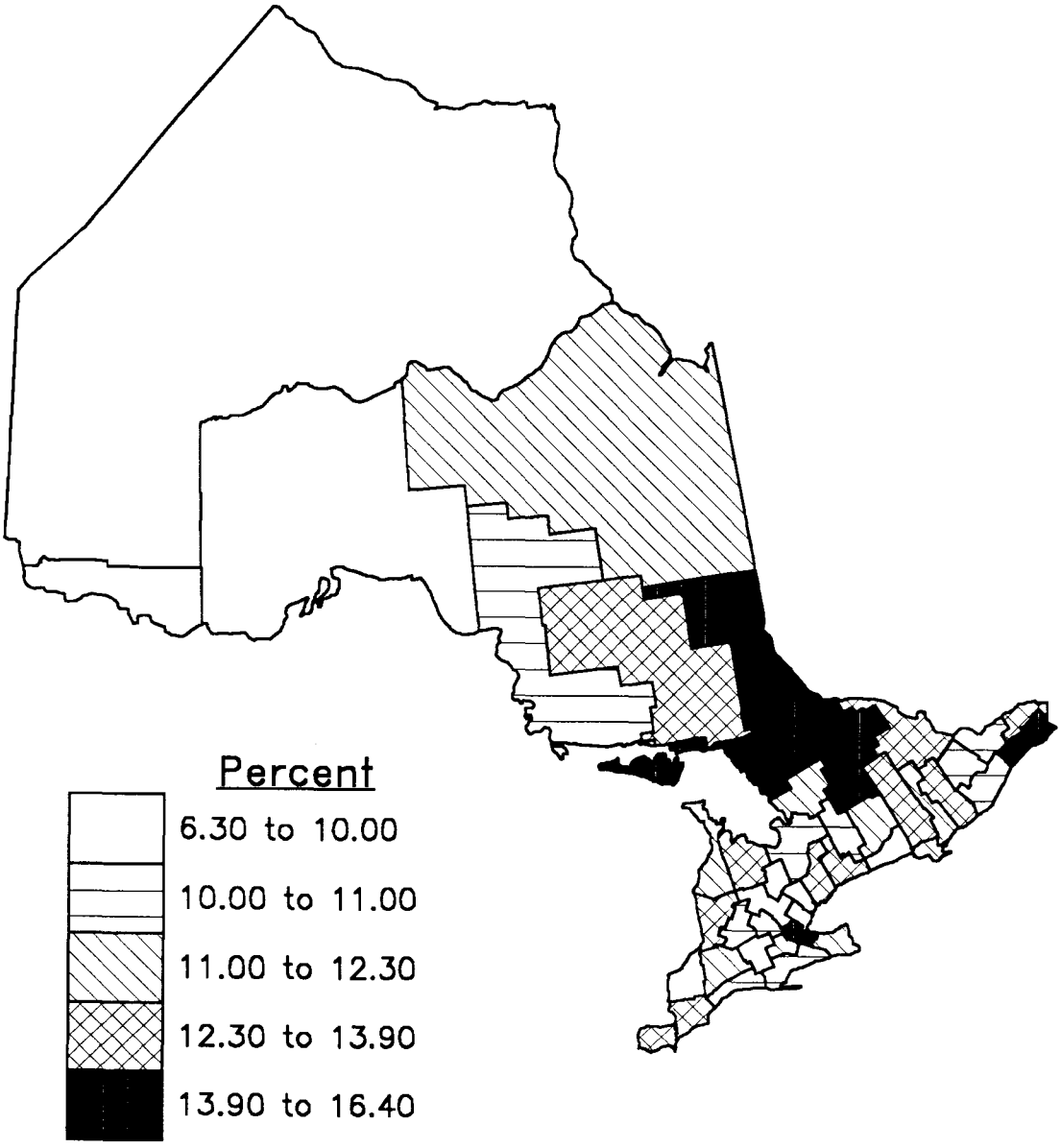
Percent with Less Than Grade Nine



Appendix 3.5

INCIDENCE OF LOW INCOME 1981

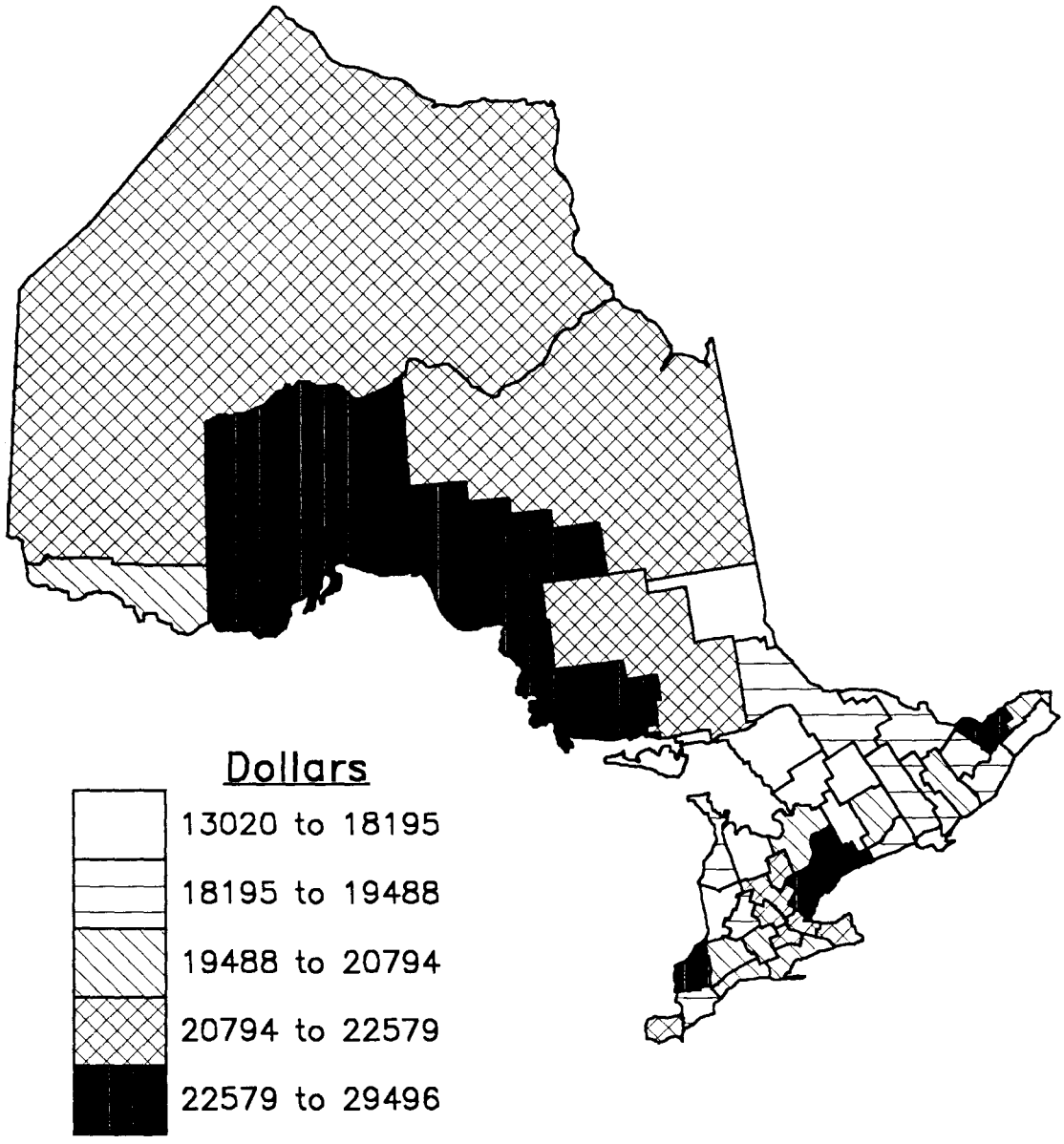
Percent of Families Below Low Income Cutoff



Appendix 3.6

MEDIAN HOUSEHOLD INCOME 1981

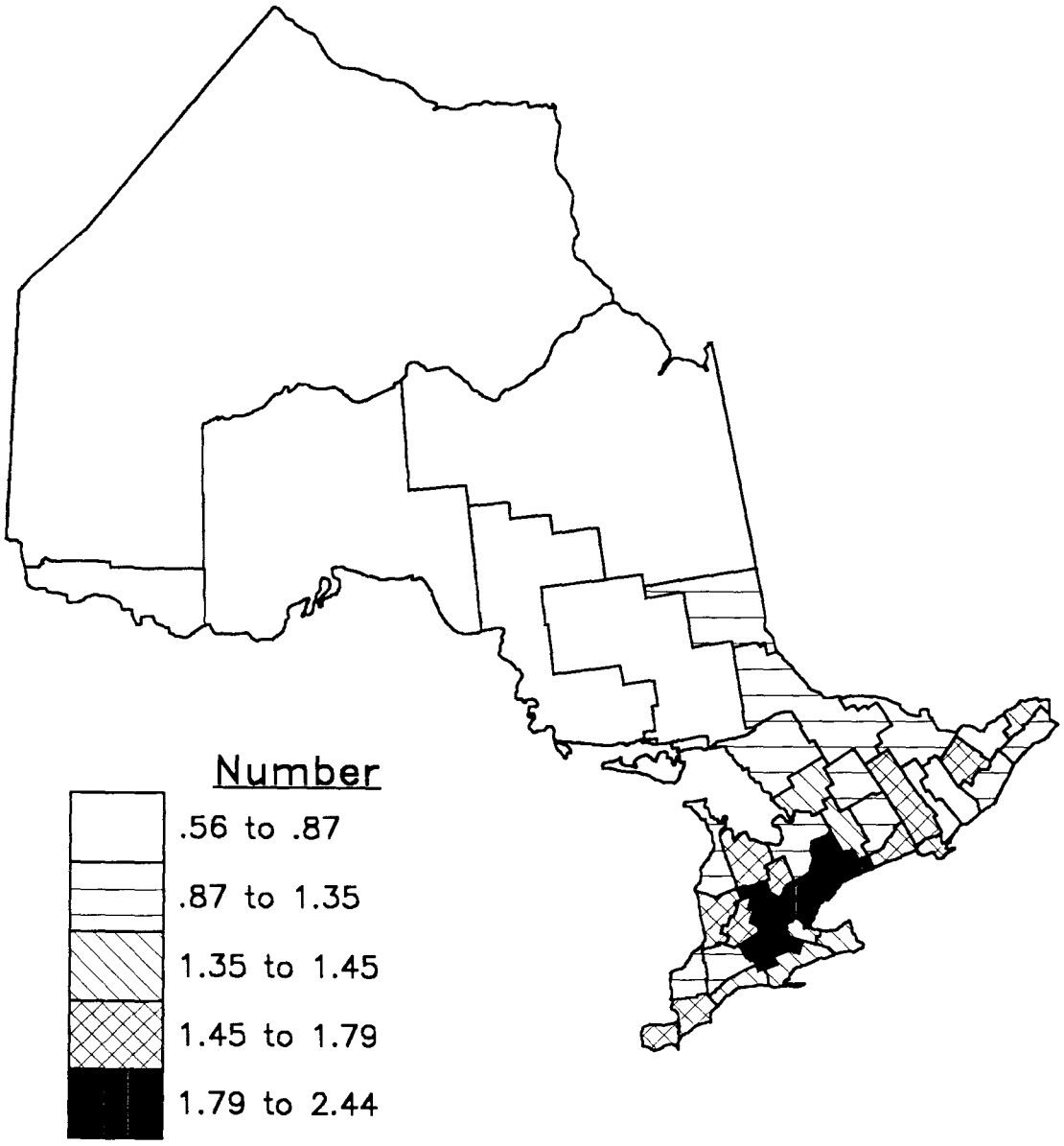
in Dollars



Appendix 3.7

MANUFACTURING ESTABLISHMENTS 1981

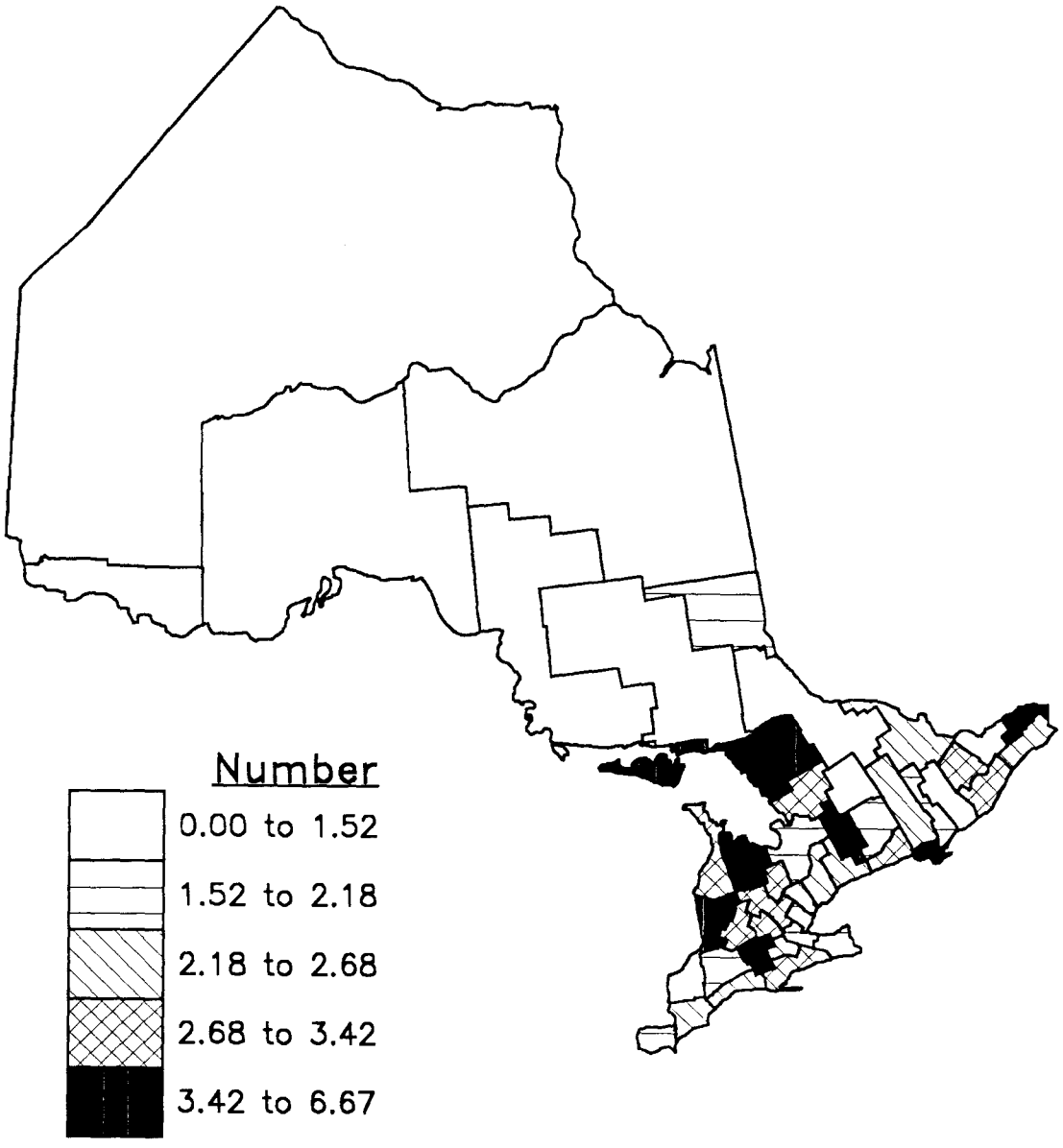
per 1,000 Population



Appendix 3.8

MANUFACTURING ESTABLISHMENTS 1981

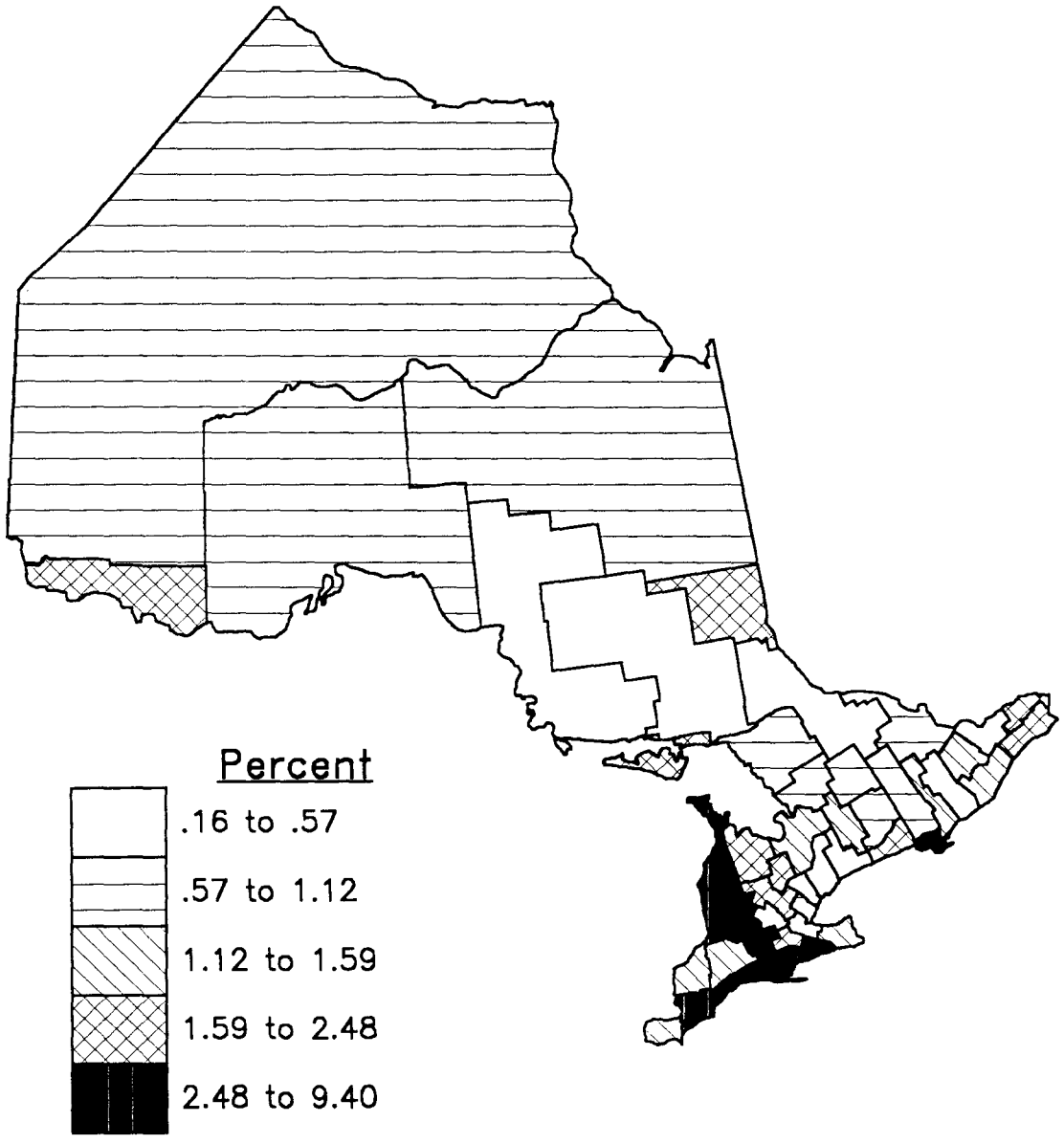
per 1,000 Urban Population



Appendix 3.9

PRIMARY INDUSTRIES 1981

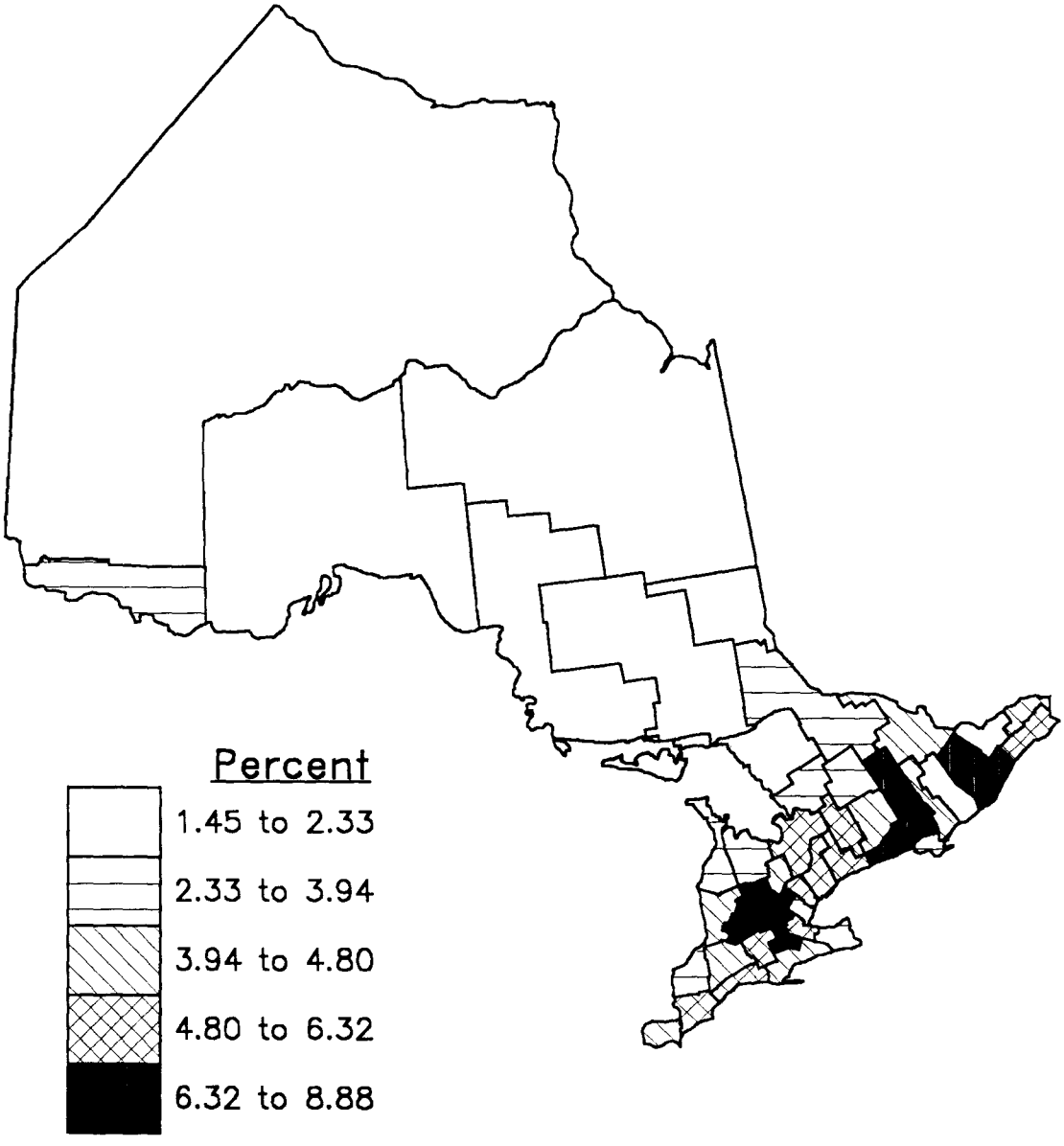
Percent Employed



Appendix 3.10

SECONDARY INDUSTRIES 1981

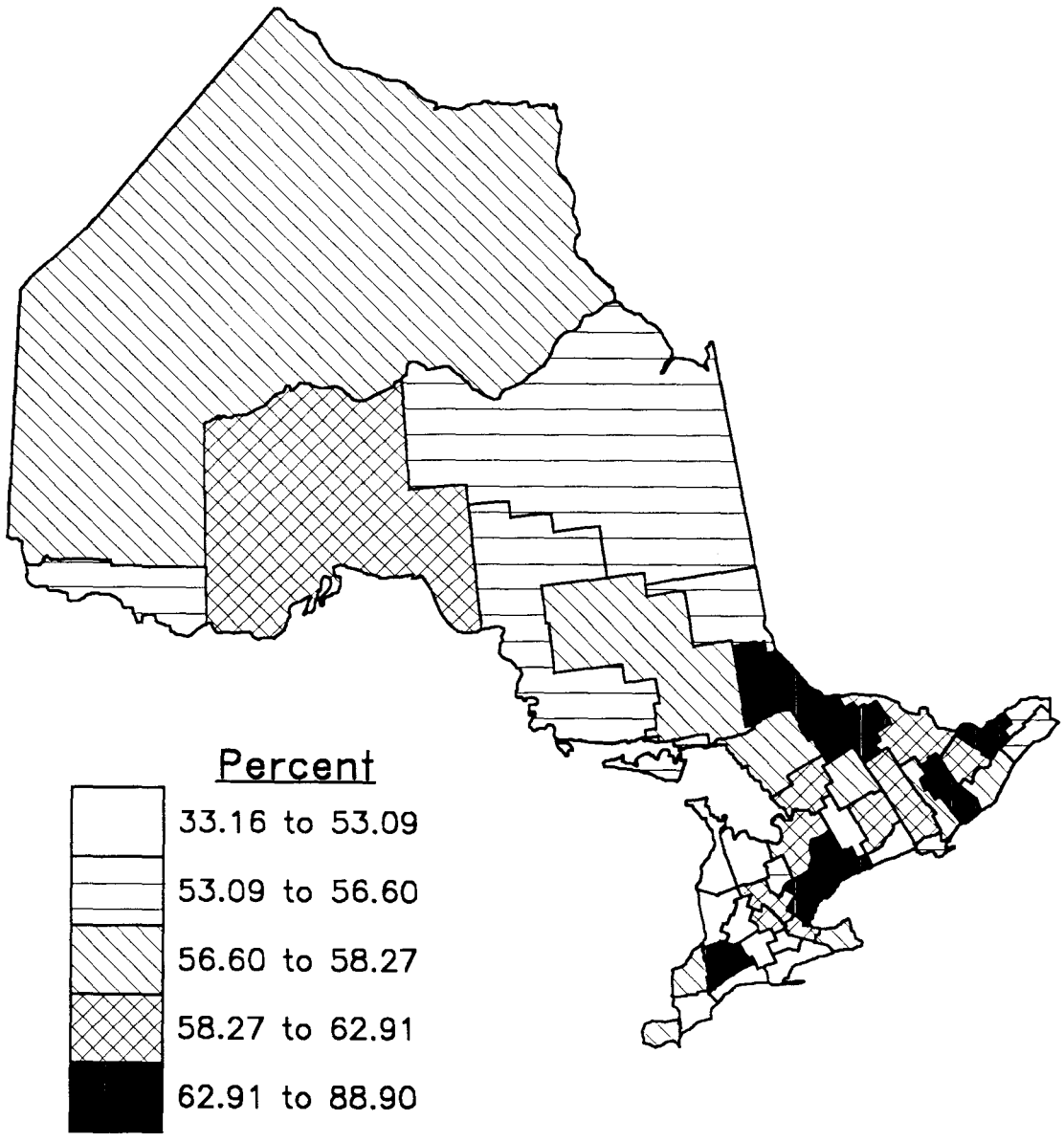
Percent Employed



Appendix 3.11

TERTIARY INDUSTRIES 1981

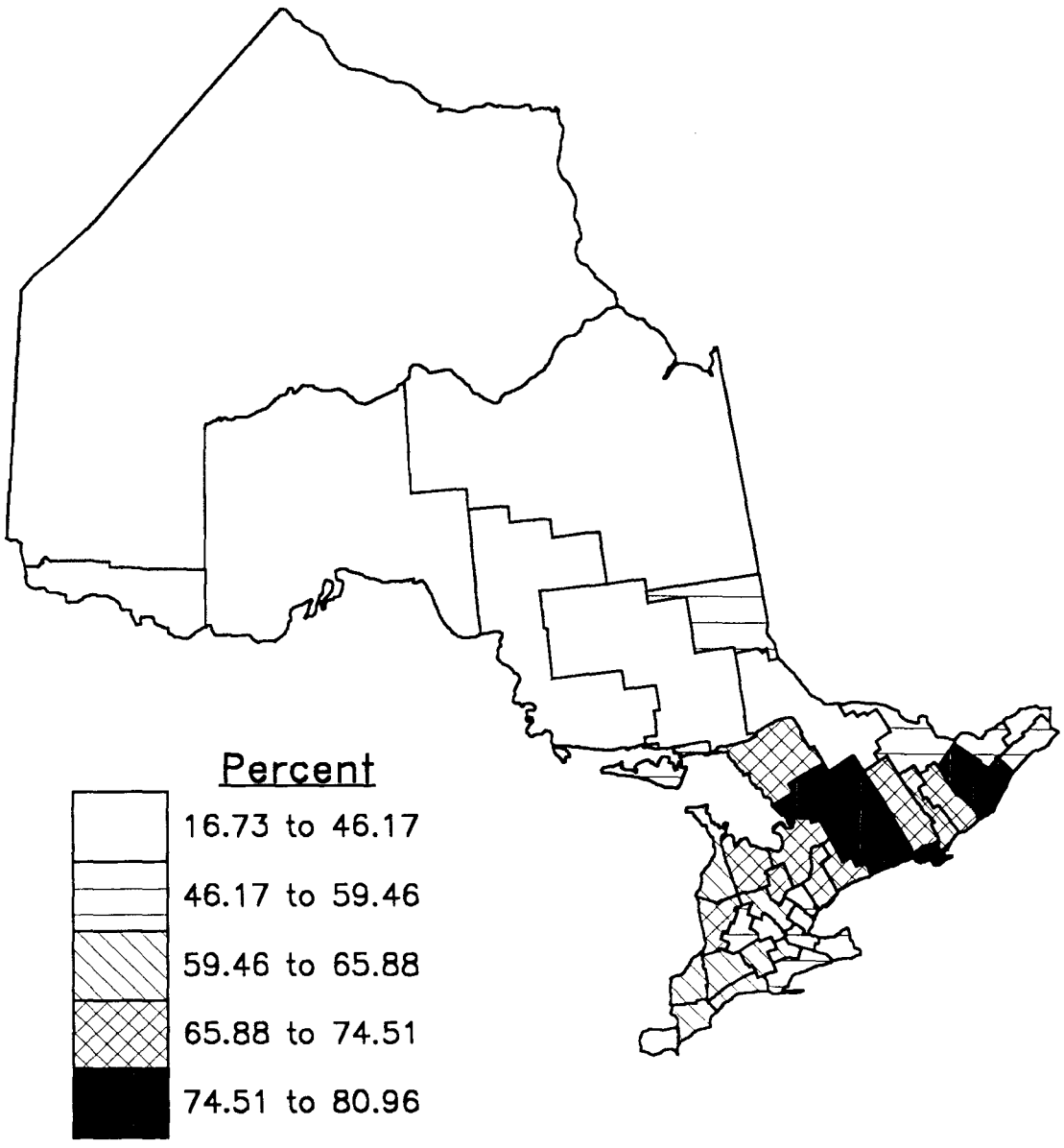
Percent Employed



Appendix 3.12

BRITISH 1981

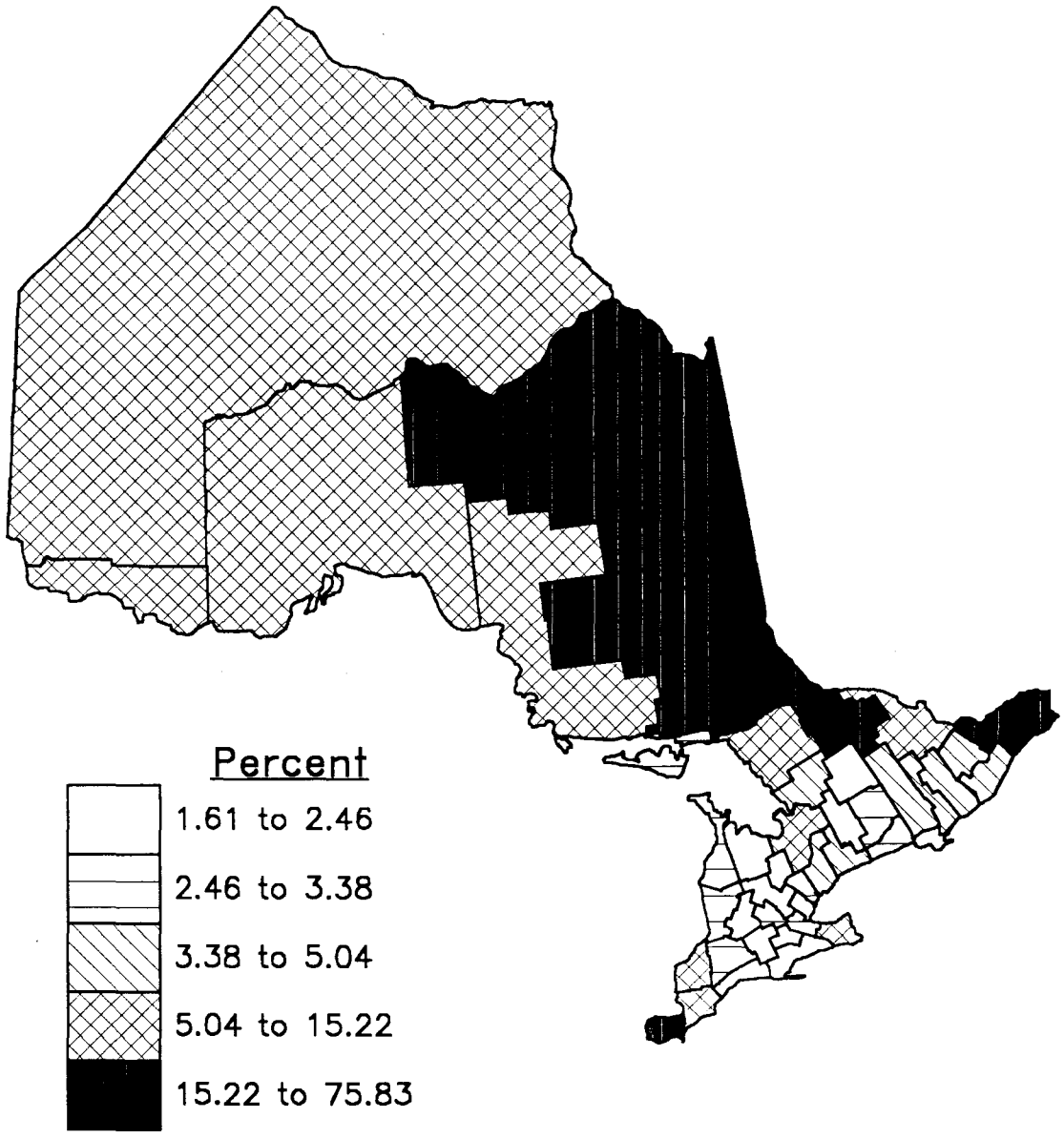
Percent of Population



Appendix 3.13

FRENCH 1981

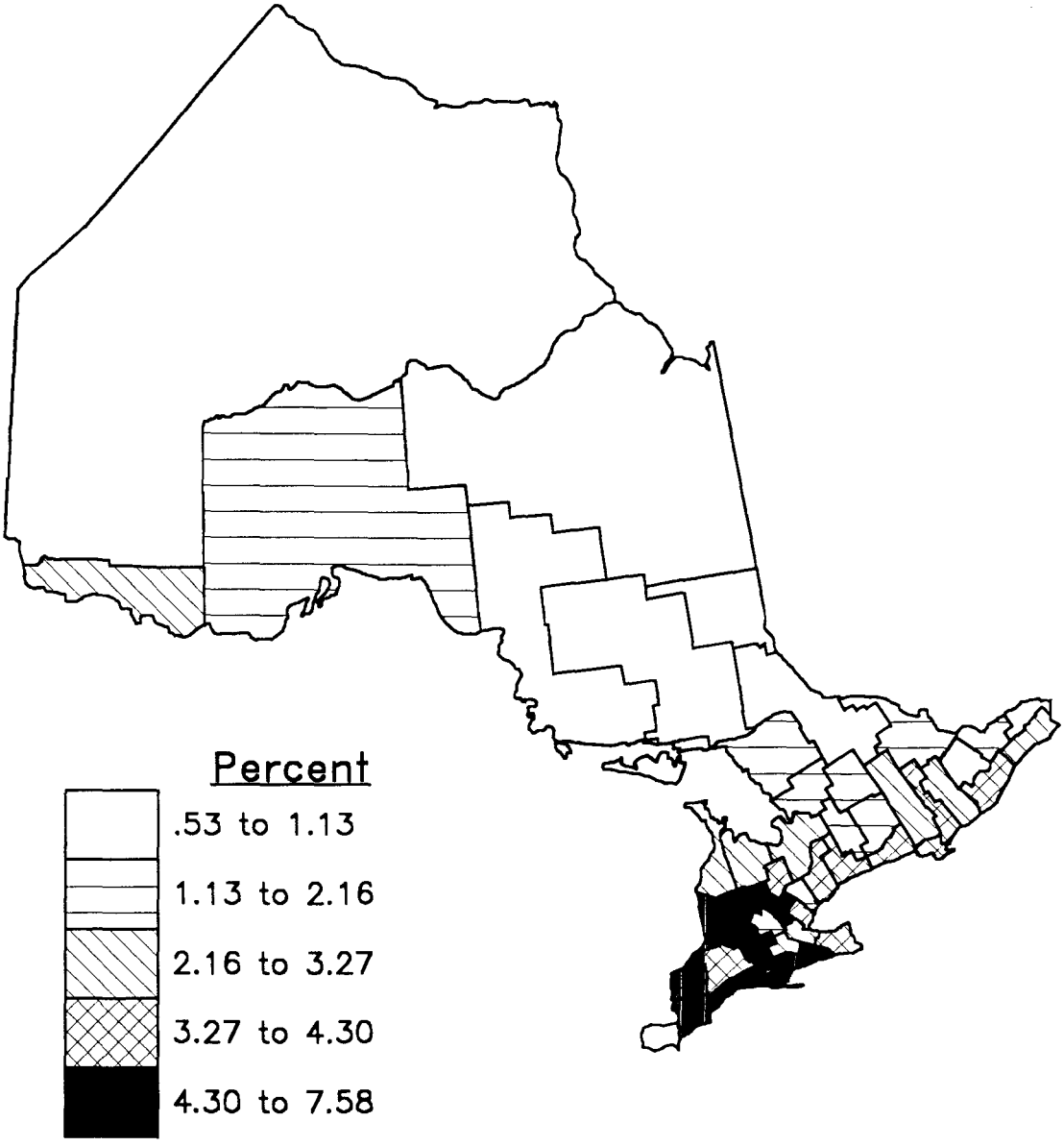
Percent of Population



Appendix 3.14

DUTCH 1981

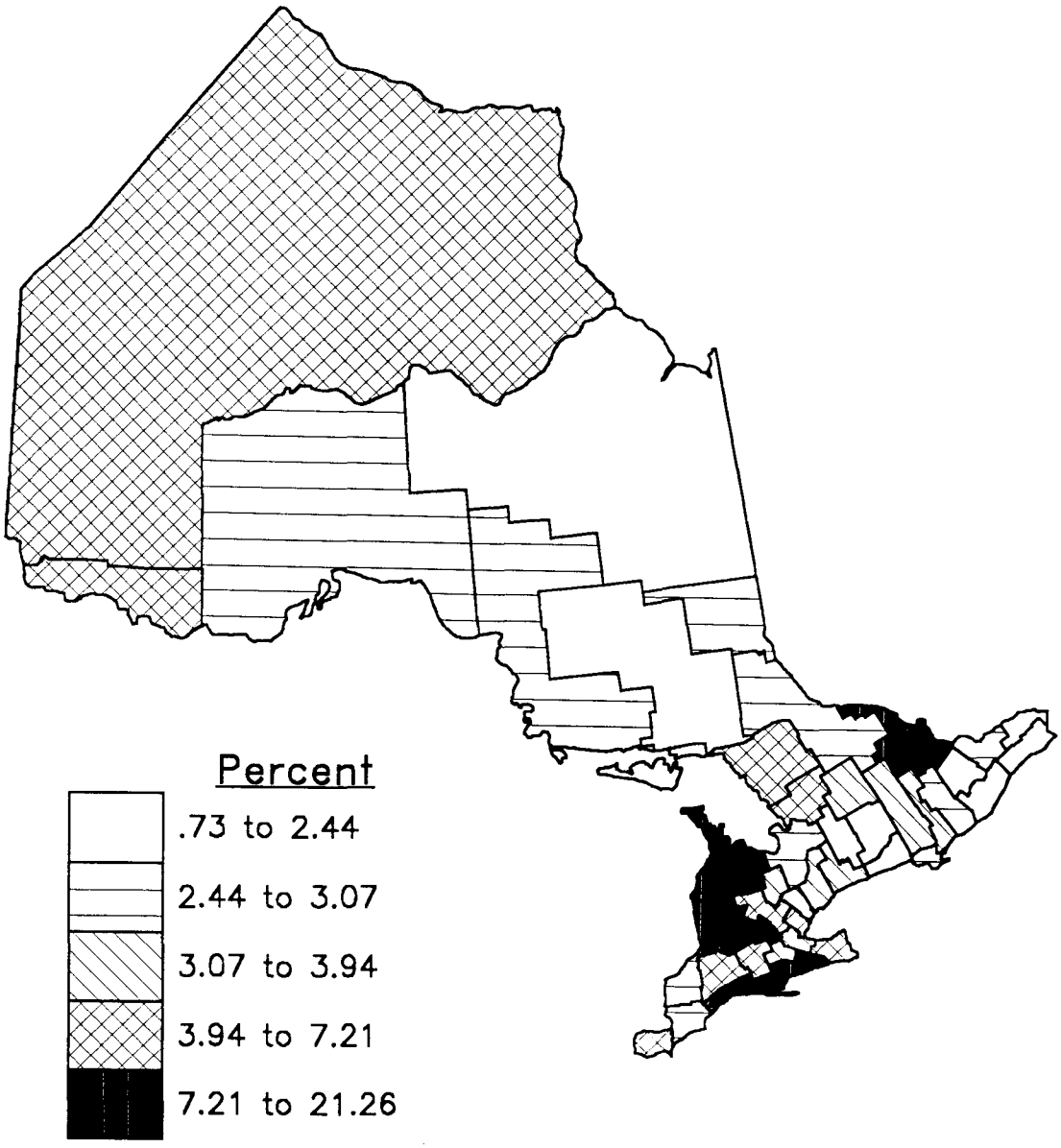
Percent of Population



Appendix 3.15

GERMAN 1981

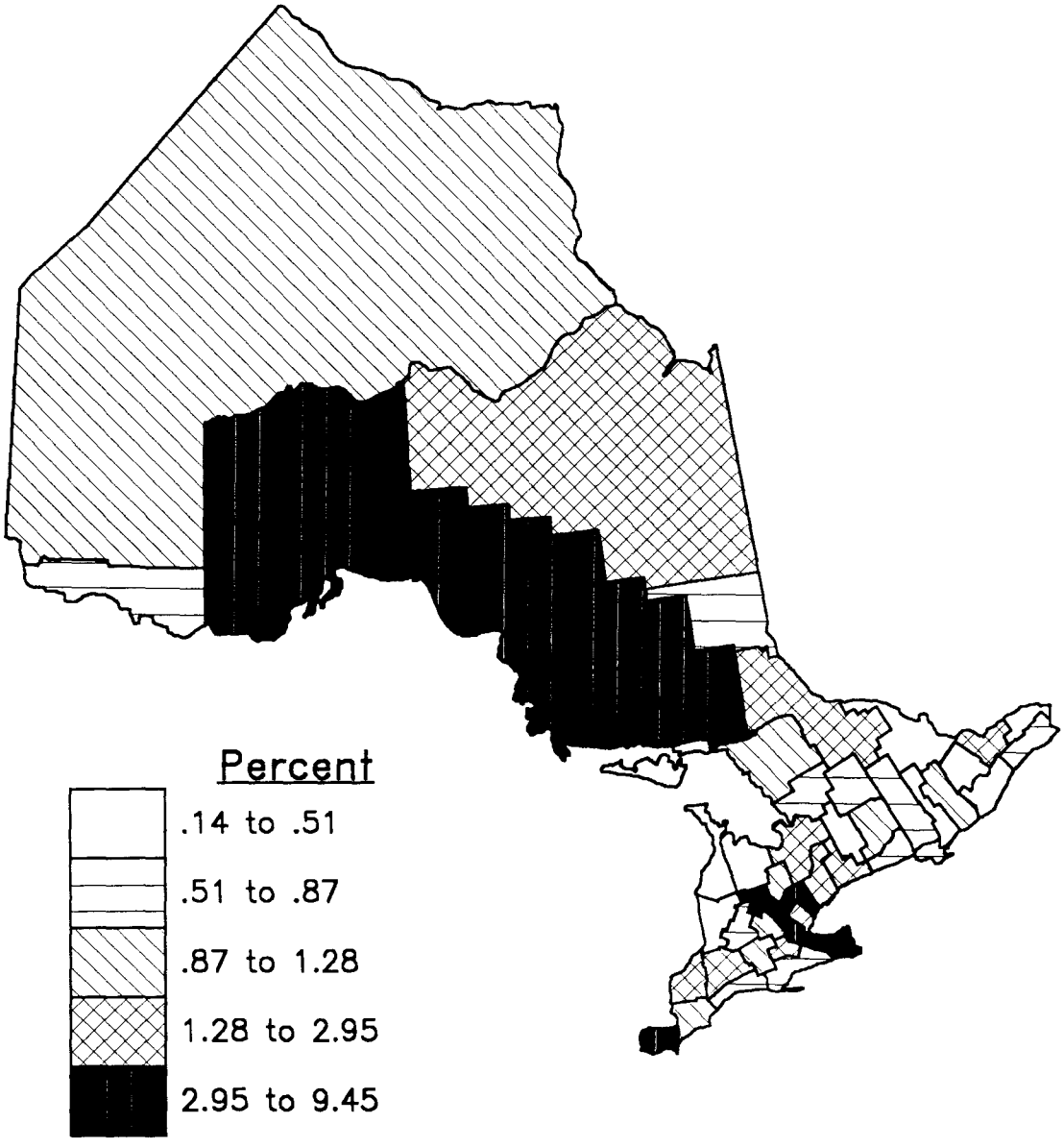
Percent of Population



Appendix 3.16

ITALIAN 1981

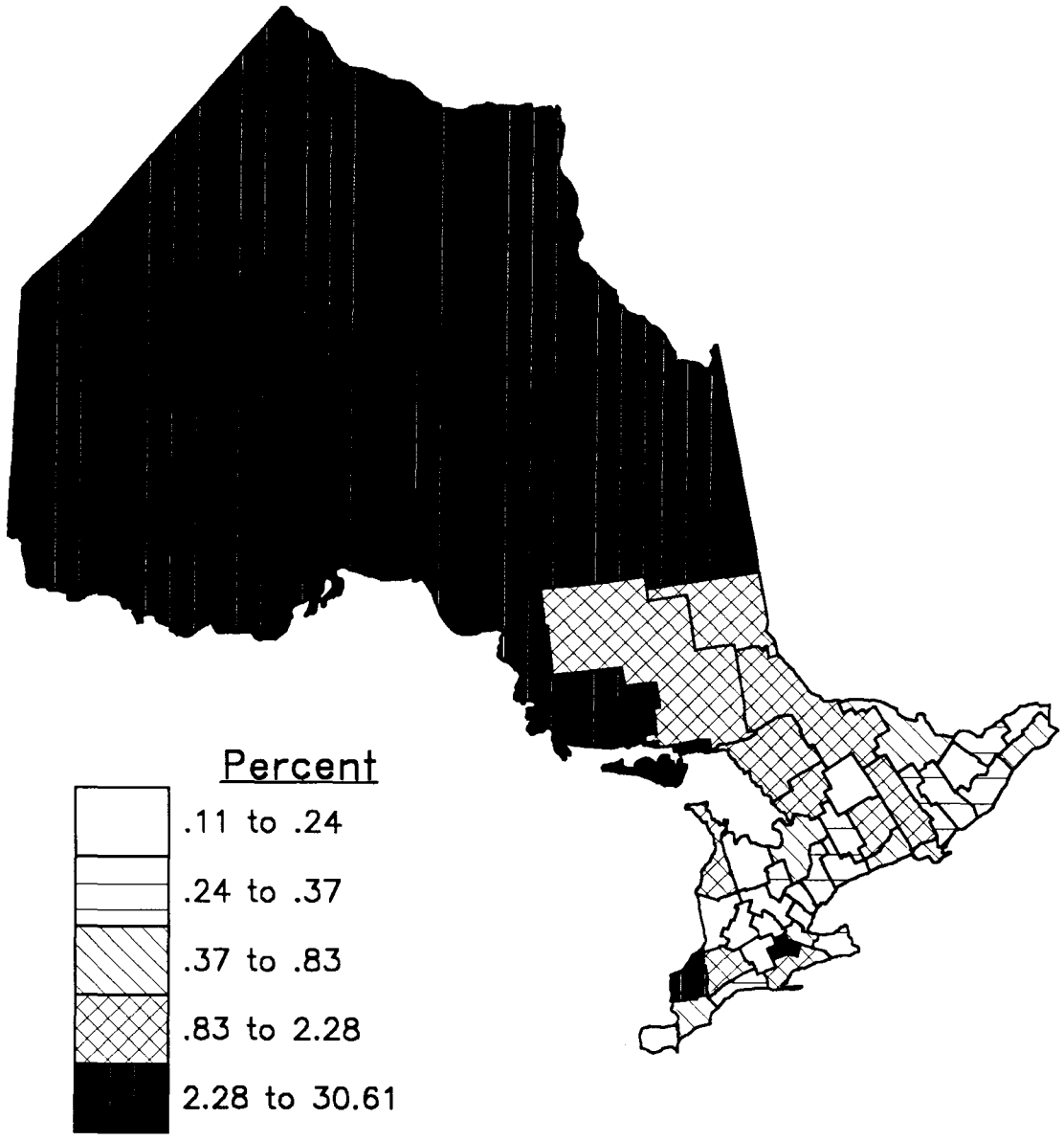
Percent of Population



Appendix 3.17

NATIVE 1981

Percent of Population



APPENDIX FOUR

Correlation Matrix - Dependent and Independent Variables

Appendix Four

Correlation Matrix

Dependent and Independent Variables

	01	02	03	04	05	06	07	08	09	10
01 Urban (5,000)	1.0	0.88	0.58	-0.53	-0.31	0.74	0.34	-0.42	-0.25	0.16
02 Urban (Census)	0.88	1.0	0.55	-0.54	-0.29	0.78	0.26	-0.50	-0.36	0.16
03 Population Density	0.58	0.55	1.0	-0.32	-0.03	0.60	0.53	-0.08	-0.28	0.22
04 Education < Grade 9	-0.53	-0.54	-0.32	1.0	0.53	-0.66	-0.26	0.27	0.30	-0.27
05 Incidence of Low Income	-0.31	-0.29	-0.03	0.53	1.0	-0.66	-0.14	0.12	-0.08	-0.23
06 Median Income	0.74	0.78	0.60	-0.66	-0.66	1.0	0.26	-0.39	-0.25	0.13
07 Manufacturing (1,000)	0.34	0.26	0.53	-0.26	-0.14	0.26	1.0	0.41	0.14	0.76
08 Manufacturing (1,000 Urban)	-0.42	-0.50	-0.08	0.27	0.12	-0.39	0.41	1.0	0.42	0.27
09 Primary Employment	-0.25	-0.36	-0.28	0.30	-0.08	-0.25	0.14	0.42	1.0	0.17
10 Secondary Employment	0.16	0.16	0.22	-0.27	-0.23	0.13	0.76	0.27	0.17	1.0
11 Tertiary Employment	0.49	0.51	0.44	-0.60	-0.12	0.48	-0.01	-0.35	-0.56	-0.09
12 British	-0.23	-0.34	0.04	-0.37	-0.01	-0.23	0.33	0.38	0.11	0.34
13 French	-0.05	0.07	-0.15	0.36	0.24	-0.01	-0.34	-0.24	-0.10	-0.29
14 Dutch	-0.04	-0.03	0.07	-0.22	-0.23	0.07	0.44	0.31	0.61	0.41
15 German	0.06	0.04	-0.03	0.08	-0.11	-0.07	0.34	0.13	0.27	0.38
16 Italian	0.59	0.59	0.26	-0.26	-0.15	0.52	-0.04	-0.41	-0.33	-0.10
17 Native	-0.26	-0.27	-0.15	0.49	0.03	-0.21	-0.38	0.14	0.07	-0.40
SMR Males	-0.12	-0.18	-0.03	0.29	0.27	-0.25	-0.31	0.10	-0.20	-0.31
SMR Females	-0.16	-0.12	-0.02	0.01	-0.05	-0.07	0.10	0.46	0.15	0.03
SMR Total	-0.18	-0.22	-0.03	0.24	0.19	-0.26	-0.17	0.34	-0.06	-0.24

Correlation Matrix

Dependent and Independent Variables

	11	12	13	14	15	16	17	SMR (M)	SMR (F)	SMR (T)
01 Urban (5,000)	0.49	-0.23	-0.05	-0.04	0.06	0.59	-0.26	-0.12	-0.16	-0.18
02 Urban (Census)	0.51	-0.34	0.07	-0.03	0.04	0.59	-0.27	-0.18	-0.12	-0.22
03 Population Density	0.44	0.04	-0.15	0.07	-0.03	0.26	-0.15	-0.03	-0.02	-0.03
04 Education < Grade 9	-0.60	-0.37	0.36	-0.22	0.08	-0.26	0.49	0.29	0.01	0.24
05 Incidence of Low Income	-0.12	-0.01	0.24	-0.23	-0.11	-0.15	0.03	0.27	-0.05	0.19
06 Median Income	0.48	-0.23	-0.01	0.07	-0.07	0.52	-0.21	-0.25	-0.07	-0.26
07 Manufacturing (1,000)	-0.01	0.33	-0.34	0.44	0.34	-0.04	-0.38	-0.31	0.10	-0.17
08 Manufacturing (1,000 Urban)	-0.35	0.38	-0.24	0.31	0.13	-0.41	0.14	0.10	0.46	0.34
09 Primary Employment	-0.56	0.11	-0.10	0.61	0.27	-0.33	-0.07	-0.20	0.15	-0.06
10 Secondary Employment	-0.09	0.34	-0.29	0.41	0.38	-0.10	-0.40	-0.31	0.03	-0.24
11 Tertiary Employment	1.0	-0.03	0.01	-0.21	-0.27	0.22	-0.07	-0.05	-0.01	-0.05
12 British	-0.03	1.0	-0.74	0.43	-0.01	-0.37	-0.28	-0.07	0.08	0.01
13 French	0.01	-0.74	1.0	-0.43	-0.30	0.03	-0.02	-0.01	-0.02	-0.03
14 Dutch	-0.21	0.43	-0.43	1.0	0.22	-0.17	-0.32	-0.28	0.16	-0.08
15 German	-0.27	-0.01	-0.30	0.22	1.0	-0.14	-0.15	-0.18	-0.05	-0.17
16 Italian	0.22	-0.37	0.03	-0.17	-0.14	1.0	-0.09	-0.03	-0.26	-0.16
17 Native	-0.07	-0.28	-0.02	-0.32	-0.15	-0.09	1.0	0.39	0.01	0.31
SMR Males	-0.05	-0.07	-0.01	-0.28	-0.18	-0.03	0.39	1.0	-0.20	0.68
SMR Females	-0.01	0.08	-0.02	0.16	-0.05	-0.26	0.01	-0.20	1.0	0.56
SMR Total	-0.05	0.01	-0.03	-0.08	-0.17	-0.16	0.31	0.68	0.56	1.0

BIBLIOGRAPHY

Arundel S.E. and L.M. Kinnier-Wilson (1986) "Parental Occupations and Cancer: A Review of the Literature", **Journal of Epidemiology and Community Health.** 40, 30 - 36.

Birch, J.M. (1983) "Epidemiology of Paediatric Cancer", **Recent Results In Cancer Research.** 88, 1 - 10.

Blot, W.J., J.F. Fraumeni, T.J. Mason and R.N. Hoover (1979) "Developing Clues to Environmental Cancer: A Stepwise Approach With the Use of Cancer Mortality Data", **Environmental Health Perspectives.** 32, 53 - 58.

Blot, W.J. and J.F. Fraumeni (1982) "Geographic Epidemiology of Cancer in the United States". In: Schottenfeld, D. and J.F. Fraumeni, **Cancer Epidemiology and Prevention**. Philadelphia: W. B. Sanders Company.

Breslow, N.E. and B. Langholz. (1983) "Childhood Cancer Incidence: Geographical and Temporal Variations", **International Journal of Cancer**. 32, 703 - 716.

Bunin, G. (1987) "Racial Patterns of Childhood Brain Cancer by Histologic Type", **Journal of the National Cancer Institute**. 78(5), 875 - 880.

Clarke, E.A., N. Kreiger and L.D. Marrett. (1985) "Cancer Incidence, Mortality and Treatment in Ontario". In **Cancer In Ontario**, 1985. Toronto: The Ontario Cancer Treatment and Research Foundation, 1986: 189 - 209.

Cleek, R.K. (1979) "Cancers and the Environment: The Effect of Scale", **Social Science and Medicine.** 13D, 241 - 247.

Corbett, S. and B.J. O'Neill (1988) "A Cluster of Cases of Lymphoma in An Underground Colliery", **The Medical Journal of Austrailia.** 149(4), 178 - 185.

Craft, A.W., S. Openshaw and J.M. Birch (1985) "Childhood Cancer in the Northern Region, 1968-82: Incidence in Small Geographical Areas", **Journal of Epidemiology and Community Health**. 39, 53 - 57.

Devesa, S.S., E.S. Pollack and J.L. Young (1984) "Assessing the Validity of Observed Cancer Incidence Trends", **American Journal of Epidemiology.** 119(2), 274 - 291.

Engel, L.W. (1980) "Accuracy of Death Certification in an Autopsied Population with Specific Attention to Malignant Neoplasms and Vascular Diseases", **American Journal of Epidemiology.** 111, 99 - 112.

Fabia, J. and T.D. Thuy (1974) "Occupation of Father at Time of Birth of Children Dying of Malignant Diseases", **British Journal of Preventative Medicine.** 28, 98 - 100.

Foster, H.D. (1986) **Reducing Cancer Mortality: A Geographic Perspective.** Western Geographical Series, University of Victoria, B.C.

Gardner, M.J. (1973) "Using the Environment to Explain and Predict Mortality", **Journal of the Royal Statistical Society.** 136(3), 421 - 427.

Glick, B.J. (1982) "The Spatial Organization of Cancer Mortality", **Annals of the Association of American Geographers.** 74(4), 471 - 481.

Gold, E.B., M.D. Diener and M. Szklo (1982) "Parental Occupations and Cancer in Children", **Journal of Occupational Medicine.** 24(8), 578 - 584.

Goodchild, M.F. (1986) Spatial Autocorrelation. London: University of Western Ontario.

Greenberg, M.R., J. Caruana, B. Holcomb, G.n Greenberg, R. Parker, J. Louis and P. White (1980) "High Cancer Mortality Rates from Childhood Leukemia and Young Adult Hodgkin's Disease and Lymphoma in the New Jersey - New York - Philadelphia Metropolitan Corridor, 1950 to 1969", **Cancer Research.** 40, 439 - 443.

_____., P.W. Preuss and R. Anderson (1980) "Clues for Case Control Studies of Cancer in the Northeast Urban Corridor", **Social Science and Medicine.** 140, 37 - 43.

_____. (1983) **Urbanization and Cancer Mortality: The United States Experience 1950 - 1975.** New York: Oxford University Press.

_____. ed. (1987) **Public Health and the Environment.** New York: The Guilford Press.

Greenberg, R.S. and J.L. Shuster. (1985) "Epidemiology of Cancer in Children", **Epidemiologic Reviews.** 7, 22 - 48.

Greenland, S. and H. Morgenstern (1987) **Ecologic Bias, Confounding and Effect Modification.** Division of Epidemiology, U.C.L.A. School of Public Health, Los Angeles, California.

Grufferman, S., H.H. Wang, E.R. DeLong (1982) "Environmental Factors in the Etiology of Rhabdomyosarcoma in Childhood", **Journal of the National Cancer Institute.** 68, 107 - 113.

Hakulinen, T., T. Salonen and L. Teppo (1976) "Cancer In the Offspring of Fathers in Hydrocarbon-Related Occupations", **British Journal of Preventative Medicine.** 30, 138 - 140.

Hammond, G.D. (1986) "The Cure of Childhood Cancers", **Cancer.** 58, 407 - 413.

Hanson, M.R. and J.J. Mulvihill (1982) "Epidemiology of Cancer in the Young". In: Levine, A.S., **Cancer in the Young.** New York: Masson Publishing Inc.

Harris, R.E., F.E. Harrell, K.D. Patil and R.Al-Rashid (1987) "The Seasonal Risk of Pediatric/Juvenile Acute Lymphocytic Leukemia in the United States", **Journal of Chronic Diseases**. 40(10), 915 - 923.

Hemminki, K., I. Saloniemi, T. Salonen, T. Partanen and H. Vainio (1981) "Childhood Cancer and Parental Occupation in Finland", **Journal of Epidemiology and Community Health.** 35, 11 - 15.

Hicks, N., M. Zack, G.C. Caldwell, D.J. Fernbach and J.M. Falletta (1984) "Childhood Cancer and Occupational Exposure in Parents", **Cancer.** 53, 1637 - 1643.

Hoover, R., T. Mason and F. McKay (1975) "Cancer By County: New Resource for Etiologic Clues", **Science.** 189, 1005 - 1007.

Horner, R.D. and T.N. Chirkos (1987) "Survirorship Differences in Geographical Comparisons of Cancer Mortality: An Urban-Rural Analysis", **International Journal of Epidemiology**. 16(2), 184 - 189.

Johnson, C.C. and M.R. Spitz. (1989) "Childhood Nervous System Tumours: an Assessment of Risk Associated with Paternal Occupations Involving Use, Repair or Manufacture of Electrical and Electronic Equipment", **International Journal of Epidemiology**. 18(4), 756 - 761.

Jones, K. and G. Moon (1987) **Health, Disease and Society: A Critical Medical Geography.** London: Routledge and Kegan Paul Limited.

Kantor, A.F., C.M. McCrea, and J.W. Meigs (1979) "Occupations of Fathers of Patients With Wilm's Tumour", **Journal of Epidemiology and Community Health.** 33, 253 - 256.

Kleinbaum, D.G., L.L. Kupper and H. Morgenstern (1982) **Epidemiologic Research: Principles and Quantitative Methods.** California: Wadsworth, Inc.

Kramer, S., A.T. Meadows, P.Jarrett and A.E. Evans (1983) "Incidence of Childhood Cancer: Experience of a Decade in a Population-Based Registry", **Journal of the National Cancer Institute**. 70, 49 -54.

Kwa, S.L.and L.J. Fine (1980) "The Association Between Parental Occupation and Childhood Malignancy", **Journal of Occupational Medicine**. 22(12), 792 - 794.

Langbein, L.I. and A.J. Lichtman (1987) **Ecological Inference.** Sage Publications, 10, 1 - 70.

Lenarsky, C. (1983) "Etiology of Childhood Malignancy". In: Lanzkowsky, P. ed. **Pediatric Oncology: A Treatise for the Clinician.** New York: McGraw-Hill Book Company.

Li, F P. (1982) "Cancers in Children." In: Schottenfeld, D. and J.F. Fraumeni, **Cancer Epidemiology and Prevention**. Philadelphia: W.B. Sanders Company.

MacMahon, B. and T.F. Pugh (1970) **Epidemiology: Principles and Methods.** Boston: Little, Brown and Company.

Mahoney, M.C., A.M. Michalek, M. Cummings, P.C. Nasca and L.J. Emrich (1989) "Cancer Surveillance in a Northeastern Native American Population", **Cancer**. 64, 191 - 195.

Marrett, L.D., E.A. Clarke, J. Hatcher and H.K. Weir (1986) "Epidemiologic Research Using the Ontario Cancer Registry", **Canadian Journal of Public Health.** 77(1), 79 - 85.

Mathews, J.D. (1988) "What can we learn from the clustering of cases of cancer?", **The Medical Journal of Australia**. 149(4), 169 - 171.

Mausner, J.S. and A.K. Bahn (1974) **Epidemiology: An Introductory Text.** Philadelphia: W.B. Saunders.

Miller, R.W. (1967) "Epidemiology of Childhood Neoplasia", **Neoplasia in Childhood**. Chicago: Year Book Medical Publishers, Inc.

_____. (1975) "Aetiology and Epidemiology". In: Bollm, H.J.G., J. Lemerle, M.K. Neidhardt and P.A. Voute, **Cancer in Children.: Clinical Management.** New York: Springer Verlag.

_____. (1978) "Environmental Causes of Cancer in Childhood", **Advances in Pediatrics.** 25, 97 - 119.

_____. and F.W. McKay (1984) "Decline in US Childhood Cancer Mortality 1950 through 1980", **Journal of the American Medical Association.** 251(12),1567 - 1570.

_____. (1988) "Geographical and Ethnic Differences in the Occurrence of Childhood Cancer", **International Incidence of Childhood Cancer**. Lyon: International Agency for Research on Cancer Scientific Publications. 87, 3 - 7.

Morgenstern, H. (1982) "Uses of Ecologic Analysis in Epidemiologic Research", **American Journal of Public Health.** 72(12), 1336 - 1343.

Mulvihill, J.J. (1982) "Ecogenetic Origins of Cancer in the Young: Environmental and Genetic Determinants". In: Levin, A.S. **Cancer In the Young.** New York: Masson Publishing Inc.

Neglia, J.P. and L.L. Robison. (1988) "Epidemiology of the Childhood Acute Leukemias", **Pediatric Clinics of North America**. 35(4), 675 - 692.

Neutel, C.I. (1970) **A Study of the Epidemiology of Childhood Cancer With Special Emphasis on Smoking During Pregnancy**. Phd Thesis, Faculty of Graduate Studies, The University of Western Ontario, London.

Newell, G.R., M.R. Spitz and J.G. Sider (1989) "Cancer and Age", **Seminars in Oncology.** 16(1), 3 - 9.

Olshan, A.F., N.E. Breslow, J.R. Daling, N.S. Weiss and A. Leviton (1986) "Childhood Brain Tumors and Paternal Occupation in the Aerospace Industry", **Journal of the National Cancer Institute.** 77(1), 17 - 19.

Percy, C., E. Stanck and L. Gloeckler (1981) "Accuracy of Cancer Death Certificates and its Effect on Cancer Mortality Statistics" **American Journal of Public Health.** 71, 242 - 250.

Peters, J.M., S. Preston-Martin and M.C. Yu (1981) "Brain Tumors in Children and Occupational Exposure of Parents", **Science.** 213, 235 - 236.

Polissar, L. (1980) "The Effect of Migration on Comparison of Disease Rates in Geographic Studies in the US", **American Journal of Epidemiology.** 111, 175 - 182.

Pratt, C.B. (1985) "Some Aspects of Childhood Cancer Epidemiology", **Pediatric Clinics of North America.** 32(3), 541 - 556.

Reynolds, D. (1989) **The Geographic Distribution of Cancer Mortality in Ontario: An Ecologic Analysis of Ethnicity, Industry and Socioeconomic Characteristics, 1976-1985.** M.Sc. Thesis, Department of Preventative Medicine and Biostatistics, University of Toronto.

Sanders, B.M., G.C. White and G.J. Draper (1981) "Occupations of Fathers of Children Dying From Neoplasms", **Journal of Epidemiology and Community Health.** 35, 245 - 250.

Savitz, D.A. and C.K. Redmond (1985) "Screening For Geographic Heterogeneity of Disease Rates: Application to Cancer Incidence in Allegheny County, Pennsylvania, 1969-71", **Journal of Chronic Diseases.** 38, 145 - 156.

_____ and D.L. Zuckerman. (1987) "Childhood Cancer in the Denver Metropolitan Area 1976 - 1983", **Cancer.** 59(8), 1539 - 1542.

Schneider, D., M.R. Greenberg and B. Stach (1990) "Pediatric Cancer Mortality Rates in New Jersey and the United States", **New Jersey Medicine.** 87 (9), 703 - 711.

Shaw, G., R. Lavey, R. Jackson and D. Austin (1984) "Association of Childhood Leukemia with Maternal Age, Birth Order and Paternal Occupation", **American Journal of Epidemiology.** 119, 788 - 795.

Spengler, R.F. (1981) "Water Contamination by Toxic Chemicals a Challenge to Cancer Registries in Assessing Population Risks", **Focus Great Lakes Water Quality.** 7(2), 1 5.

Spitz, M.R., J.G. Sider, C.C. Johnson, J.J. Butler, E.S. Pollack and G.R. Newell. (1986) "Ethnic Patterns of Hodgkin's Disease Incidence Among Children and Adolescents in the United States, 1973 - 92," **Journal of the National Cancer Institute**. 76(1), 235 - 239.

Stavraky, K.M. (1976) "The Role of Ecologic Analysis in Studies of the Etiology of Disease: A Discussion with Reference to Large Bowel Cancer", **Journal of Chronic Diseases.** 29, 435 - 444.

Statistics Canada, 1981 **Census of Canada.**

Swerdlow, A. J. (1986) "Cancer Registration in England and Wales: Some Aspects Relevent to Interpretation of the Data", **Journal of the Royal Statistical Society.** 149(2), 146 - 160.

Van Steensel-Moll, H.A., H.A. Valkenburg and G.E. Van Zanen (1985) "Childhood Leukemia and Parental Occupation", **American Journal of Epidemiology.** 121, 216 - 224.

Walter, S.D. (1988) **Ecologic Correlation Studies.** A Report to the Committee on the Assessment of the Human Health Effects of Great Lakes Water Quality, International Joint Commission.

_____. and S.E. Birnie (1991) "Mapping Mortality and Morbidity Patterns: An International Comparison", **International Journal of Epidemiology.** In Press.

_____. L.D. Marrett, S. M. Taylor, J.J. Drake, J.A. Davies and M.V. Hayes (1991) **Geographic Variation in Ontario Cancer Rates.** Report to the Ontario Ministry of Health.

Waterhouse J. ed. (1982) **Cancer Incidence in Five Continents, Volume IV.** Geneva: World Health Organization.

Wellington, D.G., E.J. Macdonald and P.F. Wolf (1979). **Cancer Mortality: Environmental and Ethnic Factors**. New York: Academic Press.

West, R. (1984) "Childhood Cancer Mortality: International Comparisons 1955 - 74", **World Health Statistics Quarterly**. 37, 98 - 127.

Wilkins J.R. and T.H. Sinks (1984) "Paternal Occupation and Wilm's Tumour in Offspring", **Journal of Epidemiology and Community Health.** 38, 7 - 11.

_____. and R.A. Koutras (1988) "Paternal Occupation and Brain Cancer in Offspring: A Mortality-Based Case-Control Study", **American Journal of Industrial Medicine.** 14(3), 229 - 318.

Young, J.L., L.G. Ries, E.Silverberg, J.W. Horm and R.W. Miller. (1986) "Cancer Incidence, Survival, and Mortality for Children Younger than Age 15 Years", **Cancer**. 58(2), 598 - 602.

Zack, M., S. Cannon and D. Lloyd (1980) "Cancer in Children of Parents Exposed to Hydrocarbon-Related Industries and Occupations", **American Journal of Epidemiology.** 111, 329 - 336.