AN ECOLOGIC STUDY OF CHILDHOOD LEUKAEMIA AND POPULATION MIXING

by

ANITA KOUSHIK

A thesis submitted to the Department of Community Health and Epidemiology
in conformity with the requirements for the degree of Master of Science

Queen's University
Kingston, Ontario, Canada
January 1999

ABSTRACT

Speculations on a possible infectious etiology of childhood leukaemia have long been existent. The 'Kinlen Hypothesis' postulates that population mixing in geographically isolated (rural) areas is conducive to the spread of infectious agents, and on rare occasions leukaemia occurs as a response to exposure to the relevant infectious agent. Population mixing in rural areas has been found to be associated with the occurrence of childhood leukaemia in several ecologic studies conducted in the United Kingdom. In particular, this relationship has been most consistent for the acute lymphocytic leukaemia (ALL) subtype in children aged 0 to 4. In order to explore this hypothesis in a new setting, an ecologic study using data from the Ontario Cancer Registry and the Census of Canada, was conducted. The association between population mixing (measured by the percentage change in population and the proportion of the population that are migrants), and childhood leukaemia incidence in Ontario was examined according to levels of geographic isolation (represented by urban-rural status, distance from an urban centre and number of paved highways). Analyses looked at both childhood leukaemia as a whole, as well as the ALL subtype in particular, and accounted for the potential confounding effect of socioeconomic status. Net increases in population in rural areas was found to be significantly associated with an increased incidence of childhood leukaemia, particularly for the ALL subtype in children aged 0 to 4 (adjusted RR=1.98, 95% CI 1.17-3.33 for greater than 20% population change, relative to no increase in population, in rural areas). A similar relationship was not observed in children aged 5 to 9 and 10 to 14. Also, there was no observed association in any age group according to the other two indicators of geographic isolation. Results from this study, in conjunction with results from other similar ecologic studies, are supportive of a role for an infectious agent in the etiology of childhood leukaemia.

ACKNOWLEDGEMENTS

I would like to thank my primary supervisor, Dr. Will King, for his exceptional guidance and sincere interest in this thesis project. I am also grateful to Dr. John McLaughlin for contributing his experience and expertise to this project.

I acknowledge the Ontario Cancer Registry as the source of leukaemia data in this study and I thank Agnes Moskowitz for facilitating data procurement. I am grateful to Janet Honig, Dr. Rowland Tinline, David Ball and Tara Vinodrai at the GIS Lab at Queen's for consultation and assistance with maps and distance data. I thank Jeff Moon at the Documents Library and Shirley Harmer at the Maps and Air Photo Collection at Queen's for their assistance in obtaining Statistics Canada data and road map information.

I gratefully acknowledge the financial support provided by the Ontario Graduate Scholarship Program.

I would like to thank faculty, staff and students at the Department of Community Health and Epidemiology for their role in making this program interesting and enjoyable. Special thanks to my classmates for the fun, the support and the sharing of your diverse research interests which have enhanced my academic experience.

Last, but not least, I am thankful to my family and friends, near and afar, whose support and encouragement I could have not done without.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ALL	Acute lymphocytic (or lymphoblastic) leukaemia
AML	Acute myeloid leukaemia
CA	Census agglomeration
CI	Confidence interval
CLL	Chronic lymphocytic leukaemia
CMA	Census metropolitan area
CML	Chronic myeloid leukaemia
CSD	Census subdivision
EA	Enumeration area
EMF	Electromagnetic field
HTLV-I	Human T-cell lymphotrophic virus type
ICD-9	International Classification of Diseases, ninth revision
MOH	Ministry of Health
OCR	Ontario Cancer Registry
RR	Rate ratio
SAS	Statistical Analysis System
SES	Socioeconomic status
SGC	Standard Geographical Classification
SMR	Standardized mortality ratio
UK	United Kingdom

1. INTRODUCTION

Leukaemia refers to a "heterogeneous group of neoplasms of the white blood cells,"[1] and is thought to arise in the bone marrow.[2] The four main subtypes, classified by cell lineage (lymphoid or myeloid) and maturity of the proliferating cells (acute or chronic),[3] include: acute lymphocytic leukaemia (ALL), acute myeloid leukaemia (AML), chronic lymphocytic leukaemia (CLL), and chronic myeloid leukaemia.[4] Each subtype differs in prognosis and age-at-incidence patterns.[1,4] In Canada, leukaemia, and specifically the ALL subtype, is the most common cancer in children aged 0 to 14.[5,6]

Recent advances in molecular biology and immunology have led to a better understanding of the steps in the leukaemogenic process.[7] As well, advances in treatment protocols and the addition of aggressive supportive care have significantly improved survival.[2] The probability of survival in some subgroups of childhood ALL is nearly 90 percent.[8] Despite this better understanding of childhood leukaemia, the etiology remains largely unknown. Certain agents (ionizing radiation, chemotherapeutic agents and inherited conditions) have been implicated as established causal factors for leukaemia in children; however, these factors explain only a very small proportion of childhood leukaemia diagnoses.[9] Thus, further research is required in order to identify other relevant factors associated with childhood leukaemia and, therefore, contribute to a better understanding of the etiology of this malignancy.

Both environmental and infectious etiologies have been suspected for childhood leukaemia.[10,11] Speculations on a viral component to the etiology have been existent since the early 1900s.[12] These speculations are supported by reported clusters, as well as on the identification of leukaemia viruses in animal species and in one form of adult leukaemia.[12] One particular hypothesis pointing to a viral causation, proposed by Leo Kinlen, focuses on the postulate that childhood leukaemia occurs as an aberrant response to a common and widespread infection.[13]

According to the Kinlen hypothesis, people living in geographically isolated areas are not exposed to common infectious agents at usual ages, as compared to individuals living in urban centres. Immunity is the direct result of infection (through natural immunization processes).[14] Thus, as a result of a lack of (or delay in) exposure, a large proportion of individuals living in a geographically isolated community are susceptible to infection by these agents. In other words, the herd immunity to common infectious agents can be expected to be low in unexposed rural communities. A large number of susceptibles and a high rate of contact between susceptible and infected individuals is conducive to transmission of an infectious agent and consequently an epidemic. This condition can arise when there is a large population influx into a geographically isolated rural community. Thus, when susceptible individuals are exposed to infected people (or carriers) through population mixing, there is a resultant epidemic of the infection, and childhood leukaemia can occur as a rare response to this infection.

Using data from the Ontario Cancer Registry and from the Census of Canada, the Kinlen hypothesis focusing on the association between population mixing and childhood leukaemia incidence in Ontario was examined for the years 1978 to 1992. This relationship was examined for various levels of geographic isolation, while accounting for the potential confounding effect of socioeconomic status. An ecologic study design was employed, with the unit of analysis being the census subdivision in a 5-year period. Analyses looked at both childhood leukaemia as a whole, as well as the ALL subtype in particular.

The following chapter summarizes current knowledge regarding the etiology of childhood leukaemia, and describes evidence pertaining to a viral etiology. A description of past studies specifically examining population mixing is included. In chapter 3, the objectives are presented and the study methodology is described in detail. Issues with data quality of both the Census of Canada and the Ontario Cancer Registry are discussed in chapter 4. The results of the analysis are presented in chapter 5, followed by a discussion of the findings in chapter 6.

2. BACKGROUND

2.1 Descriptive Epidemiology

The most recent data for Canada indicates that leukaemia accounts for approximately one-third of all cancers in children aged 0 to 14, with an age-standardized incidence rate of 48.10 per 1,000,000 in 1993.[6] Approximately 80 percent of new cases are of the ALL subtype, with the next most frequent being AML, accounting for about 12 percent of new cases.[6] In Western countries, CML accounts for one to 3 percent of new cases while CLL is extremely rare under the age of 30.[7] Approximately 3 percent of all childhood leukaemias occur in infants (less than one year of age), with roughly equal proportions of ALL and AML occurring.[15]

Features of childhood leukaemia include a slightly increased frequency of incidence in males than in females[16] and, common to both sexes, a distinct peak in incidence between the ages of two and three.[17] The incidence peak is attributable to the ALL subtype[16], and specifically, to B-cell precursor ALL, also referred to as 'common ALL'.[18] Distinct from ALL, the peak in incidence for childhood AML occurs in the first year of life, decreases through age 4, and finally stays constant throughout the remaining childhood years.[16]

2.2 Established Risk Factors

Although the majority of childhood leukaemia cases are unexplained, environment and heredity are related to the etiology in a small proportion. Specifically, ionizing radiation, chemotherapeutic agents, and inherited conditions have been established as causal factors in a minority of cases.

2.2.1 Ionizing Radiation

Although not entirely understood, the best substantiated biological mechanism associated with leukaemia is through ionizing radiation.[19] Molecular and cytogenetic research have clearly shown that

many leukaemias and lymphomas are associated with chromosomal rearrangements that are consistent with radiation-induced single-strand DNA breaks.[20] The majority of studies supporting a leukaemogenic role for ionizing radiation comes from studies of environmentally and occupationally exposed groups such as Japanese atomic bomb survivors, individuals exposed to fallout from nuclear weapons testings, people undergoing radiation therapy, prenatal exposure to diagnostic x-rays, medical radiation workers and nuclear industry workers.[20-22] Research on these populations reveal an increased risk of leukaemia with radiation exposure. In particular, ionizing radiation is most closely associated with CML, AML and ALL[20], while CLL shows no association.[22]

2.2.2. Chemotherapeutic Agents

Secondary cancers include two or more independent cancers in the same patient, which are not metastases or recurrences of the original primary cancer.[23] Secondary leukaemias have been shown to occur in response to chemotherapy (in addition to radiotherapy).[19] Similar to the case with radiation, biological mechanisms involve DNA strand breaks.[24] The most common leukaemia subtype to occur following chemotherapy is AML.[25] In children, therapy-related leukemia is rare[25], however, children being treated for ALL, lymphoma or Hodgkin's disease are at increased risk of AML.[19]

2.2.3. Inherited Conditions

An increased risk of leukaemia in children with conditions involving constitutional chromosomal abnormalities lends support to a role for inherited factors in the etiology of childhood leukaemia.[27] The most striking observation is the association between trisomy 21 (or Down's syndrome) and acute leukaemia (both ALL and AML), where compared to the general population, the risk is 10- to 20-fold increased in children with this condition.[16] Fanconi's anemia, ataxia-telangiectasia and Bloom's syndrome are inherited genetic instability syndromes that are associated

with an increased risk of leukaemia.[28] Other inherited conditions that increase the susceptibility to childhood leukaemia are Klinefelter's syndrome, trisomy G syndrome, neurofibromatosis, and Shwachman's syndrome.[27] Further support for a genetic role in the etiology of childhood leukaemia is provided by evidence of leukaemia clustering in families with excess cancer incidence, and by the high degree of concordance amongst monozygotic twins for leukaemia.[25] However, this evidence additionally supports the role of an environmental factor to which family members are commonly exposed.

2.3. Suspected Risk Factors

Given the known role of relatively high doses of ionizing radiation, it is thought that background (natural) ionizing radiation, including external gamma radiation and internal alpha particles, may play a role in childhood leukaemia incidence.[29] In one ecologic study on the effect of background radiation, it was found that outdoor gamma radiation was significantly associated with childhood leukaemia; however, in a more recent study it was found that once radon was accounted for, the effect of both indoor and outdoor gamma radiation was non-significant.[30] In addition to ionizing radiation, non-ionizing electromagnetic field (EMF) radiation has been suspected of increasing the risk of leukaemia in children. The role of EMF in the etiology of childhood leukaemia remains controversial,[22,31] however, this is likely due to problems associated with exposure measurement.[16] Research focusing on parental exposure to chemicals such as pesticides[32] and benzene[25] has provided evidence supporting an association with childhood leukaemia, and in particular with AML, however, the results have been contradictory.[16] Maternal smoking and use of alcohol during pregnancy has been shown to be associated with childhood leukaemia, however, the evidence is inconsistent.[25,33,34] Recently it has been suggested that paternal smoking prior to conception may be a relevant risk factor.[35] Maternal factors that have been reported to be associated with an increased risk of ALL include

advanced maternal age and a history of fetal loss.[36] High birth weight has also been found to be associated with childhood leukaemia, although in a few studies this relationship was not observed.[37] A recent study showing an association between childhood leukaemia and anemia during pregnancy lends support to the suspected relationship between childhood leukaemia and birth weight since anemia during pregnancy is associated with an increased placental weight.[38] Pet ownership was implicated as a risk factor in this same study, although, as the authors state, this result could have been due to chance.[38] Because of the importance of dietary factors in relation to adult cancers, it has been suggested that child and/or maternal diet may be etiologically relevant, however, to date very few studies have investigated this factor.[25] Another environmental factor which has received considerable attention in research is infectious disease.[29]

2.4. Potential Viral Etiology

Speculations on a possible viral component to childhood leukaemia have been existent since the early 1900s.[12] One of the most striking findings suggesting a possible infectious etiology was reported in 1963.[39] A cluster of eight childhood leukaemia cases was observed during a span of four years (1957-1960) in the town of Niles, Illinois following a period of rapid population growth. Common to the eight cases were: a particular school, where 7 of the cases either attended or had a sibling attending; the time pattern of their onset, which corresponded with the 1957-58 and 1959-60 school years; and the coinciding appearance of a rheumatic-like illness amongst other children during the same school year periods. These observations, as well as the fact that space and time clustering is a characteristic common to diseases caused by infectious agents,[17] suggested a possible microbiological (and perhaps viral) origin to childhood leukaemia.

Numerous cluster analyses, reviewed in 1993 by Alexander,[12] have examined childhood leukaemia with respect to space, time and both space and time. Results are inconsistent, however,

formal statistical analyses were not always performed, adjustments for multiple testing were not always undertaken, and the definitions for space and time have varied among studies. In any case, because of the relatively long latency period[12] and the possibility that leukaemia is a rare response to the relevant infection[40], the observed lack of clustering does not contradict an infectious etiology.

Some recent studies that assessed clustering of childhood leukaemia with respect to space and time[9,41] using formal statistical methods and similar space and time definitions have reported evidence of space-time interactions for childhood leukaemia. However, it has been suggested that the analysis of spatial clustering, which assumes that the relevant risk factors have existed in the study area for a considerable part of the study period,[42] is more appropriate than space-time clustering when studying infectious exposures with relatively long latent periods.[12] Spatial clustering was evident in studies carried out in Great Britain[43], Greece[42] and Hong Kong[18] while studies conducted in the United States[44] and Sweden[45] found no evidence of spatial clustering.

Further research which has addressed a potential viral etiology of childhood leukaemia has focused on such factors as socioeconomic status (SES), birth order and the influence of childhood contacts.[17,25] The underlying rationale, consistent with the Kinlen hypothesis, relates to timing of exposure to infectious agents.[25] Specifically, it is thought that these factors are associated with delayed exposure to common infectious agents[46] which can consequently result in childhood leukaemia upon later exposure.

A number of studies have shown a higher risk of childhood leukaemia associated with a high SES at both the ecologic[46-49] and individual level.[36,38,50] As well, studies have shown a decrease in ALL risk with increasing birth order.[17] It is thought that high SES is associated with a higher standard of hygiene, thereby resulting in a delay in exposure to common infections.[40] Similarly, delayed exposure to common infections is expected in children of earlier birth order[17] while early exposure is more likely with later birth order through older siblings.[37] Although several studies have supported a link between

SES and birth order with childhood leukaemia risk, others have not.[36,44]

Additional factors of interest, related to childhood contacts, include household crowding and day care attendance. A study conducted in Greece found that early attendance at nurseries for more than three months in the first two years of life provided significant protection from the risk of childhood leukaemia incidence.[50] This decreased risk of leukaemia was hypothesized to be due to a young age of exposure to relevant infectious agents. A more recent Greek study also found an inverse association between childhood leukaemia and day care attendance, as well as with household crowding, however, the results were statistically non-significant.[38]

The role of a possible viral agent in the etiology of childhood leukaemia is further supported by the known role of viruses in animal leukaemias, and additionally, in two human hematopoietic cancers.[19] Human T-cell lymphotropic virus type I (HTLV-I) is a retrovirus known to cause adult T-cell leukaemia and the human herpesvirus, Epstein-Barr virus, is associated with Burkitt's lymphoma.[51]

2.5. Population Mixing and Geographic Isolation

Further evidence supporting a role for viruses in the etiology of childhood leukaemia come from ecologic studies examining population mixing and geographic isolation. These studies improve upon descriptive cluster analyses by including an exposure variable, and therefore, an analytical component to the study. The original studies were conducted by Leo Kinlen in situations of extreme population mixing in isolated areas of the United Kingdom (UK). Similar studies have been conducted by other researchers, both in the UK as well as in other parts of the world.

2.5.1 Kinlen Studies

In Britain in the 1980s, clusters of childhood leukaemia were reported in two separate sites,

Sellafield and Dounreay, both located near nuclear reprocessing plants.[13] The increased incidence of childhood leukaemia in the residential towns near Sellafield and Dounreay led to suspicions that radiation from the nuclear plants were causing the cancers. Several independent studies, however, indicated that the levels of radiation in these areas were far below that necessary to account for the excess of leukaemia. In determining alternative factors that were common to these areas, it was noted that both sites were geographically isolated and underwent a large population increase upon the construction of the nuclear facilities. Thus, the hypothesis of a viral etiology, with respect to the relationship between population mixing in geographically isolated areas and childhood leukaemia, was postulated by Leo Kinlen.

To test this hypothesis, Kinlen identified a New Town created in Scotland which was relatively similar to the towns near Sellafield and Dounreay, with respect to geographic isolation and the experience of a large population influx, yet did not contain any source of nuclear radiation.[13] Glenrothes was designated as a New Town in 1948 and the population doubled between 1951 and 1967. It was found that in the age group 0 to 4 years particularly, a significant increased mortality rate due to childhood leukaemia occurred as the population of Glenrothes doubled. The results of this ecologic study thus provided initial evidence for the role of population mixing on the incidence of childhood leukaemia thereby pointing to a possible viral etiology.

A second study was conducted which investigated all British New Towns designated around the same time as Glenrothes (1946-1950).[52] These New Towns fell into the two categories of 'rural', which could be considered isolated, and 'overspill', which were towns near large cities like London and Glasgow and were, therefore, relatively not isolated. Two 20 year periods following the designation period were analyzed (1946-1965 and 1966-1985). It was found that in rural New Towns there was a significantly increased rate of leukaemia deaths in the 20 year period immediately following the designation of the new town, specifically in the 0 to 4 year age group. On the other hand,

in the second 20 year period in rural New Towns, as well as in both periods in the 'overspill' New Towns no excess was found.

Several subsequent ecologic studies, which have examined other circumstances of population mixing in Great Britain, have been conducted by the same group of investigators.[40] Situations and populations that have been investigated include: the highly mobile groups of military servicemen and construction workers, where population mixing in rural areas resulted in subsequent increased rates of childhood leukaemia; changes in commuting level, which showed a significant excess of leukaemia in children in the decile with the greatest change in commuting; and British wartime evacuations, where childhood leukaemia in rural areas with increasing proportions of evacuees showed a significantly positive trend. The consistent results from these studies thus provided further evidence for the role of population mixing on childhood leukaemia in the UK, particularly in the age group 0 to 4.

Kinlen's initial studies were conducted on localized and temporary situations of population mixing in the UK. In an attempt to investigate a population mixing situation that was more prolonged and covered a larger area, a study was conducted which examined childhood leukaemia mortality among several different countries (particularly rural).[53] The measure of population mixing used was internal rural to urban migration in the 1950s and 1960s. Among 33 countries studied, it was found that Greece and Italy had the highest level of rural to urban population movement while also having high childhood leukaemia mortality rates between 1958 and 1972. In fact, Greece had the highest childhood leukaemia mortality rate among all the countries studied.

Recently, five of Kinlen's original studies have been reanalyzed to test the hypothesis that among excesses of childhood leukaemia in areas with extreme population mixing, the incidence is higher for the children of men in occupations involving contact with many individuals.[54] For each study, cases in the categories of highest population mixing were identified and then categorized by paternal occupation contact level thus determining the observed number of cases by contact level.

Expected numbers were derived by determining the distribution of paternal occupations by contact level among cases in the categories of lowest population mixing. The data across the studies were pooled and a significant trend across low to very high exposure categories was found. Within the high contact group, the increased risk was found to be most associated with occupations involving transport and construction. Thus, in areas where population mixing is associated with childhood leukaemia, paternal occupations involving many contacts is associated with an excess risk.

2.5.2. Other UK Studies

Several other researchers have examined population mixing and/or geographic isolation at the ecologic level using Kinlen's hypothesis as a model. One such study used census data on electoral wards of England and Wales to investigate factors related to isolation.[46] The incidence of ALL from 1984 to 1988 was examined and it was found that greater distance to urban areas was significantly associated with increased risk, even when adjusted for urban-rural status and socioeconomic status. Another study hypothesized that the degree of population change in a community may be more relevant than the initial remoteness of the area.[55] Rather than restricting the analysis to remote areas, as in Kinlen's original studies, this study examined the percent change in population between 1961 and 1971 in all local authority areas of England and Wales. An excess risk of childhood leukaemia mortality between 1969 and 1973 was found in areas experiencing greater than 50% increase in population, relative to areas experiencing less than 50% increase.

In 1991, a report on the geographical and epidemiological features of childhood leukaemia and non-Hodgkin's lymphoma in Great Britain from 1966 to 1983 was published.[56] Among several analyses, population mixing, measured by the ratio of the total child population in 1981 to the total child population in 1971 and categorized into quintiles, was examined.[49] There was no increase in rates of leukaemia from 1966 to 1983 with increasing child population ratio, even when controlled for

SES and urban-rural status. However, when the analysis was restricted to rural areas and population mixing was analyzed as dichotomous, it was found that census tracts with a child population ratio of 140% or more had a higher risk of lymphocytic and unspecified leukaemia in children aged 0 to 14, relative to census tracts with less than 140% ratio, particularly in the highest SES groups.

A recently published study suggesting that population mixing, even at relatively low levels, may be important in the etiology of childhood leukaemia, analyzed the effect of the proportion of the total population, as well as the proportion of children under age 15, who had not been resident in that county in the previous year (i.e. total migrants and child migrants).[47] For both the 0 to 4 and 5 to 9 age groups, child migration was significantly associated with childhood leukaemia incidence. Total migration, on the other hand, was significantly associated with leukaemia incidence only in the 0 to 4 year age group. However, in the 5 to 9 year age group, total migration accounting for diversity of the incomers was significantly associated with ALL risk.

Overall, the studies conducted in the UK by Kinlen and by other researchers have shown that childhood leukaemia rates are associated with population mixing. In particular, this finding has been consistently seen in the 0 to 4 year age group.

2.5.3. Studies Outside of the UK

An ecologic analysis was conducted on data from the Aegean and Ionian Greek islands, which had been relatively isolated until an increase in tourism.[57] It was expected that population influxes due to tourism on the islands would result in an increase in childhood leukaemia. However, it was found that mortality rates between 1976 and 1989 for childhood leukaemia in these island groups were not significantly different from the other Greek regions. An ecologic study conducted in New Zealand investigated the effect of population increase and the establishment of new towns as a result of developments in the forestry industry.[58] Leukaemia incidence rates in children in these previously non-

populated areas were compared with rates in the rest of New Zealand. It was found that incidence rates in the study areas were not significantly increased during or after the greatest population increase, as would be expected according to the population mixing hypothesis. Thus, these studies, both conducted outside the UK, did not provide supportive evidence for the Kinlen's hypothesis.

A recent study examined clustering as well as the effect of population growth in new towns in the New Territories of Hong Kong, built in response to the expanding population and imbalanced population density of the 'Hong Kong protectorate'.[18] Analyses were conducted on total leukaemia, ALL and common ALL. In areas where extreme population mixing occurred an excess of ALL, and particularly common ALL, was found in young age groups, although, this result was not quite statistically significant.

2.6. Summary and Methodological Considerations

The most consistent finding in past studies is the excess of leukaemia in the 0 to 4 year age group, which is likely related to the incidence peak at ages 2 to 3. Because children in this age group would have had the least opportunity for early exposure to the relevant virus, it is likely that the proportion of susceptible children in this age group is larger than the proportion of susceptibles in the older age groups.[40] It can therefore be expected that the risk of leukaemia would be greater in this age group. It has also been observed that the excess risk in the 0 to 4 year age group associated with population mixing, and the childhood peak, is concentrated in the ALL subtype.[18,40] Accordingly, it has been suggested that the etiology of ALL in the early age group may differ from that for other ages.[46] Most studies have examined all childhood leukaemia subtypes as one entity, and some have even included non-Hodgkin's lymphoma.[59] Recent studies, however, suggest that the etiology of ALL may differ from that of other subtypes.[18,49] The study conducted in Hong Kong separately examined both childhood leukaemia (all subtypes) and ALL in relation to population mixing and in doing so

emphasized the importance of separate analyses of ALL in the childhood peak.[18]

Previous studies are for the most part consistent with the hypothesis of a viral etiology to childhood leukaemia. However, the majority that support the Kinlen hypothesis were ecologic analyses conducted within the UK. Only one of the three studies conducted outside the UK was supportive of this hypothesis and the results were not statistically significant.[18] The use of mortality data is of concern in some of the past studies, particularly with respect to cases that have occurred recently. Because treatment modalities have improved survival following diagnosis, mortality cannot approximate incidence. Thus, although the study in the Aegean and Ionian islands of Greece[57] was not supportive of the Kinlen hypothesis, this could have been a result of the use mortality data. For example, it is possible that population mixing due to tourism could have actually resulted in an increased incidence of childhood leukaemia but this was not captured in the analysis of mortality.

The majority of past studies have examined the population mixing hypothesis in study areas considered to be isolated. According to the Kinlen hypothesis, herd immunity to common infections is low in geographically isolated areas compared to urban areas, either because the opportunity for infectious transmission is fewer, or because small populations do not allow for the disease to be maintained in an endemic form.[13] In any case, areas with a low herd immunity are expected to be highly susceptible to an epidemic of infection (and subsequent increase in leukaemia cases) upon population mixing. On the other hand, in areas with a relatively high herd immunity, a large population influx would not be expected to cause a high rate of infection since most individuals would be immune. In attempting to consider various levels of population mixing, in contrast to Kinlen's studies which focused on more extreme levels, a few of the UK studies examined the effect of population mixing in all geographic regions of England and Wales.[47,49,55] However, a limitation common to these studies is the failure to account for the varying levels of geographic isolation included in the study areas. Regardless, in two of the studies population mixing was found to be related to rates

of childhood leukaemia.[47,55] However, in one study, an association was found only when restricting the analysis to rural (i.e. isolated) areas.[49] If herd immunity does affect the potential for an epidemic of infection causing an increase in childhood leukaemia, then it is possible that the relationship between population mixing and childhood leukaemia differs in isolated areas compared to non-isolated areas.

This study examined the possible viral etiology with respect to the Kinlen hypothesis using incidence rather than mortality data. In contrast to most previous studies, particularly those of Kinlen, which focused on situations of very high levels of population mixing in the UK, this study conducted in Canada offered the opportunity to explore the association of childhood leukaemia with more moderate levels of population mixing, as suggested in the most recent UK study.[47] In addition to studying variations in population mixing, varying levels of geographic isolation were measured and accounted for in the analyses thus offering an opportunity to examine possible differences between isolated and non-isolated areas with respect to the relationship between population mixing and childhood leukaemia. Given that the existing literature indicates that the Kinlen hypothesis is most plausible for ALL among children in the youngest age group, analyses were conducted separately on each specific age group and for the ALL subtype.

3. DESIGN AND METHODOLOGY

3.1. Objectives

The objective of this study was to determine the ecologic association between population mixing and the incidence rate of leukaemia among children, with respect to varying levels of geographic isolation. To address possible etiological differences by age, and in particular among young children, risk was examined in the three specific age groups of 0 to 4, 5 to 9, and 10 to 14. In addition, risk was examined separately for the ALL subtype.

3.2. Overview of Study Design

This study employed an ecologic design in order to assess the effect of population mixing on the incidence rates of childhood leukaemia. The ecologic unit of analysis was defined by the census subdivision (CSD) in a 5-year period. Incident leukaemia cases, including all diagnostic subtypes, occurring in children aged 0 to 14 between the years 1978 to 1992 in Ontario were examined. The study included three 5-year periods: 1978-82, 1983-87 and 1988-92. Variables used to represent population mixing were changes in population size and migration levels between census years, which were both available from census data. The relationship between each population mixing indicator and childhood leukaemia incidence was examined for two levels of geographic isolation, measured by: urban-rural status, proximity to an urban centre, and highway accessibility. In addition, the possibility of confounding by nuclear radiation and socioeconomic status were assessed. Specifically, CSD characteristics of distance from a nuclear facility along with employment, education and income measures were examined.

3.3. Study Period

Incident leukaemia cases between the years 1978 to 1992 were examined. This study period

reflected the constraints of the census years and the availability of census migration data. Also, the recognition of a possible latency period was considered in determining the study period. Given the age peak in incidence of approximately 2 to 3, it would be reasonable to assume a latency period between exposure to population mixing and incidence of childhood leukaemia of about 2 to 3 years. However, according to the Kinlen hypothesis, it is a lack of exposure (to population mixing) that results in susceptibility to the effect of population mixing. Thus, accounting for a period during which there is an absence of exposure (likely during infancy) and also the age peak, a very short latency period of perhaps 1 to 2 years is possible. In this study, a latency period of 1½ years was applied utilizing mid-year population and demographic estimates from each census, and annual counts of leukaemia cases based on the year beginning in January and ending in December. For example, population mixing occurring between the census years of 1976 to 1981 corresponded to childhood leukaemia incidence between 1978 and 1982. Likewise, population mixing between the census years of 1986 and 1991 corresponded to leukaemia cases between 1988 and 1992. Census information was obtained from the census year included in each of the 5-year periods, i.e. 1981, 1986 and 1991.

3.4. Census Geography

Several levels of geography, relating to census data, are referred to in this study and are described in detail in census publications.[60] Briefly, the smallest geographic area for which census data are available is the enumeration area (EA). EAs are delineated based on the number of dwellings that will be canvassed by one Census Representative and varies between roughly 100 to 400 households. All larger census areas are made up of aggregated EAs. The next larger census area is the census subdivision (CSD) which generally refers to municipalities and includes Indian reserves, Indian settlements and unorganized territories. Census metropolitan areas (CMA) and census agglomerations (CA) are made up of CSDs. A CA is the economic and socially integrated area surrounding and

including an urbanized area (core) with a population of 10,000, based on the previous census. When the population of the urbanized core is 100,000 or greater, then the area is called a CMA.

3.5. Ecologic Units

The unit of analysis for this ecologic study was the CSD in a 5-year period. CSDs are identified by a Standard Geographical Classification (SGC) code by Statistics Canada.[61] In Ontario, there were 932 CSDs in 1981, 956 in 1986 and 951 in 1991. Exposure information on population mixing, geographic isolation and other variables of interest were obtained for these 2,839 CSDs from census data and other sources and linked to an SGC code.

The differing numbers of CSDs among census years is primarily due to geographical boundary changes which have occurred with some CSDs, including such events as amalgamations of municipalities, the splitting of municipalities, and the shifting of boundaries such that the size of one CSD increases while the size of an adjacent CSD decreases. Because the population mixing variables (changes in population size and migration levels) reflect changes between census years, it is necessary that the geography of a CSD remain constant between any two census years so that an observed change in population in fact reflects a true population change rather than a change in geographical boundary. Thus, each CSD associated with the three time periods were treated as independent geographical units and population mixing indicators were calculated using values that corresponded to a constant census geography based on a given census year. For example, for the 5-year period 1983 to 1987 (which corresponds to information on population mixing between the 1981 and 1986 censuses), if the boundary of a CSD changed between 1981 and 1986, then the population in 1981, adjusted to reflect the population if the boundaries were the same as that in 1986, was used in calculations. The adjusted populations were directly available from census data.

In Ontario, CSDs are also identified by Ministry of Health (MOH) codes. In the data received

from the Ontario Cancer Registry, MOH codes were provided as residence information for leukaemia cases. For the most part, the geography identified by an MOH code and any boundary changes that may occur are identical to that identified by the SGC code. The only differences are where some geographic areas associated with an MOH code include more than one CSD as identified by an SGC code. These are generally CSDs with small total populations (<5,000) and some of the large unorganized territories. There were 54 CSDs identified by MOH codes that corresponded to 130 CSDs identified by SGC codes. In this study, ecologic units corresponded to the 2,763 CSDs identified by MOH codes. For areas identified by an MOH code that included more than one CSD (as identified by an SGC code), exposure data from the comprising CSDs were combined (i.e. populations were summed, income was averaged).

3.6. Leukaemia Cases

Information on new diagnoses of leukaemia (ICD-9 site codes 204-208) between 1978 and 1992 among the population of children aged 0 to 14 in Ontario was obtained in computerized form from the Ontario Cancer Registry (OCR), which includes cancer patients diagnosed since 1964.[62] Specific information that was obtained for each case included: year of diagnosis, sex, age group at diagnosis (0 to 4, 5 to 9, or 10 to 14), leukaemia site (coded according to ICD-9), histology (corresponding to ICD-O codes), residence (CSD) at diagnosis, 6-digit postal code (when available), and the sources of information for case's registration in the OCR.

There were 1418 cases of childhood leukaemia registered at the OCR during the study period, of which 1178 were of the ALL subtype. A total of 31 cases were not linked to a CSD for such reasons as: the case was not assigned to a CSD, the case was assigned to a county, or the residence code was incorrect (either a code that had never existed or a code that had changed between census years). For unlinked cases that were assigned a residence code that had changed, a listing of residence

code changes, provided by the Ontario Ministry of Health, was used to assign that case to the correct CSD. Also, there were some unlinked cases, corresponding to the 1986 and 1991 census years, for which postal codes were available. For these cases, the Statistics Canada Postal Code Conversion files (which converts 6-digit postal codes to CSDs) were used to assign that case to the correct CSD.[63,64] Of the 31 unlinked cases, 10 were subsequently assigned to a CSD. Thus, there were 1397 childhood leukaemia cases, of which 1160 were ALL, available for analysis (98.5% of all registered cases). For each CSD and 5-year period, the number of cases by age group and sex was determined.

3.7. Population At Risk

Population denominator data for each CSD were obtained from the Census of Canada for the years 1981,[65] 1986[66] and 1991,[67] and manually entered into a database. For each CSD and census year, these mid-year estimates of male and female population in age groups 0 to 4, 5 to 9, and 10 to 14 were multiplied by five years to calculate the person-years at risk in each age-sex group over a 5-year period.

3.8. Variables Analyzed

The main exposure variables related to the Kinlen hypothesis were population mixing and geographic isolation. Covariates that were accounted for were socioeconomic status and exposure to the nuclear industry. A summary of the independent variables, including components used in calculating the variable and source of information, is provided in table 3.1.

3.8.1. Population Mixing

Population mixing, as it relates to the Kinlen hypothesis, refers to a population influx in a geographically isolated area. Two factors based on census data, were used to represent population

mixing over a 5-year period (i.e. between two censuses).

Percent Change in Population The percentage change in population was used to represent population mixing as a result of a net increase in population which, through an increase in population density, would increase the potential for contact between susceptible and infected people. This variable was calculated for each ecologic unit using the population in a given census year, and the population in the preceding census year given that the boundaries were the same as in that census year. Both population estimates were available from census data and obtained in computerized form using the computer software package PCensus,[68] provided by the Documents Library at Queen's University. For the main analysis, percent change in population was analyzed in categories of equal increments relative to a category of negative population change. Categories were: ≤0%, >0% to ≤10%, >10% to ≤20% and >20%.

Migrant Population The proportion of the total population of a CSD that are migrants represented population mixing where a high population growth may have been balanced by a high level of out-migration, such that the area did not experience a net increase in total population, yet the original population may have been exposed to new infectious agents. Migrants include the population 5 years of age and older in a CSD who were not a resident of that CSD 5 years before and they can originate from the same province, from a different province or from outside Canada.[69] The number of migrants and the total population greater than 5 years of age were available from PCensus. Similar to the situation with the percent change in population, the proportion of the population that are migrants was intended to be classified into four categories with 10% increments beginning at 0%. However, due to the small number of CSDs and the absence of leukaemia cases in the lowest (referent) category (>0% to ≤10%), this variable was analyzed in categories of: 0% to ≤20%, >20% to ≤30% and >30%.

3.8.2. Geographic Isolation

The Kinlen hypothesis specifies that population mixing results in an increased incidence of childhood leukaemia within geographically isolated areas. Thus, the level of geographic isolation of a CSD was accounted for. Three variables were employed as indicators of geographic isolation. Since population influx can conceivably alter the isolation of a community (e.g. an influx of population can increase the population density, more paved highways may be built in response to a larger population), geographic isolation was measured for the census year starting the 5-year period over which population mixing was measured (i.e. the first census of the two relevant census years).

Urban-rural status Since it is assumed that the herd immunity to the relevant infection would be high in an urban centre, where a high population density would facilitate the transmission of infections, relative to a rural location, where frequency of contact between individuals would be lower, one measure of geographic isolation used was the degree of urbanization of a CSD. Similar to Statistics Canada, an urban centre was defined as a CSD with a population of at least 1,000 and a population density of at least 400 per square kilometre based on the previous census.[69-71] For a given census year, populations in the previous census year for each CSD was available from PCensus. Land areas (corresponding to the current census) were obtained from Statistics Canada Census publications[65-67] and manually entered into the database. Urban-rural status was treated as a dichotomous variable.

Proximity to an urban centre Travel to an urban centre by individuals living in a rural CSD in close proximity to an urban centre can be expected to be frequent. Thus, although this CSD may be considered rural, the herd immunity would be more similar to that of the urban centre, where frequency of contact between individuals are relatively high. For this study, urban centre was defined as a Census Metropolitan Area (CMA) or a Census Agglomeration (CA). In order to determine proximity

of a CSD to an urban centre, the distance between the centroid of each CSD and the centroid of the nearest CMA or CA was calculated. A centroid refers to the geographical point that represents the central location of an area.[71] Centroids, in latitude-longitude coordinates, were available for all CSDs, CMAs and CAs from PCensus. Because the earth's surface is curved and Ontario covers roughly a range of 15 degrees in latitude and 21 degrees of longitude, the shortest distance between two centroids could not be simply calculated using Pythagorus's theorem without resulting in significantly distorted distance estimates.[72] Rather, distances were calculated using measurements along the 'great circle arc', which is the arc produced when a plane passes through the centre of the earth.[72] With the assistance of the Geographic Information Systems Laboratory at Queen's University, the distances between centroids were calculated in this way. CSDs were then classified into two categories based on whether it fell within 50 kilometres or greater than 50 kilometres from an urban centre.

Highway accessibility There are some CSDs, particularly in northern Ontario, which have only one highway going through it. Additionally, some areas do not have any highways and are accessible either through seasonal roads (e.g. ice roads) or through airplane. These areas are, therefore, not subject to exposure to a high traffic of people through such things as commuting or traveling and can be considered isolated. An approximate central point within a CSD was identified and the number of King's Highways and two-lane hard surface secondary and other highways (as defined by the Ontario Ministry of Transportation and Communications Road Maps) was counted for the years 1976, 1981 and 1986.[73-75] 'Number of highways' was intended to be categorized as zero, one to two, and three or more. However, due to small numbers in the 'zero' category, CSDs were classified as having zero to two or three or more highways.

3.8.3. Covariates

Other suspected risk factors for childhood leukaemia which were of interest as potential confounders were socioeconomic status and proximity to nuclear facilities.

Socioeconomic Status (SES) Previous studies have shown a higher risk of childhood leukaemia associated with a high SES and this has been observed at both the individual and ecologic level.[47] SES is also potentially related to population mixing, in that areas associated with a higher SES, for example as measured by low unemployment, would likely attract more incoming individuals and families. Thus, it is necessary to account for the potential confounding effect of SES. Variables related to employment, income and education were considered as indicators of SES in this study.

It has been shown that when a factor (with several levels) is a confounder because of its association with the exposure and outcome at the individual level, a single summary measure, such as the proportion in one category, may not provide adequate control of that confounder at the ecologic level and that it is best to use more than just a single summary.[76] When a covariate has several possible levels at the individual level, it is recommended that the covariate distribution of an ecologic unit be summarized by a set of variables representing the proportion of individuals in each level. For example, rather than analyzing the proportion of individuals 15 years or older in a CSD with a highest level of education at grade 13, the proportions of individuals in all education categories can be determined. Categories considered for highest level of education were: less than grade 9, grade 9-13 (with and without a secondary certificate), trade certificate/diploma, and university degree. Education was thus analyzed as a set of continuous variables that were considered simultaneously. Due to limitations in the availability of data, income could not be similarly analyzed. Income was therefore represented by the average individual income for a CSD, standardized to a 1986 constant dollar value based on the Ontario Consumer Price Index.[77] Because employment has only two levels at the

individual level, the percent in one category, i.e. unemployed, was analyzed. All SES information was obtained from PCensus.

Proximity to Nuclear Facilities Another potential confounder that was considered was proximity to nuclear facilities. Research on population mixing was in fact initiated by the excess childhood leukaemia observed near nuclear facilities in England and Scotland. At this time, radiation from these facilities was suspected of being the responsible environmental risk factor. Although the role of ionizing radiation with respect to childhood leukaemia is well substantiated, the evidence regarding nuclear facilities is weak.[19] However, since the nuclear industry can be associated with population mixing, (e.g through the availability of employment leading to a population inflow into an area) it is reasonable to consider its potential confounding effect. CSDs within a 25 kilometre distance from a nuclear facility was designated as 'near a nuclear facility'. The choice of a 25 kilometre distance was based on previous studies which have examined nuclear facilities as a potential point source of environmental risk.[78,79] This information was obtained from a previous study conducted in Ontario where CSDs were classified based on their location relative to circles with a 25 kilometre radius drawn around each nuclear facility.[80] All nuclear facilities in Ontario were already in operation prior to the beginning of the study period.[80]

3.9. Data Analysis

Once the data were reviewed for accuracy and completeness, analyses were conducted to describe case data and population exposure data, as well as to determine the association between the main variables and leukaemia rates in children. The statistical programs SAS[81] and EGRET[82] were used.

Table 3.1. Summary of independent variables

INDEPENDENT VARIABLE	VARIABLES USED IN CALCULATION	SOURCE
Population Mixing		
Percent Change in Population	• Total population in current census year	Census of Canada
	• Total population in previous census year	Census of Canada
Percent Migrant Population	• Total population aged 5 years and over	Census of Canada
	• Population 5 years of age or older that did not live in CSD in previous census	Census of Canada
Geographic Isolation		
Urban-rural status	• Total population in previous census year	Census of Canada
	• Total population in current census year	Census of Canada
	• Land area (current census year)	Census of Canada
Proximity to an urban centre	• Centroid coordinates for each CSD	Census of Canada
	• Centroid coordinates for each CMA or CA	Census of Canada
Highway accessibility	• N/A	Ontario Ministry of Transportation and Communications Road Maps
Covariates		
Unemployment Rate	• Total labour force	Census of Canada
	• Number of people unemployed	Census of Canada
Highest level of schooling	• Total population aged 15 years and over	Census of Canada
	• Population aged 15 years and over with highest level of schooling less than or equal to: grade nine	Census of Canada
	grade 13	Census of Canada
	trade, non-university	Census of Canada
	university	Census of Canada
Average Individual Income	• Consumer price indices based on 1986 $	CANSIM database
Proximity to a nuclear facility	• N/A	AECB study[80]

3.9.1. Description of Outcome and Exposure Data

In order to describe cases of childhood leukaemia diagnosed between 1978 and 1992 in Ontario, the distribution of incident cases by age group, sex, leukaemia subtype and 5-year time period was determined. For all levels of each of the variables, the number of cases and percent of the total cases included in the study (1397 cases) was calculated.

For each exposure variable, the number of CSDs and the population at risk, in person-years, was determined for each level of the variable. For descriptive purposes, continuous SES variables were dichotomized based on the Ontario average for each variable over the study period.

Analyses were conducted to give an indication of the nature of the relationship between independent variables. Chi-square tests and correlation coefficients were used to examine the relationships between the three indicators of SES, the relationships between the various geographic isolation indicators and the relationship between indicators of population mixing and SES.

3.9.2. Poisson Regression Modeling

In order to evaluate the association between population mixing and incidence rates for childhood leukaemia, while controlling for the effects of covariates, multiple regression modeling was used. A traditional linear regression approach is one way to assess this relationship. However, because the population at risk in each ecologic unit is of different sizes, the variability in the rates differ among each unit.[83] One method to account for this is to conduct a weighted regression where weights are assigned to each observation based on the amount of information they contribute.[84] In both the weighted and unweighted regression approaches the dependent variable can take on both negative and positive values. However, incidence rates can only be positive, therefore, a traditional regression may subsequently produce incidence rates which take on negative and therefore meaningless values.[84] An alternative regression method, which has been used in more recent ecologic studies of childhood leukaemia,[46-49] is Poisson regression analysis. The Poisson model is suitable for situations where the outcome variable is a count and as well when the outcome is a rare event such that the Poisson distribution can be used to make inferences.[85,86] Thus, since leukaemia is a rare event in the population aged 0 to 14, the observed number of incident cases was regressed on the population mixing indicator and covariates with a log link function, logarithm of the person-years at risk as the offset, and a

Poisson error.

Numerators for incidence rates were the number of cases, by age group and sex, of all forms of childhood leukaemia that occurred in each CSD for the 5-year periods of 1978-82, 1983-87 and 1988-92. Denominators were the person-years at risk, by age group and sex, for that same time period. Analyses were conducted on the whole age group and the three specific age groups for all childhood leukaemia subtypes combined as well as on ALL separately. Each CSD associated with the 3 time periods, was treated as an independent observation, or in other words, as 3 separate ecologic units of analysis.

3.9.3. Covariate Analysis

Since appropriate data were available, the education distribution of each CSD representing four levels of education, was analyzed as a set of three continuous variables representing the proportion of the population in each of the three highest education categories (the proportion of the population in the lowest education category was treated as a referent category and therefore omitted).[84] Unfortunately, similar information was not available for income. Thus, before conducting the covariate analysis, it was necessary to determine the best form to represent the variable average individual income. For average individual income, as well as unemployment rate, the original continuous form, a quadratic representation and a categorical representation were considered. The form associated with the smallest p-value, when fit in a model with the outcome, was subsequently used in the covariate analysis. This analysis was carried out with the data set containing all leukaemia cases as well as that restricted to ALL.

Analogous to the case with individual level studies, a variable is an ecologic confounder if it is associated with the outcome, and additionally, is ecologically associated with the exposure of interest across groups.[87] To select covariates to be included in the adjusted analyses, the 'confounder-disease

association' was evaluated for each of the three SES indicators and proximity to nuclear facilities. The aim was to identify variables that were associated with the outcome and consequently had the potential to confound the relationship between population mixing and the rate of childhood leukaemia. Thus, unemployment rate, average income, the education distribution, and proximity to a nuclear facility were examined for their association with the incidence of childhood leukaemia. A backward Poisson regression deletion procedure was conducted to determine which set of SES variables was associated with the incidence of childhood leukaemia and could therefore be considered as potential confounders. To ensure adequate control for confounding, the alpha level for deletion was set at 0.20.[84] This covariate analysis was repeated for the subset of data containing only ALL cases.

3.9.4. Main Analysis

Once the covariate model was established, the main variables related to the Kinlen hypothesis were examined. Population mixing in an area with a low herd immunity, and therefore a greater number of susceptible individuals, is expected to result in an epidemic of infection (and subsequent increase in childhood leukaemia cases). On the other hand, population mixing in areas with a high herd immunity would not result in the same degree of infection since more individuals are immune. Thus, herd immunity, measured by geographic isolation, can be considered an effect modifier of the relationship between population mixing and the incidence rate of childhood leukaemia. All analyses, therefore, examined the relationship between population mixing and the incidence rate of childhood leukaemia among individual levels of geographic isolation. Each population mixing variable was separately analyzed with each geographic isolation variable . A test for interaction between population mixing and geographic isolation was conducted by including product terms between population mixing and geographic isolation in the Poisson regression model containing the main variables and the selected covariates. The significance of the interaction parameters was tested by the score-test, which tests the

null hypothesis that the coefficients of the parameters associated with the interaction terms are zero.[82] P-values were calculated by comparing the score-test statistic to a chi-square distribution with the appropriate degrees of freedom. For each population mixing variable, crude and adjusted rate ratios at each level of population mixing for both isolated and non-isolated areas were determined, along with the associated 95% confidence interval. For the group of children for which the Kinlen hypothesis is most relevant (i.e. ALL in children aged 0 to 4), the overall association between population mixing and ALL, adjusted for SES, was examined. Also, to determine whether there was any gradient of effect, a test for trend was carried out.

3.9.5. Secondary Analyses

Since the Toronto Regional Municipality includes six CSDs (less than 1 percent of both the total number of CSDs and the land area in Ontario) yet contains roughly 20 percent of the population, analyses using the same model as in the main analysis were conducted on the set of data with the Toronto Regional Municipality excluded. This was done to evaluate the influence of Toronto-area data on the observed results.

Over-reporting of childhood cancer has been a cause of concern regarding data quality with the OCR.[88] The main sources of data for registration in the OCR are hospital discharge summaries, pathology reports, death certificates and records of referrals to the Regional Cancer Centres or the Princess Margaret Hospital.[62] Cases registered exclusively by hospital records are most often associated with over-reporting.[88] Since 1975, efforts have been taken to minimize over-reporting due to hospital-only registered case (i.e. confirmation of diagnoses are sought). Additionally, in a previous study examining childhood leukaemia around nuclear facilities in Ontario, hospital chart reviews were conducted to confirm diagnoses for incident cases occurring from 1964 to 1986.[80] Confirmation of diagnosis for recent cases registered in the OCR (1993 to present) has not yet been completed,

however, these cases are not included in this study. Thus, over-reporting has been resolved for cases included in this study. However, there still existed some hospital-only registered cases. Therefore, to evaluate the influence of possible over-reporting, analyses were carried out using the same model as in the main analysis excluding hospital-only registered cases.

Since confirmation of diagnosis began only a short time prior to the beginning of the study period, it was possible that over-reporting may have influenced results predominantly in the earlier time periods. To address this issue, as well as to examine a period effect, subanalyses were conducted on each 5-year period separately.

3.10. Ethical Issues

The study protocol was approved by the Queen's University Health Sciences & Affiliated Teaching Hospitals Research Ethics Board and a confidentiality agreement was undertaken with Cancer Care Ontario, prior to release of data by the OCR.

The original data file from the OCR was used to conduct descriptive analyses and to assign individuals to the correct ecologic unit. Upon completion of this task, all identifying information was removed and the resultant data file was used for subsequent analyses. All study results are displayed in the form of aggregate statistics.

4. DATA QUALITY

The two main sources of data in this study were the Ontario Cancer Registry and the Census of Canada. Quality of the data, particularly that of the Census of Canada, had an effect on the study area included as well as the statistical analyses.

4.1. Census of Canada Data

Census of Canada information was in fact obtained directly in computerized form from the software package PCensus. Similar to any data collection effort, census data are subject to errors. The most common types of errors include: coverage errors, non-response errors, response errors, processing errors and sampling errors.[65-67] Several methods of data checking are employed and studies are conducted to evaluate the quality of responses. In addition to these errors, several other issues with census data are cause for concern over data quality in this study. A brief description based on information from census publications follows.[65-67]

4.1.1. Incomplete Enumeration

In some CSDs, specifically on some Indian reserves and Indian settlements, enumeration was not completed. This resulted in missing data for all variables for these areas. Therefore, these CSDs were not available for analysis. Figures 4.1 and 4.2 present the distribution of CSDs from the 1991 census according to whether or not they were available for inclusion.

Because the variable 'percent population change' depends on population estimates for two census years, if incomplete enumeration occurred in the previous census, then this variable could not be calculated. Thus, although information on population denominators and covariates may be available, 'percent population change' is missing for some of the CSDs.

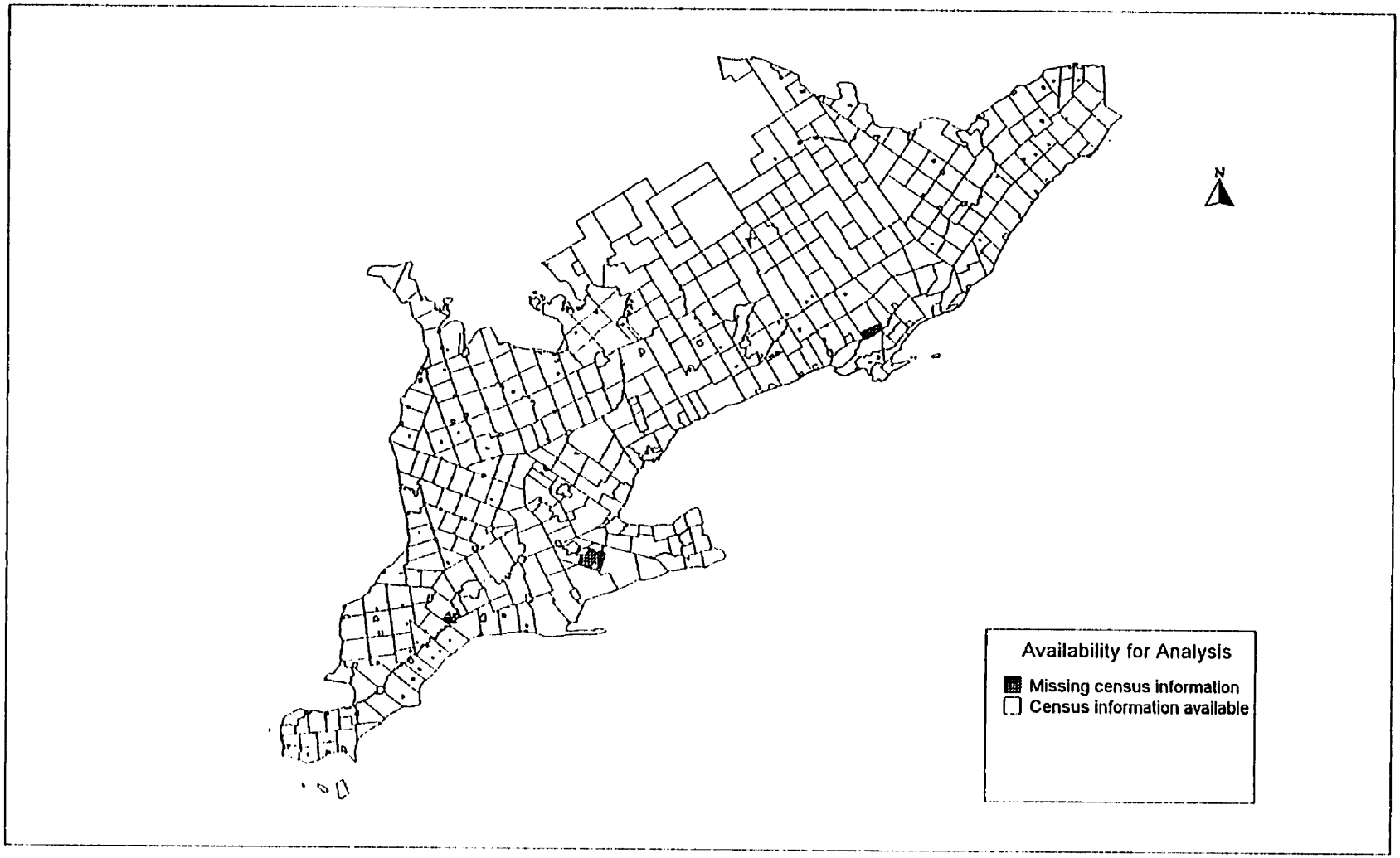

Figure 4.1. Census subdivisions according to availability for analysis, Southern Ontario

Source: Statistics Canada, 1991

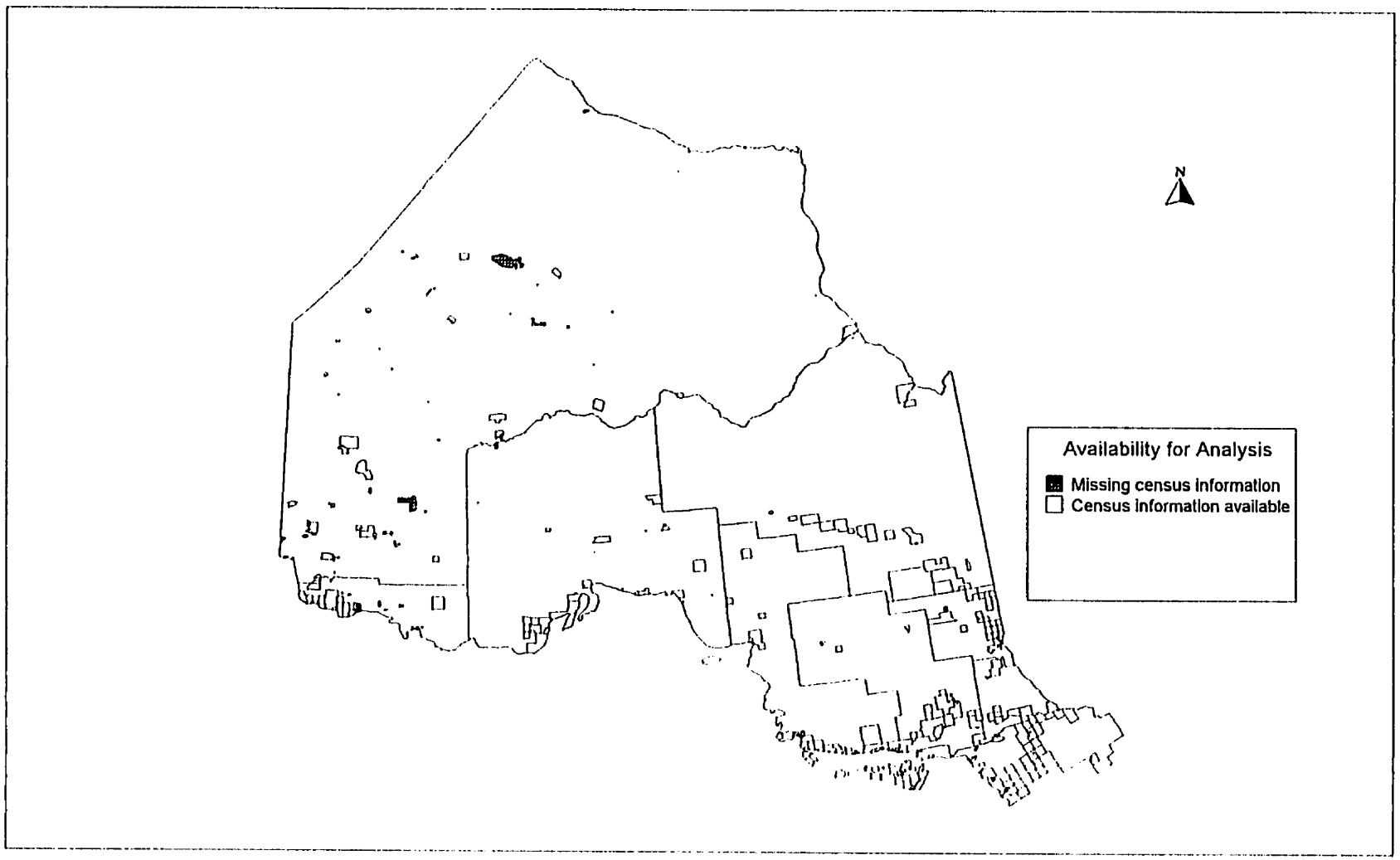

Figure 4.2. Census subdivisions according to availability for analysis, Northern Ontario

Source: Statistics Canada, 1991

4.1.2. Random Rounding

To provide a measure of confidentiality, numbers in census data are randomly rounded up or down to a multiple of five. Regardless of level of geography, this is done at the final stage after individual data have been used to determine the population count. Thus, it is possible that the sum of individually rounded numbers in distributions do not necessarily equal the total. For example, the number of people in each education group may not add up to the total population aged 15 years and over since each number has been randomly rounded independently.

The raw data that were used to create the PCensus database were counts at the enumeration area (EA) level, and therefore, had been randomly rounded at this level of geography. To obtain counts for CSDs, the counts for the comprising EAs were summed. Thus in most cases, the CSD values obtained from PCensus did not match CSD values in census publications exactly, since the published values used individual data and then randomly rounded at the level of CSD. The variables used in this study which were subject to random rounding were: the male and female populations in each of the three age groups, the number of people in each of the education groups and the total population over age 15, the number of people in the labour force and the number unemployed, and the number of migrants and the number of people aged 5 years or older.

Because some of the male and female populations in each of the age groups were quite small and therefore could be quite distorted (e.g. a count of 6 could be rounded to either 5 or 10), and because these estimates were used directly as denominator data in the Poisson regression analysis, these numbers were obtained from census publications and manually entered into the database, rather than using PCensus estimates. All entries were double checked for accuracy of entry.

For all other variables (see table 4.1), counts were used to calculate the variable that would be analyzed. Calculated values based on PCensus data were compared with calculated values based on published data for areas of differing size and found to be similar. For large populations, both

estimates agreed to the first decimal place. For smaller areas, agreement between estimates was not as precise, however, relative differences were not significantly distorted. For example, the relative proportions of people in each education group calculated with PCensus data compared to published data was similar. For CSDs where the total education distribution was not within the range of 90% to 110%, which were CSDs with very small total populations, counts for these areas were obtained from published values and manually entered into the database. Areas with a labour force population of 5 people were assigned a missing value for unemployment rate since numerators could only take on a value of 0 or 5 thus resulting in either a 0% or 100% unemployment rate.

Table 4.1. Calculated variables affected by random rounding

VARIABLE AND INDICATOR	CENSUS VARIABLE USED IN CALCULATION
Population Mixing	
Percent Migrants	• Number of people 5 years of age or older that did not live in CSD in previous census • Total population aged 5 years and over
Covariates	
Unemployment Rate	• Total labour force • Number of people unemployed
Highest Level of Schooling	• Total population aged 15 years and over • Number of people with highest level of education in each category

4.1.3. Area Suppression

Also in the interest of individual confidentiality, area suppression is done, where all characteristic data (including data on education, unemployment, migration, and income) for geographic areas with populations below a certain level are excluded. In this situation, the variable percent population change as well as population at risk information may be available, however, information on covariates and migrant population would be missing. This occurred for seven CSDs.

Income is further suppressed for areas with populations below 250 people. Thus, although

only seven of the CSDs available for analyzing population change were missing characteristic data based on area suppression, an additional 89 CSDs were missing information only on income. Thus, for 96 CSDs, information was available for population change but not for income. For 121 CSDs, information was available for percent migrants but not for income. This has implications with respect to adjusted analyses since the number of CSDs analyzed if income was included would be different than the number of CSDs included in the crude analysis. Because area suppression is conducted on the basis of small population size, it can be expected that the population at risk affected is very small. In fact, the analyses conducted with income in the model were missing 0.3% of the total population at risk (see table 4.2).

Table 4.2. Summary of ecologic units and population at risk

	Number of CSDs	Population At Risk (ages 0 to 14)
1981 Census	932	
1986 Census	956	(cannot calculate due to missing denominator data)
1991 Census	951	
Total	2839	
MOH identified	2763	(cannot calculate due to missing denominator data)
Missing Denominators		
Incomplete enumeration	75	
Area Suppression	41	
Total	2647	29,003,900
Percent Population Change		
Crude Analysis	2607	28,991,125
Adjusted Analysis	2511	28,905,175
Percent Missing Income		0.3%
Percent Migrant Population		
Crude Analysis	2645	29,002,850
Adjusted Analysis	2524	28,915,800
Percent Missing Income		0.3%

4.1.4. Non-permanent Residents

Non-permanent residents are individuals who reside in Canada through student or employment authorization, through a Minister's permit or who are refugee claimants.[89] In 1991, non-permanent residents were included in the census for the first time. Because of this change, total population counts, as well as counts for all other variables, were affected. The inclusion of these residents had the potential to artificially inflate estimates of 'percent population change' and the 'percent of the population that are migrants' for the period of population mixing between the 1986 and 1991 censuses. This effect would predominantly occur in the major CMAs (i.e. the Toronto CMA) where there is a concentration of non-permanent residents.

Of the total Ontario population, 1.3 percent were non-permanent residents in the 1991 census.[90] The percent of the population that were non-permanent residents in each of the CMAs in Ontario (excluding the Toronto CMA) ranged from 0.1 to 0.9 percent, affecting population sizes which ranged from 238,025 to 912,095. In the Toronto CMA, 2.5 percent of the population of 3,863,110 were non-permanent residents. Since in each CMA the total population was large and the percent that were non-permanent residents was quite small, it is unlikely that CSDs were considerably misclassified according to population mixing categories.

The inclusion of non-permanent residents could inflate denominators in the 1987-92 period causing rates to be underestimated. Again, since the proportion of non-permanent residents is small and this population would probably consist of mostly adults (e.g. people holding student or employment authorizations), denominators are likely not substantially affected.

4.2. Ontario Cancer Registry Data

As described in the previous chapter, over-reporting of cancer cases may be a cause for concern over data quality from the Ontario Cancer Registry. There were 30 cases of childhood

leukaemia registered exclusively on hospital records during the study period. 5 occurred between 1978-1982, 11 between 1983-1987 and 14 between 1988-1992. Of the 30, 13 were of the ALL subtype.

5. RESULTS

5.1. Description of Leukaemia Cases

There were 1397 cases of leukaemia diagnosed in children aged 0 to 14 between the years 1978 and 1992 and linked to a CSD. The distribution of cases by gender, age group, subtype and 5-year period is presented in table 5.1. Slightly more males than females were diagnosed, and the majority occurred in the youngest age group and the ALL subtype. Roughly equal proportions occurred in each of the three 5-year periods.

Table 5.1. Description of cases included in the study

VARIABLE	CATEGORY	NUMBER OF CASES	PERCENT
GENDER	Male	785	56.2
	Female	612	43.8
	Total	1397	100.0
AGE	0 to 4	766	54.8
	5 to 9	371	26.6
	10 to 14	260	18.6
	Total	1397	100.0
SUBTYPE	ALL	1160	83.0
	AML	164	11.7
	Other	73	5.2
	Total	1397	100.0
5-YEAR PERIOD	1978-82	451	32.2
	1983-87	464	33.2
	1988-92	482	34.5
	Total	1397	100.0

5.2. Description of Independent Variables

There were 2647 CSDs available for analysis which included 29,003,900 person-years at risk. In table 5.2, the independent variables are summarized according to number of CSDs and person-years at risk. Each of the independent variables are ecologic measures and therefore are characteristics of CSDs. Population mixing and geographic isolation indicators were grouped into predefined categories.

SES variables were dichotomized based on Ontario average values for the census years included over the study period. Also included are the proportion of CSDs and the proportion of the total population at risk that are missing values for each variable. Although a relatively high proportion of CSDs were missing values for income and urban-rural status (4.6% and 3.3% respectively), this affected 0.3% or less of the population at risk.

It was found that the variables distributed differently on CSDs and individuals. For example, although 87% of CSDs are characterized by having at least 55% of the total CSD population with a highest level of education less than or equal to grade 13, only 58% of the total population at risk reside in such areas. Similarly, the majority of the CSDs of Ontario are rural (77.9%), however, the majority of the total population at risk reside in urban CSDs (63.9%). In figures 5.1 to 5.6, maps showing the distribution of CSDs in 1991 according to measures of geographic isolation are presented.

5.3. Covariate Analysis

5.3.1. Relationships between SES indicators

Table 5.3 shows the correlations between the three SES indicators for CSDs in the three census years. Average individual income was positively correlated at the ecologic level with highest level of schooling being post-secondary education and university. Unemployment rate was only weakly correlated with income, post-secondary education and university degree.

5.3.2. Variable Representations

Education was analyzed as a set of continuous variables representing the proportion of the total population aged 15 and over in each of 3 categories of highest level of schooling. Unemployment rate and average individual income were considered in the original continuous form, a quadratic representation and as a categorical variable. Table 5.4 presents rate ratios for childhood leukaemia

and p-values associated with each form of unemployment rate and average individual income. Also, the rate ratios and p-values associated with the education variables that were analyzed are included. For both average individual income and unemployment rate, it was found that the continuous form of the variable provided the best representation. When the different representations were analyzed with the ALL data set, the same representations were selected (results not shown).

Table 5.2. Description of independent variables according to CSDs and population at risk

CSD CHARACTERISTIC	CATEGORY	% OF CSDs*	% OF POPULATION AT RISK†
HIGHEST LEVEL OF EDUCATION ≤ GRADE 13	> 55% of population aged 15+	87.0	58.0
	≤ 55% of population aged 15+	11.9	42.0
	Missing	1.0	0.0
UNEMPLOYMENT RATE	> 7%	50.5	44.8
	≤ 7%	48.3	55.2
	Missing	1.2	0.0
AVERAGE INDIVIDUAL INCOME	≤ $19,000	79.4	51.4
	> $19,000	15.9	48.3
	Missing	4.6	0.3
PROXIMITY TO A NUCLEAR FACILITY	Greater than 25 km	96.8	88.4
	Within 25 km	3.2	11.6
NUMBER OF PAVED HIGHWAYS	2 or less	50.0	8.7
	3 or more	50.0	91.3
DISTANCE FROM AN URBAN CENTRE	Greater than 50 km	30.0	5.2
	Within 50 km	70.0	94.8
URBAN-RURAL STATUS	Rural	77.9	35.9
	Urban	18.9	63.9
	Missing	3.3	0.2
PERCENTAGE CHANGE IN POPULATION	≤ 0%	30.2	18.9
	> 0% and ≤ 10%	41.6	48.0
	> 10% and ≤ 20%	16.4	16.2
	> 20%	10.3	16.9
	Missing	1.5	0.0
MIGRANT POPULATION	≥ 0% and ≤ 20%	38.2	41.2
	> 20% and ≤ 30%	42.0	40.1
	> 30%	19.6	18.7
	Missing	0.1	0.0

* Total CSDs = 2647

† Total person-years = 29,003,900

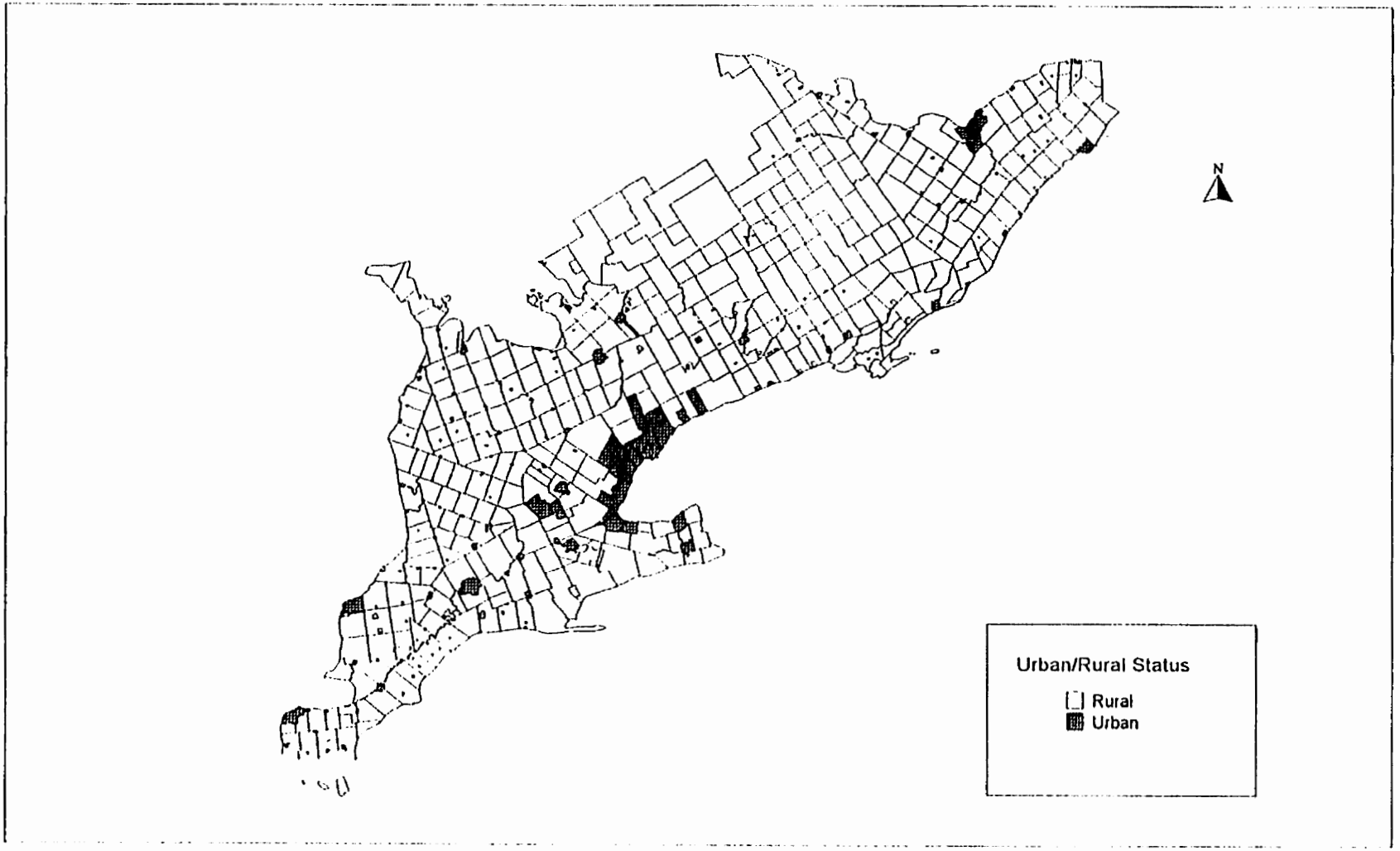

Figure 5.1. Census subdivisions by urban/rural status, Southern Ontario

Source: Statistics Canada, 1991

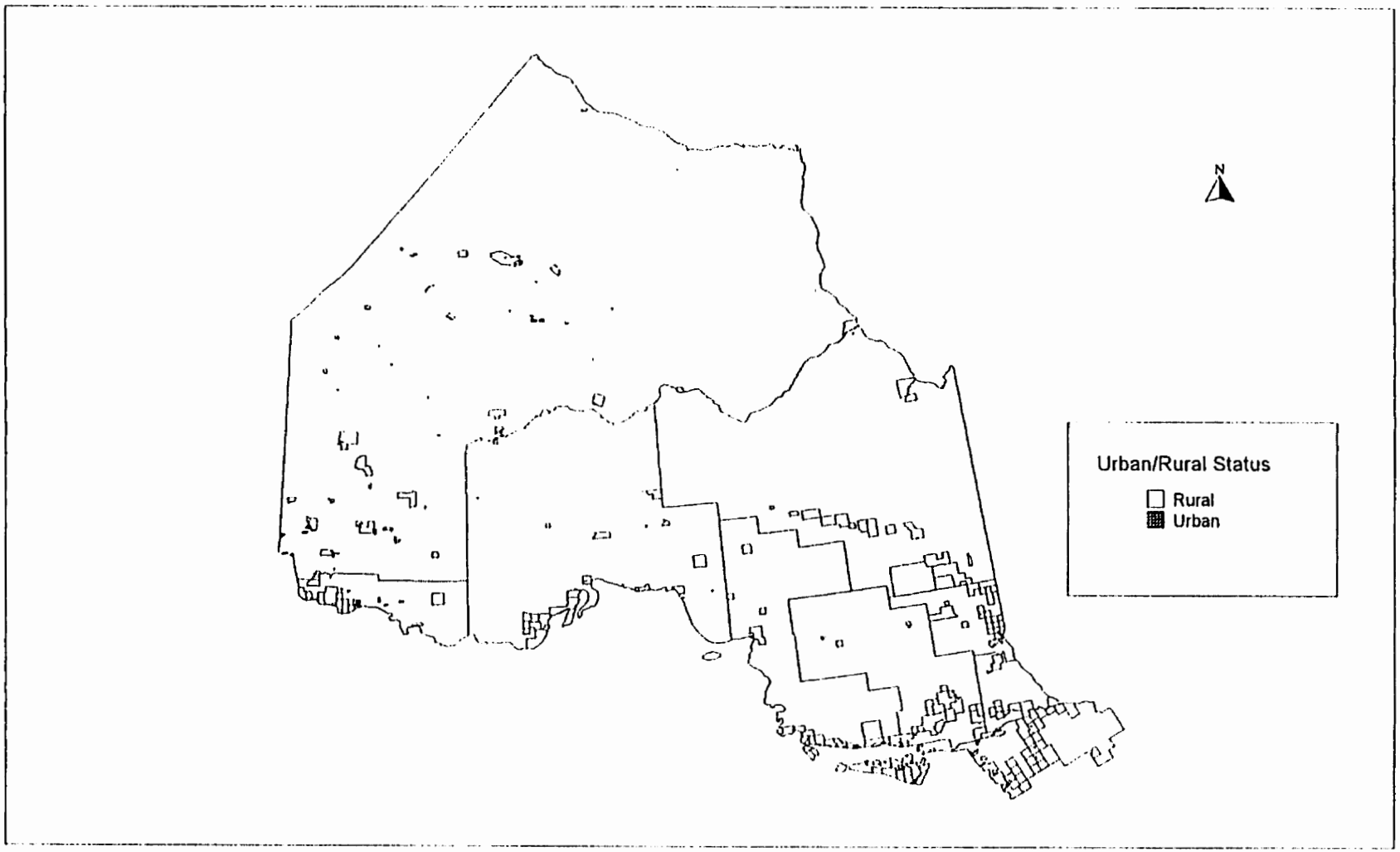

Figure 5.2. Census subdivisions by urban/rural status, Northern Ontario

Source: Statistics Canada, 1991

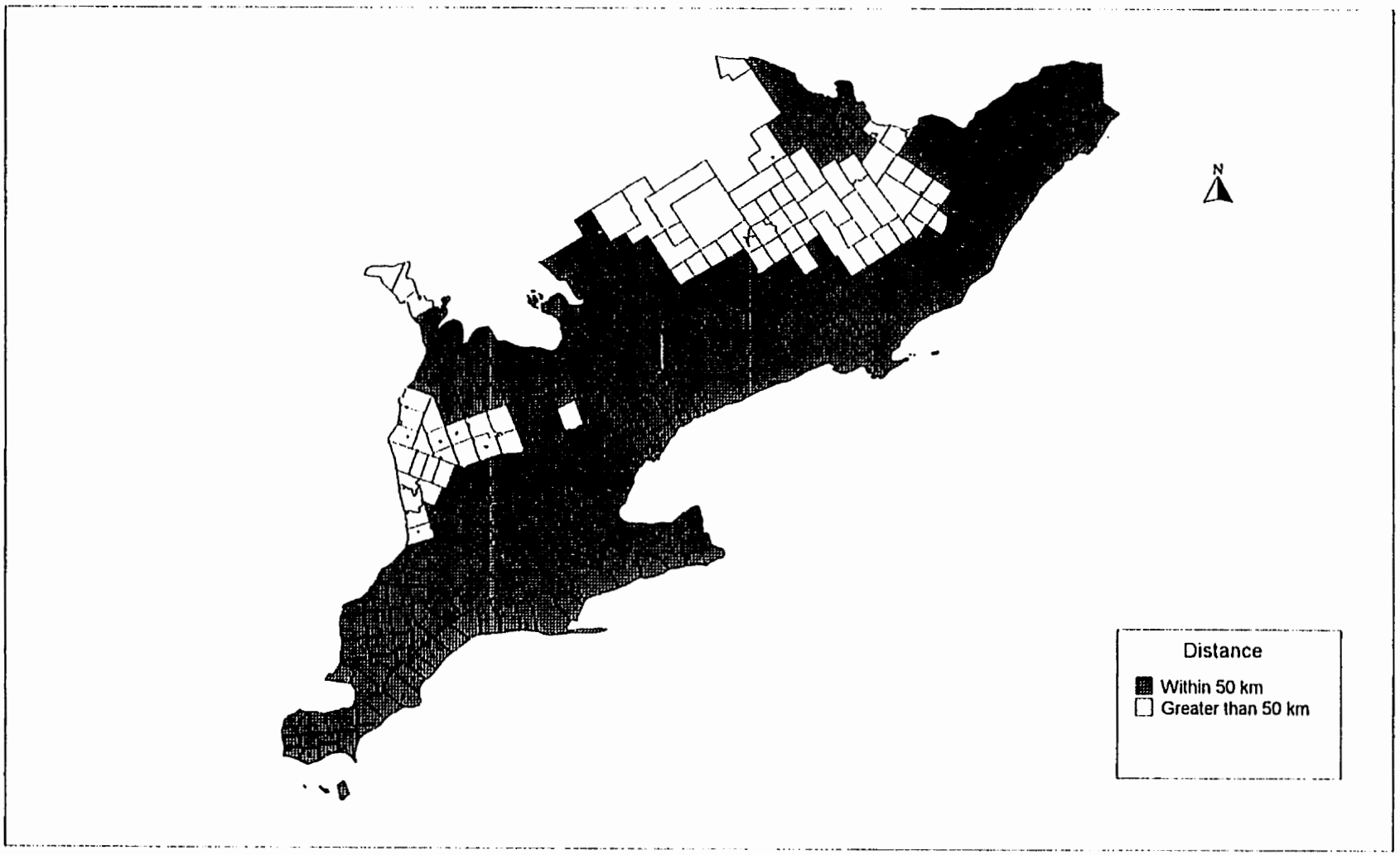

Figure 5.3. Census subdivisions by distance from nearest urban centre, Southern Ontario

Source: Statistics Canada, 1991

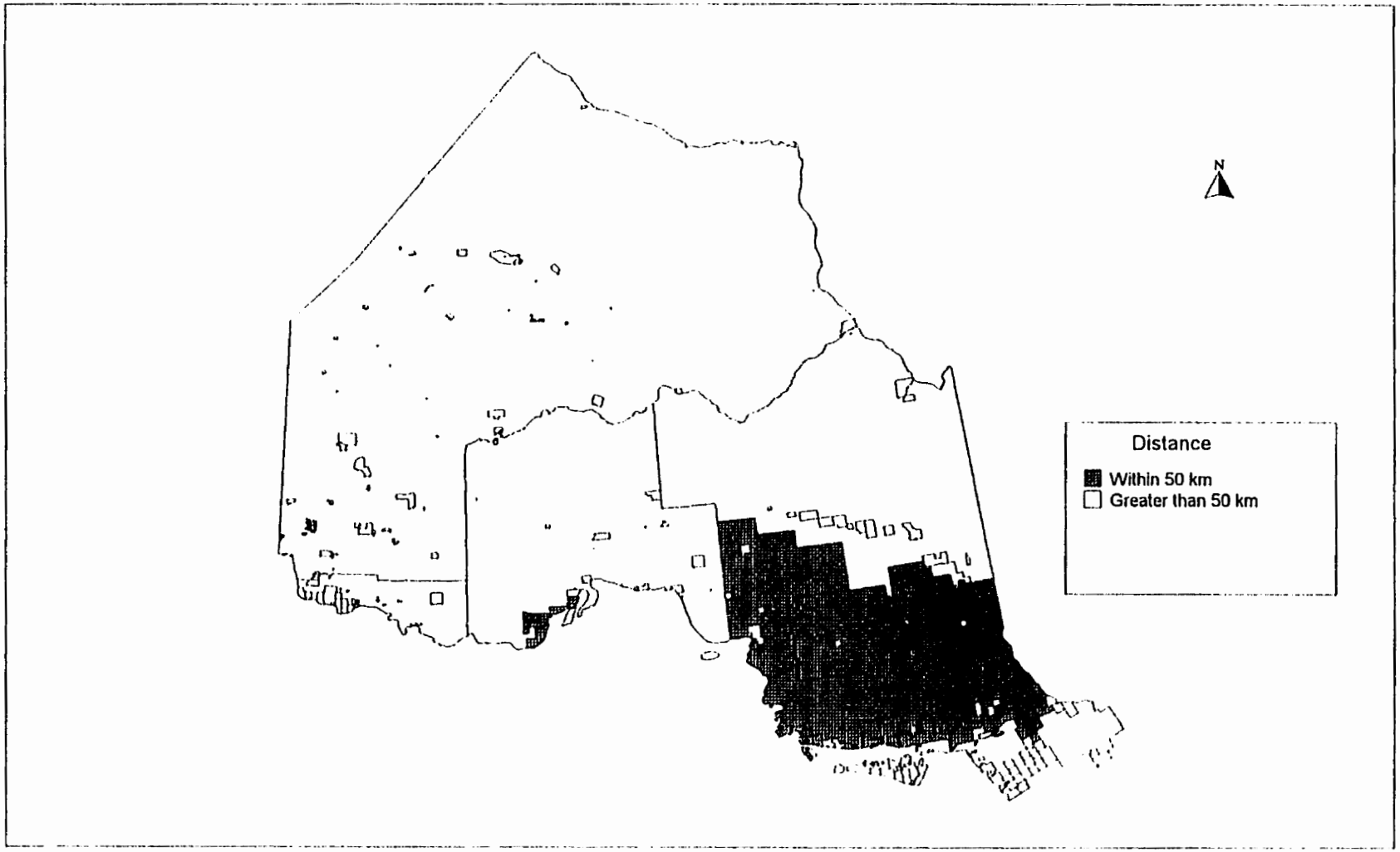

Figure 5.4. Census subdivisions by distance from nearest urban centre, Northern Ontario

Source: Statistics Canada, 1991

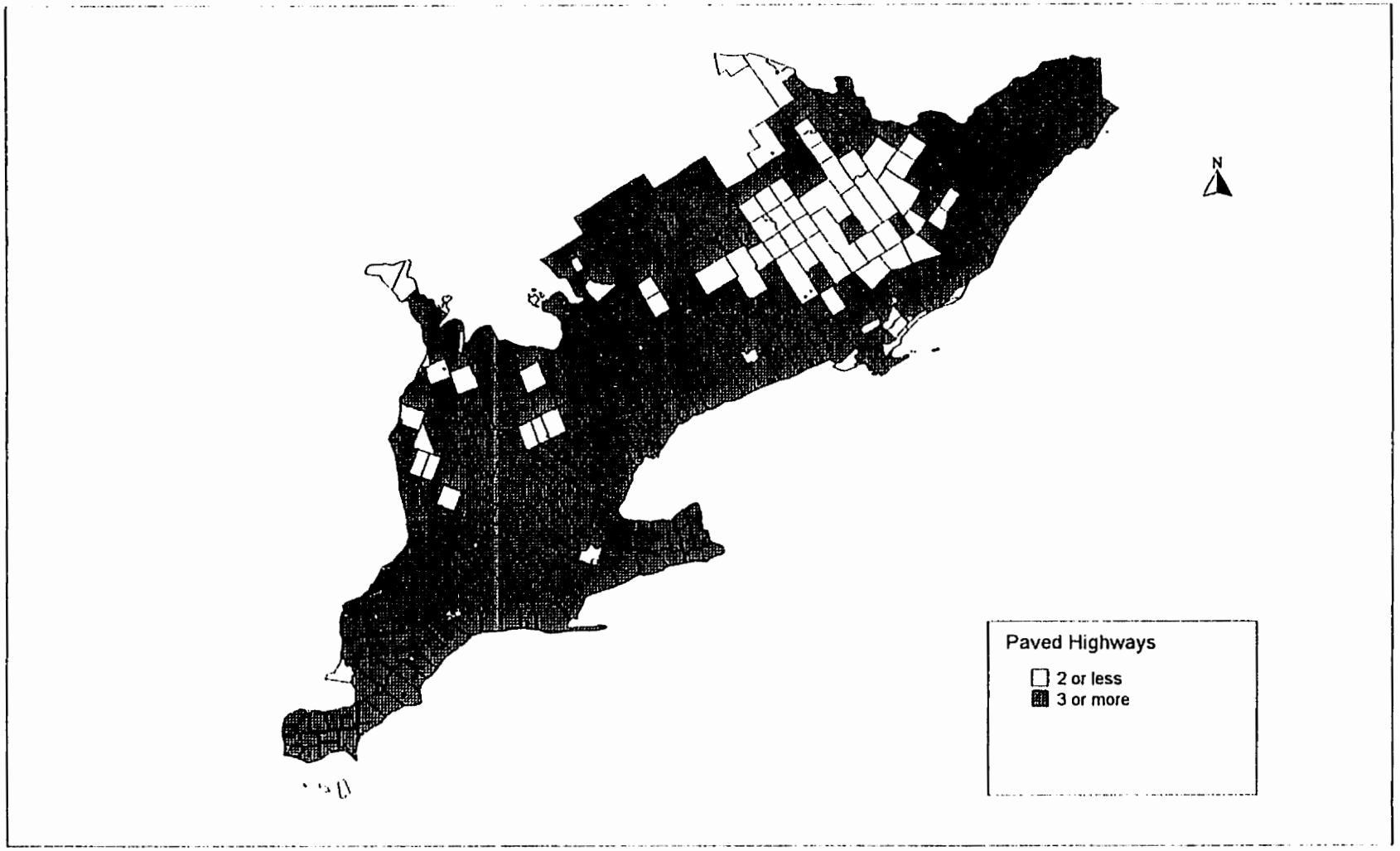

Figure 5.5. Census subdivisions by number of paved highways, Southern Ontario

Source: Statistics Canada, 1991

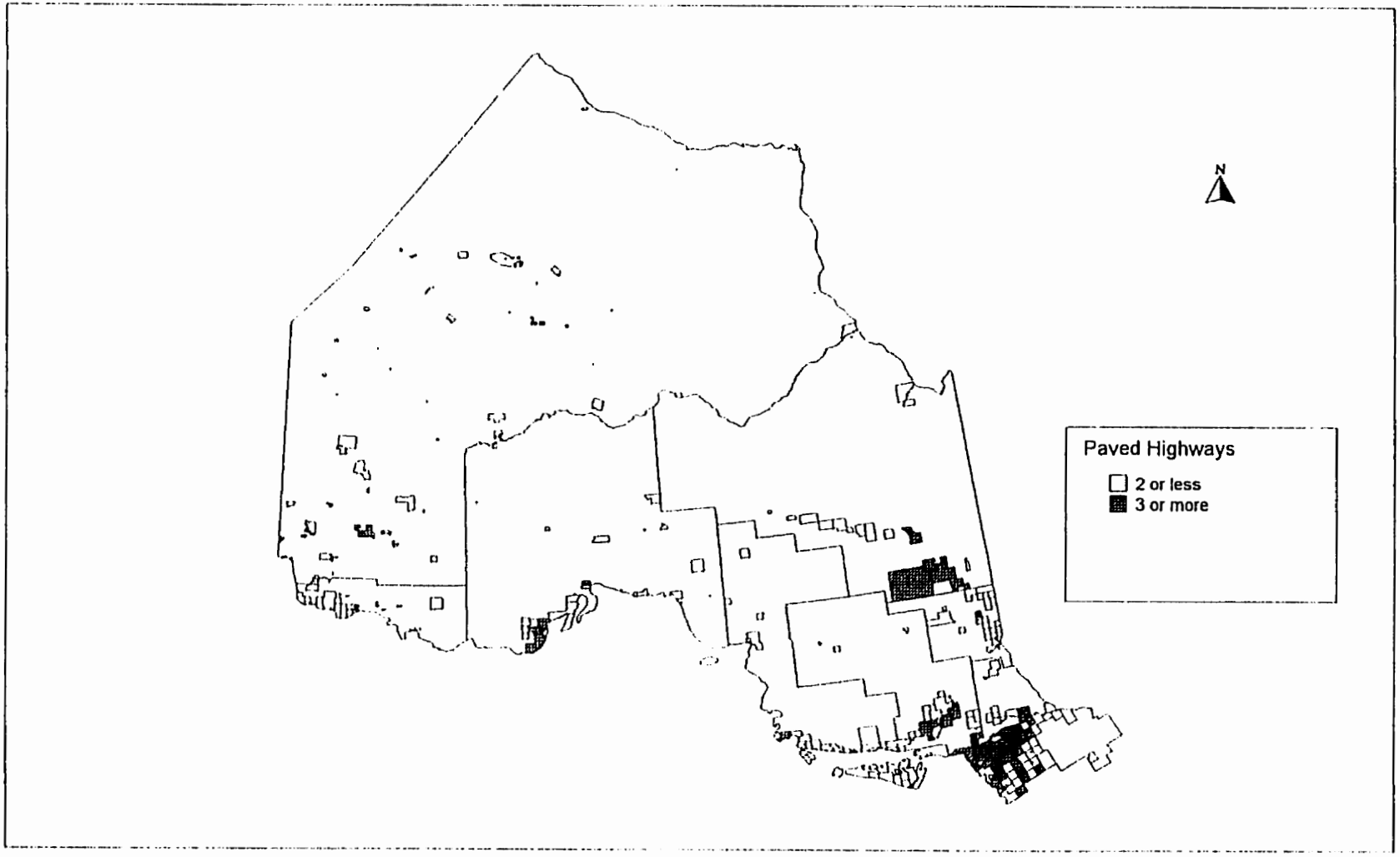

Figure 5.6. Census subdivisions by number of paved highways, Northern Ontario

Source: Statistics Canada, 1991

Table 5.3. Correlations between SES indicators for CSDs in Ontario in census years 1981, 1986 and 1991

Correlation coefficient*	Income	Unemployment	Grade 13	Post-secondary	University
Income	1.00	-0.32	-0.04	0.53	0.63
Unemployment		1.00	-0.00	-0.27	-0.26
Grade 9 or less			-0.38	-0.79	-0.51
Grade 13			1.00	-0.15	-0.30
Post-secondary				1.00	0.44
University					1.00

* Unit of analysis is the CSD; correlations calculated from observations which were not missing any values for any of the SES variables included (2522 CSDs)

Table 5.4. Rate ratios for the different representations of average individual income and unemployment rate and for the education variables

Variable	Representation	Rate Ratio*†	p-value	p-value (overall)‡
Average Individual Income	Continuous	0.92	0.304	—
	Quadratic			
	Income	0.85	0.724	0.579
	Income^2	0.87	0.865	
	Categorical			
	≤$10,000	1.00	referent	
	>$10,000 to ≤$15,000	1.13	0.715	0.732
	>$15,000 to ≤$20,000	1.01	0.971	
	>$20,000	1.03	0.930	
Unemployment Rate	Continuous	1.01	0.170	—
	Quadratic			
	Unemployment	1.02	0.330	0.372
	Unemployment^2	1.00	0.724	
	Categorical			
	≤5%	1.00	referent	
	>5% to ≤10%	1.08	0.223	0.488
	>10% to ≤15%	1.15	0.159	
	>15%	1.15	0.613	
Highest level of education	% with grade 13	0.96	0.680	
	% with post-secondary	0.84	0.007	0.013
	% with university degree	1.10	0.361	

* Rate ratio for continuous representation of income is per $10,000

† Rate ratios for each education variable are per 10% of population; rate ratios for education are adjusted for each other

‡ Where more than one term is used to represent variable (e.g. quadratic and categorical), p-value refers to the simultaneous addition of all terms

5.3.3. Variable Selection

Table 5.5 presents results pertaining to the backward deletion procedure. Observations which had information on all three SES indicators were used in this analysis. There were 2522 CSDs and 28,915,425 person-years at risk included. At an alpha level of 0.20, unemployment rate was not significant while the education distribution and average individual income were. Thus, unemployment rate was deleted and average individual income and the education distribution were refit. Both average individual income (p=0.129) and the education distribution (p=0.007) remained in the model. When fitting the covariate model including the observations that were originally excluded due to missing values for unemployment rate, 2524 CSDs and 28,915,800 person-years had information and the p-values did not change. Therefore, no more SES variables were deleted.

The variable regarding proximity to nuclear facilities was entered into a model containing average individual income and the education distribution. The p-value for proximity to nuclear facilities was 0.720 and thus non-significant at an alpha level of 0.20. The final covariate model therefore included average individual income and the education distribution. The variable selection procedure was also carried out using the data set that included only ALL cases and the final covariate model was identical.

Table 5.5. Results of backwards deletion procedure

Variable	p-values	
	Model including all SES variables*	After unemployment rate deleted†
Unemployment rate	0.424	–
Average individual income	0.175	0.007
Education distribution‡	0.010	0.129

* 2522 CSDs, 28,915,425 person-years
† 2524 CSDs, 28,915,800 person-years
‡ For the education distribution, p-values refer to addition of all three education indicators simultaneously

5.4. Results from Main Analysis

Following is a presentation of results regarding the main objective of this study. Each section presents results separately for the two indicators of population mixing used (percent population change and migrant population). Crude and adjusted rate ratios (RR) are presented along with the associated 95% confidence interval (CI). The discussion of results mainly focuses on the adjusted rate ratios, since in many cases there was no substantial confounding observed. Similarly, p-values for interaction are presented only for adjusted analyses since differences in p-values between crude and adjusted analyses were small. Due to area suppression of average individual income (see chapter 4), adjusted analyses included a fewer number of CSDs. As well, a maximum of two cases may be missing from adjusted analyses.

5.4.1. Results for Childhood Leukaemia Among Children Aged 0 to 14

Tables 5.6 to 5.8 present results for analyses of percent population change. Rate ratios are presented for increasing levels of percentage change in population, relative to a referent category of negative to zero population change, and stratified by geographic isolation (as represented by urban-rural status, distance from an urban centre and number of paved highways). Tests for interaction between each of the three indicators of geographic isolation and percent change in population (referring to analyses adjusted for SES) were all statistically non-significant.

According to the variable urban-rural status, there was some indication that percent population change influenced leukaemia risk. In particular, a slightly elevated risk was observed for all categories of population change, relative to negative population change, in rural areas. In urban areas, there was no association between percent population change and risk of childhood leukaemia after adjustment for SES. Rate ratios associated with percent population change in areas both distant and near an urban centre did not differ significantly from the null. Similarly, in both isolated and non-isolated

areas, in terms of number of paved highways, percent population change did not appear to influence risk of developing childhood leukaemia.

Table 5.6. Rate ratios for childhood leukaemia according to percent change in population, by urban-rural status for children aged 0 to 14

% Δ IN POPULATION	RURAL			URBAN		
	# OF CASES*	CRUDE RR (95% CI)	ADJUSTED RR†‡ (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR† (95% CI)
≤0	88	1.00 (reference)	1.00 (reference)	181	1.00 (reference)	1.00 (reference)
>0 to 10	186	1.15 (0.89,1.49)	1.21 (0.94, 1.76)	492	0.87 (0.74, 1.04)	0.92 (0.77, 1.10)
>10 to 20	72	1.13 (0.83,1.54)	1.27 (0.92, 1.76)	141	0.81 (0.65, 1.01)	0.92 (0.73, 1.17)
>20	85	1.20 (0.89,1.62)	1.38 (1.00, 1.92)	149	0.85 (0.69, 1.06)	1.04 (0.81, 1.33)
p (interaction)§ = 0.273						

* Three cases missing from analyses of urban-rural status
† RR adjusted for average individual income and the education distribution
‡ Two additional cases missing in adjusted analysis
§ P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 3 degrees of freedom

Table 5.7. Rate ratios for childhood leukaemia according to percent change in population, by distance from nearest urban centre for children aged 0 to 14

% Δ IN POPULATION	>50 KILOMETRES			≤50 KILOMETRES		
	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR*† (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
≤0	31	1.00 (reference)	1.00 (reference)	238	1.00 (reference)	1.00 (reference)
>0 to 10	24	0.83 (0.49,1.42)	0.85 (0.50, 1.46)	654	1.00 (0.86, 1.16)	1.04 (0.90, 1.22)
>10 to 20	8	1.01 (0.46,2.19)	1.00 (0.46, 2.17)	208	0.94 (0.78, 1.13)	1.06 (0.87, 1.30)
>20	7	0.93 (0.41,2.10)	0.71 (0.28, 1.83)	227	0.97 (0.81, 1.17)	1.16 (0.93, 1.44)
p (interaction)‡ = 0.744						

* RR adjusted for average individual income and the education distribution
† Two cases missing in adjusted analysis
‡ P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 3 degrees of freedom

Table 5.8. Rate ratios for childhood leukaemia according to percent change in population, by number of paved highways for children aged 0 to 14

% Δ IN POPULATION	≤2 PAVED HIGHWAYS			>2 PAVED HIGHWAYS		
	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR*† (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
≤0	47	1.00 (reference)	1.00 (reference)	222	1.00 (reference)	1.00 (reference)
>0 to 10	49	1.02 (0.68, 1.52)	1.04 (0.70,1.55)	629	0.99 (0.85,1.15)	1.03 (0.89,1.20)
>10 to 20	20	1.07 (0.64, 1.81)	1.10 (0.65,1.86)	196	0.92 (0.76,1.12)	1.04 (0.84,1.28)
>20	12	1.12 (0.60, 2.12)	1.01 (0.51,2.01)	222	0.96 (0.80,1.16)	1.13 (0.90,1.40)
p (interaction)‡ = 0.981						

* RR adjusted for average individual income and the education distribution
† Two cases missing in adjusted analysis
‡ P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 3 degrees of freedom

Tables 5.9 to 5.11 present results for analyses examining the population mixing variable of migrant population. Rate ratios are presented for increasing concentrations of migrant populations, relative to a referent category of 20% or less, and stratified by geographic isolation (as represented by urban-rural status, distance from an urban centre and number of paved highways). Tests for interaction between each of the three indicators of geographic isolation and percent migrant population were all statistically non-significant, thereby implying no effect modification by geographic isolation.

Overall, the level of migrant population did not appear to have a strong influence on leukaemia risk in either isolated or non-isolated areas, with respect to all three indicators of geographic isolation. In isolated areas, none of the rate ratios differed significantly from the null. However, in areas classified as urban where migrants constituted greater than 30% of the total population, a slightly increased risk of childhood leukaemia was observed in the adjusted analysis (RR=1.23, 95% CI 1.00-1.51).

Table 5.9. Rate ratios for childhood leukaemia according to percent of the population that are migrants, by urban-rural status for children aged 0 to 14

% MIGRANT POPULATION	RURAL			URBAN		
	# OF CASES*	CRUDE RR (95% CI)	ADJUSTED RR†‡ (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR† (95% CI)
0 to 20	166	1.00 (reference)	1.00 (reference)	431	1.00 (reference)	1.00 (reference)
>20 to 30	149	0.83 (0.67, 1.04)	0.88 (0.70, 1.11)	378	0.94 (0.82, 1.08)	0.99 (0.86, 1.15)
>30	116	0.98 (0.77, 1.24)	1.09 (0.84, 1.42)	154	1.06 (0.88, 1.27)	1.23 (1.00, 1.51)
p (interaction)§ =0.618						

* Three cases missing from analyses of urban-rural status
† RR adjusted for average individual income and the education distribution
‡ Two additional cases missing in adjusted analysis
§ P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 2 degrees of freedom

Table 5.10. Rate ratios for childhood leukaemia according to percent of the population that are migrants, by distance from nearest urban centre for children aged 0 to 14

% MIGRANT POPULATION	>50 KILOMETRES			≤50 KILOMETRES		
	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR*† (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
0 to 20	39	1.00 (reference)	1.00 (reference)	560	1.00 (reference)	1.00 (reference)
>20 to 30	22	0.72 (0.43, 1.22)	0.78 (0.46, 1.33)	506	0.92 (0.81, 1.04)	0.94 (0.83, 1.07)
>30	9	0.88 (0.43, 1.82)	0.78 (0.34, 1.74)	261	1.00 (0.86 1.16)	1.11 (0.94, 1.32)
p (interaction)‡ =0.613						

* RR adjusted for average individual income and the education distribution
† Two cases missing in adjusted analysis
‡ P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 2 degrees of freedom

Table 5.11. Rate ratios for childhood leukaemia according to percent of the population that are migrants, by number of paved highways for children aged 0 to 14

% MIGRANT POPULATION	≤2 PAVED HIGHWAYS			>2 PAVED HIGHWAYS		
	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR*† (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
0 to 20	54	1.00 (reference)	1.00 (reference)	545	1.00 (reference)	1.00 (reference)
>20 to 30	49	0.94 (0.64, 1.39)	0.99 (0.67, 1.47)	479	0.90 (0.80, 1.02)	0.92 (0.80, 1.05)
>30	25	1.09 (0.68, 1.76)	1.08 (0.66, 1.78)	245	0.98 (0.85, 1.14)	1.08 (0.91, 1.29)
p (interaction)‡ =0.917						

* RR adjusted for average individual income and the education distribution
† Two cases missing in adjusted analysis
‡ P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 2 degrees of freedom

5.4.2. Results for Childhood Leukaemia Among Specific Age Subgroups

Tables 5.12 to 5.14 present results for analyses of percent population change in children aged 0 to 4. Tests for interaction between the three indicators of geographic isolation and percent change in population indicated statistically significant effect modification by urban-rural status only (p=0.016). In areas classified as rural, an increased risk was observed among all levels of population change. The adjusted risk estimate was statistically significant in areas with a level of population change between 0% and 10% (RR=1.60, 95% CI 1.10-2.33). Conversely, a slightly decreased risk was observed among all three categories of population change in urban areas. For the analyses stratified by number of paved roads and distance from an urban centre, rate ratios in both isolated and non-isolated areas did not differ significantly from the null.

Table 5.12. Rate ratios for childhood leukaemia according to percent change in population, by urban-rural status for children aged 0 to 4

% Δ IN POPULATION	RURAL			URBAN		
	# OF CASES*	CRUDE RR (95% CI)	ADJUSTED RR*† (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
≤0	37	1.00 (reference)	1.00 (reference)	107	1.00 (reference)	1.00 (reference)
>0 to 10	109	1.57 (1.08, 2.28)	1.60 (1.10, 2.33)	273	0.79 (0.63, 0.98)	0.80 (0.64, 1.01)
>10 to 20	37	1.31 (0.83, 2.07)	1.38 (0.86, 2.22)	77	0.74 (0.55, 0.99)	0.78 (0.56, 1.07)
>20	49	1.49 (0.97, 2.28)	1.58 (0.99, 2.51)	77	0.71 (0.53, 0.96)	0.81 (0.58, 1.14)
p (interaction)‡ =0.016						

* RR adjusted for average individual income and the education distribution
† One case missing in adjusted analysis
‡ P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 3 degrees of freedo

Table 5.13. Rate ratios for childhood leukaemia according to percent change in population, by distance from nearest urban centre for children aged 0 to 4

% Δ IN POPULATION	>50 KILOMETRES			≤50 KILOMETRES		
	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR*† (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
≤0	17	1.00 (reference)	1.00 (reference)	127	1.00 (reference)	1.00 (reference)
>0 to 10	13	0.80 (0.39, 1.64)	0.78 (0.38, 1.60)	369	1.01 (0.83, 1.24)	1.01 (0.82, 1.25)
>10 to 20	4	0.89 (0.30, 2.64)	0.83 (0.28, 2.48)	110	0.90 (0.70, 1.16)	0.92 (0.69, 1.22)
>20	4	0.87 (0.29, 2.58)	0.66 (0.19, 2.27)	122	0.92 (0.72, 1.18)	0.98 (0.73, 1.32)
p (interaction)‡ =0.879						

* RR adjusted for average individual income and the education distribution
† One case missing in adjusted analysis
‡ P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 3 degrees of freedom

Table 5.14. Rate ratios for childhood leukaemia according to percent change in population, by number of paved highways for children aged 0 to 4

% Δ IN POPULATION	≤2 PAVED HIGHWAYS			>2 PAVED HIGHWAYS		
	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR*† (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
≤0	25	1.00 (reference)	1.00 (reference)	119	1.00 (reference)	1.00 (reference)
>0 to 10	31	1.17 (0.69, 1.99)	1.16 (0.68, 1.96)	351	0.98 (0.80, 1.21)	0.98 (0.79, 1.22)
>10 to 20	8	0.80 (0.36, 1.77)	0.76 (0.34, 1.68)	106	0.90 (0.69, 1.17)	0.92 (0.69, 1.22)
>20	6	0.96 (0.39, 2.34)	0.82 (0.31, 2.16)	120	0.91 (0.71, 1.18)	0.96 (0.71, 1.30)
p (interaction)‡ =0.791						

* RR adjusted for average individual income and the education distribution
† One case missing in adjusted analysis
‡ P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 3 degrees of freedom

When examining the effect of migrant population in children aged 0 to 4, tests for interaction indicated no significant effect modification by any of the three indicators of isolation (tables 5.15 to 5.17). In both isolated and non-isolated areas, according to all three indicators of isolation, most of the rate ratios were near the null value.

Table 5.15. Rate ratios for childhood leukaemia according to percent of the population that are migrants, by urban-rural status for children aged 0 to 4

% MIGRANT POPULATION	RURAL			URBAN		
	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR*† (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
0 to 20	86	1.00 (reference)	1.00 (reference)	241	1.00 (reference)	1.00 (reference)
>20 to 30	83	0.89 (0.66, 1.20)	0.92 (0.67, 1.25)	209	0.91 (0.75, 1.09)	0.92 (0.75, 1.12)
>30	63	0.95 (0.69, 1.32)	1.01 (0.70, 1.44)	84	1.02 (0.80, 1.31)	1.14 (0.86, 1.51)
p (interaction)‡ =0.810						

* RR adjusted for average individual income and the education distribution
† One case missing in adjusted analysis
‡ P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 2 degrees of freedom

Table 5.16. Rate ratios for childhood leukaemia according to percent of the population that are migrants, by distance from nearest urban centre for children aged 0 to 4

% MIGRANT POPULATION	>50 KILOMETRES			≤50 KILOMETRES		
	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR*† (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
0 to 20	21	1.00 (reference)	1.00 (reference)	306	1.00 (reference)	1.00 (reference)
>20 to 30	11	0.66 (0.32, 1.38)	0.67 (0.32, 1.40)	281	0.91 (0.78, 1.07)	0.90 (0.75, 1.07)
>30	6	1.04 (0.42, 2.58)	0.91 (0.34, 2.44)	141	0.96 (0.79, 1.17)	1.00 (0.79, 1.27)
p (interaction)‡ =0.744						

* RR adjusted for average individual income and the education distribution
† One case missing in adjusted analysis
‡ P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 2 degrees of freedom

Table 5.17. Rate ratios for childhood leukaemia according to percent of the population that are migrants, by number of paved highways for children aged 0 to 4

% MIGRANT POPULATION	≤2 PAVED HIGHWAYS			>2 PAVED HIGHWAYS		
	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR*† (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
0 to 20	31	1.00 (reference)	1.00 (reference)	296	1.00 (reference)	1.00 (reference)
>20 to 30	26	0.85 (0.51, 1.44)	0.84 (0.50, 1.43)	266	0.90 (0.77, 1.07)	0.88 (0.73, 1.05)
>30	13	0.92 (0.48, 1.76)	0.83 (0.42, 1.64)	134	0.97 (0.79, 1.18)	0.99 (0.78, 1.26)
p (interaction)‡ =0.889						

* RR adjusted for average individual income and the education distribution
† One case missing in adjusted analysis
‡ P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 2 degrees of freedom

Results for analyses of percent population change in children aged 5 to 9 is presented in tables 5.18 to 5.19. There was no significant effect modification by any of the three indicators of isolation. Risk estimates for percent population change according to levels of all three isolation indicators were not significantly different from the null. There was some indication of a decreased risk in the category of 0% to 10% change in population in areas with two or less highways (RR=0.56, 95% CI 0.24-1.32), and in the highest category of population change in areas greater than 50 kilometres from an urban centre (RR=0.46, 95% CI 0.06-3.67), however, the confidence intervals associated with these estimates were wide and included the null value.

Table 5.18. Rate ratios for childhood leukaemia according to percent change in population, by urban-rural status for children aged 5 to 9

% Δ IN POPULATION	RURAL			URBAN		
	# OF CASES*	CRUDE RR (95% CI)	ADJUSTED RR†‡ (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
≤0	21	1.00 (reference)	1.00 (reference)	25	1.00 (reference)	1.00 (reference)
>0 to 10	37	0.73 (0.45, 1.16)	0.79 (0.49, 1.28)	84	0.87 (0.63, 1.21)	1.00 (0.71, 1.40)
>10 to 20	15	0.90 (0.51, 1.59)	1.14 (0.63, 2.06)	26	0.78 (0.51, 1.19)	1.04 (0.66, 1.64)
>20	17	0.78 (0.44, 1.38)	0.99 (0.53, 1.87)	33	0.79 (0.52, 1.20)	1.06 (0.65, 1.72)
p (interaction)§ =0.755						

* One case missing for analysis of urban-rural status
‡ RR adjusted for average individual income and the education distribution
* One additional case missing in adjusted analysis
§ P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 3 degrees of freedom

Table 5.19. Rate ratios for childhood leukaemia according to percent change in population, by distance from nearest urban centre for children aged 5 to 9

% Δ IN POPULATION	>50 KILOMETRES			≤50 KILOMETRES		
	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR*† (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
≤0	9	1.00 (reference)	1.00 (reference)	70	1.00 (reference)	1.00 (reference)
>0 to 10	7	0.84 (0.31, 2.25)	0.92 (0.34, 2.48)	168	0.87 (0.66, 1.14)	0.95 (0.71, 1.26)
>10 to 20	2	0.85 (0.18, 3.92)	0.82 (0.18, 3.83)	57	0.85 (0.60, 1.20)	1.08 (0.74, 1.58)
>20	2	0.90 (0.20, 4.18)	0.46 (0.06, 3.67)	56	0.79 (0.56, 1.12)	1.02 (0.67, 1.54)
p (interaction)‡ =0.887						

* RR adjusted for average individual income and the education distribution
† One case missing in adjusted analysis
‡ P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 3 degrees of freedom

Table 5.20. Rate ratios for childhood leukaemia according to percent change in population, by number of paved highways for children aged 5 to 9

% Δ IN POPULATION	≤2 PAVED HIGHWAYS			>2 PAVED HIGHWAYS		
	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR*† (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
≤0	15	1.00 (reference)	1.00 (reference)	64	1.00 (reference)	1.00 (reference)
>0 to 10	8	0.52 (0.22, 1.22)	0.56 (0.24, 1.32)	167	0.90 (0.68, 1.20)	0.98 (0.73, 1.32)
>10 to 20	6	0.99 (0.38, 2.56)	1.08 (0.42, 2.78)	53	0.84 (0.59, 1.21)	1.06 (0.71, 1.57)
>20	4	1.16 (0.39, 3.50)	0.92 (0.26, 3.18)	54	0.79 (0.55, 1.13)	1.00 (0.65, 1.52)
p (interaction)‡ =0.623						

* RR adjusted for average individual income and the education distribution
† One case missing in adjusted analysis
‡ P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 3 degrees of freedom

Similar results were observed in the analysis of the effect of migrant population in children aged 5 to 9 (tables 5.21 to 5.23). Tests for interaction were consistent with no effect modification by geographic isolation. As well, rate ratios for the effect of migrant population were not significantly different from the null. In areas greater than 50 kilometres from an urban centre in the highest category of migrant population concentration, a decreased risk was observed, however, the rate ratio was associated with a wide confidence interval (RR=0.43, 95% CI 0.05-3.36).

Table 5.21. Rate ratios for childhood leukaemia according to percent of the population that are migrants, by urban-rural status for children aged 5 to 9

% MIGRANT POPULATION	RURAL			URBAN		
	# OF CASES*	CRUDE RR (95% CI)	ADJUSTED RR†‡ (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR† (95% CI)
0 to 20	43	1.00 (reference)	1.00 (reference)	126	1.00 (reference)	1.00 (reference)
>20 to 30	41	0.87 (0.57, 1.34)	1.01 (0.65, 1.57)	96	0.81 (0.62, 1.05)	0.91 (0.68, 1.21)
>30	25	0.80 (0.49, 1.31)	0.99 (0.58, 1.70)	39	0.88 (0.62, 1.26)	1.06 (0.71, 1.59)
p (interaction)§ =0.849						

* One case missing from analyses of urban-rural status
† RR adjusted for average individual income and the education distribution
‡ One additional case missing in adjusted analysis
§ P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 2 degrees of freedom

Table 5.22. Rate ratios for childhood leukaemia according to percent of the population that are migrants, by distance from nearest urban centre for children aged 5 to 9

% MIGRANT POPULATION	>50 KILOMETRES			≤50 KILOMETRES		
	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR*† (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
0 to 20	11	1.00 (reference)	1.00 (reference)	159	1.00 (reference)	1.00 (reference)
>20 to 30	7	0.81 (0.31, 2.08)	1.00 (0.38, 2.60)	130	0.82 (0.65, 1.03)	0.88 (0.69, 1.13)
>30	2	0.68 (0.15, 3.09)	0.43 (0.05, 3.36)	62	0.81 (0.60, 1.08)	0.94 (0.67, 1.32)
p (interaction)‡ =0.688						

* RR adjusted for average individual income and the education distribution
† One case missing in adjusted analysis
‡ P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 2 degrees of freedom

Table 5.23. Rate ratios for childhood leukaemia according to percent of the population that are migrants, by number of paved highways for children aged 5 to 9

% MIGRANT POPULATION	≤2 PAVED HIGHWAYS			>2 PAVED HIGHWAYS		
	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR*† (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
0 to 20	14	1.00 (reference)	1.00 (reference)	156	1.00 (reference)	1.00 (reference)
>20 to 30	12	0.88 (0.41, 1.91)	1.03 (0.47, 2.24)	125	0.81 (0.64 1.03)	0.88 (0.68, 1.13)
>30	7	1.17 (0.47, 2.90)	1.24 (0.47, 3.30)	57	0.77 (0.57, 1.05)	0.90 (0.63, 1.27)
p (interaction)‡ =0.810						

* RR adjusted for average individual income and the education distribution
† One case missing in adjusted analysis
‡ P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 2 degrees of freedom

Rate ratios for percent change in population in children aged 10 to 14 are presented in tables 5.24 to 5.26. No significant effect modification by geographic isolation was observed. In areas classified as rural, a slightly increased non-significant risk due to percentage population change was observed. As well, elevated risks were observed in urban areas, which appeared to be stronger when adjusted for SES. In urban areas experiencing greater than 20% change in population, a statistically significant adjusted rate ratio of 1.84 (95% CI 1.01-3.34) was observed. In areas greater than 50 kilometres from an urban centre rate ratios were generally surrounded by wide confidence intervals. In areas less than 50 kilometres from an urban centre, an elevated risk was observed in all categories of population change. The risk in areas experiencing greater than 20% change in population was 1.90 (95% CI 1.16-3.12). Rate ratios according to number of paved highways were elevated in both isolated and non-isolated areas, although in the isolated areas the confidence intervals were relatively wide. In the category of greater than 20% population change in areas with more than two paved highways, the observed elevated risk was statistically significant (adjusted RR 1.77, 95% CI 1.08-2.93).

Table 5.24. Rate ratios for childhood leukaemia according to percent change in population, by urban-rural status for children aged 10 to 14

% Δ IN POPULATION	RURAL			URBAN		
	# OF CASES*	CRUDE RR (95% CI)	ADJUSTED RR† (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR† (95% CI)
≤0	21	1.00 (reference)	1.00 (reference)	25	1.00 (reference)	1.00 (reference)
>0 to 10	37	0.98 (0.58, 1.68)	1.05 (0.61, 1.81)	84	1.14 (0.73, 1.78)	1.17 (0.74, 1.84)
>10 to 20	15	1.05 (0.54, 2.04)	1.20 (0.60, 2.40)	26	1.14 (0.66, 1.97)	1.25 (0.69, 2.25)
>20	17	1.13 (0.60, 2.14)	1.41 (0.70, 2.86)	33	1.48 (0.88, 2.49)	1.84 (1.01, 3.34)
p (interaction)‡ =0.932						

* Two cases missing for analysis of urban-rural status
† RR adjusted for average individual income and the education distribution
‡ P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 3 degrees of freedom

Table 5.25. Rate ratios for childhood leukaemia according to percent change in population, by distance from nearest urban centre for children aged 10 to 14

% Δ IN POPULATION	>50 KILOMETRES			≤50 KILOMETRES		
	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
≤0	5	1.00 (reference)	1.00 (reference)	41	1.00 (reference)	1.00 (reference)
>0 to 10	4	0.89 (0.24, 3.30)	0.94 (0.25, 3.51)	117	1.09 (0.76, 1.56)	1.19 (0.82, 1.71)
>10 to 20	2	1.66 (0.32, 8.54)	1.73 (0.33, 8.92)	41	1.13 (0.74, 1.75)	1.41 (0.88, 2.25)
>20	1	0.92 (0.11, 7.90)	1.10 (0.13, 9.42)	49	1.34 (0.89, 2.03)	1.90 (1.16, 3.12)
p (interaction)† =0.919						

* RR adjusted for average individual income and the education distribution
† P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 3 degrees of freedom

Table 5.26. Rate ratios for childhood leukaemia according to percent change in population, by number of paved highways for children aged 10 to 14

% Δ IN POPULATION	≤2 PAVED HIGHWAYS			>2 PAVED HIGHWAYS		
	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
≤0	7	1.00 (reference)	1.00 (reference)	39	1.00 (reference)	1.00 (reference)
>0 to 10	10	1.44 (0.55, 3.77)	1.49 (0.57, 3.92)	111	1.04 (0.72, 1.50)	1.12 (0.77, 1.63)
>10 to 20	6	2.23 (0.75, 6.64)	2.47 (0.82, 7.38)	37	1.06 (0.67, 1.65)	1.27 (0.78, 2.07)
>20	2	1.40 (0.29, 6.75)	1.67 (0.35, 8.09)	48	1.31 (0.86, 1.99)	1.77 (1.08, 2.93)
p (interaction)† = 0.694						

* RR adjusted for average individual income and the education distribution
† P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 3 degrees of freedom

Results relating to migrant population and the risk of childhood leukaemia in children aged 10 to 14 are presented in tables 5.27 to 5.29. Tests for interaction provided some indication of effect modification by urban-rural status (p=0.060). A slight excess risk was observed in rural areas where greater than 30% of the population are migrants (RR=1.32, 95% CI 0.77-2.28). Elevated risks were also observed in all categories of percent migrant population, relative to the category of 0% to 20%. In the highest category of concentration of migrant population, the rate ratio was significant at 1.81 (95% CI 1.13-2.91). There was no significant effect due to concentration of migrant population in areas considered isolated based on distance from an urban centre. However, in areas less than 50

kilometres from an urban centre, the rate ratio associated with greater than 30% of the population being migrants was significantly elevated (RR=1.73, 95% CI 1.18-2.52). Increased rate ratios were found in isolated and non-isolated areas, with respect to number of paved highways. In areas having greater than 30% of the population that are migrants and more than two paved highways, a statistically significant risk estimate of 1.66 (95% CI 1.12-2.45) was observed.

Table 5.27. Rate ratios for childhood leukaemia according to percent of the population that are migrants, by urban-rural status for children aged 10 to 14

% MIGRANT POPULATION	RURAL			URBAN		
	# OF CASES*	CRUDE RR (95% CI)	ADJUSTED RR† (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
0 to 20	37	1.00 (reference)	1.00 (reference)	64	1.00 (reference)	1.00 (reference)
>20 to 30	25	0.63 (0.38, 1.06)	0.65 (0.39, 1.10)	73	1.29 (0.92, 1.80)	1.36 (0.95, 1.95)
>30	28	1.16 (0.71, 1.89)	1.32 (0.77, 2.28)	31	1.51 (0.98, 2.32)	1.81 (1.13, 2.91)
p (interaction)‡ = 0.060						

* Two cases missing from analyses of urban-rural status
† RR adjusted for average individual income and the education distribution
‡ P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 2 degrees of freedom

Table 5.28. Rate ratios for childhood leukaemia according to percent of the population that are migrants, by distance from nearest urban centre for children aged 10 to 14

% MIGRANT POPULATION	>50 KILOMETRES			≤50 KILOMETRES		
	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
0 to 20	6	1.00 (reference)	1.00 (reference)	95	1.00 (reference)	1.00 (reference)
>20 to 30	4	0.74 (0.22, 2.53)	0.84 (0.25, 2.91)	95	1.05 (0.79, 1.40)	1.13 (0.83, 1.53)
>30	1	0.58 (0.07, 4.74)	0.69 (0.08, 5.64)	58	1.39 (1.01, 1.93)	1.73 (1.18, 2.52)
p (interaction)† =0.659						

* RR adjusted for average individual income and the education distribution
† P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 2 degrees of freedom

Table 5.29. Rate ratios for childhood leukaemia according to percent of the population that are migrants, by number of paved highways for children aged 10 to 14

% MIGRANT POPULATION	≤2 PAVED HIGHWAYS			>2 PAVED HIGHWAYS		
	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
0 to 20	7	1.00 (reference)	1.00 (reference)	93	1.00 (reference)	1.00 (reference)
>20 to 30	11	1.31 (0.54, 3.16)	1.46 (0.60, 3.56)	88	1.01 (0.75, 1.35)	1.07 (0.78, 1.46)
>30	5	1.42 (0.48, 4.23)	1.62 (0.54, 4.93)	54	1.35 (0.97, 1.89)	1.66 (1.12, 2.45)
p (interaction)† =0.754						

* RR adjusted for average individual income and the education distribution
† P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 2 degrees of freedom

5.4.3. Results for ALL Among Children Aged 0 to 14

Tables 5.30 to 5.32 present results for the risk of ALL among increasing levels of percentage population change in children included in the whole age group (0 to 14). Tests for interaction provided some indication of effect modification by urban-rural status (p=0.084). Rate ratios for percent population change at all levels of isolation, as indicated by all three isolation variables, were not significantly different from the null except for areas classified as rural. A statistically significant elevated risk was observed for rural areas experiencing between 0% and 10% (RR=1.35, 95% CI 1.01-1.81) and greater than 20% (RR=1.59, 95% CI 1.10-2.29) changes in population.

Table 5.30. Rate ratios for ALL according to percent change in population, by urban-rural status for children aged 0 to 14

% Δ IN POPULATION	RURAL			URBAN		
	# OF CASES*	CRUDE RR (95% CI)	ADJUSTED RR†‡ (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR† (95% CI)
≤0	66	1.00 (reference)	1.00 (reference)	157	1.00 (reference)	1.00 (reference)
>0 to 10	154	1.27 (0.95, 1.70)	1.35 (1.01, 1.81)	408	0.84 (0.70, 1.01)	0.89 (0.74, 1.08)
>10 to 20	58	1.21 (0.85, 1.73)	1.40 (0.97, 2.02)	121	0.80 (0.63, 1.02)	0.93 (0.72, 1.21)
>20	70	1.32 (0.94, 1.85)	1.59 (1.10, 2.29)	124	0.82 (0.65, 1.03)	1.02 (0.78, 1.34)
p (interaction)§ =0.084						

* Two cases missing from analyses of urban-rural status
† RR adjusted for average individual income and the education distribution
‡ One additional case missing in adjusted analysis
§ P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 3 degrees of freedom

Table 5.31. Rate ratios for ALL according to percent change in population, by distance from nearest urban centre for children aged 0 to 14

% Δ IN POPULATION	>50 KILOMETRES			≤50 KILOMETRES		
	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR*† (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
≤0	24	1.00 (reference)	1.00 (reference)	199	1.00 (reference)	1.00 (reference)
>0 to 10	22	0.99 (0.55, 1.76)	1.02 (0.57, 1.81)	540	0.99 (0.84, 1.16)	1.03 (0.87, 1.22)
>10 to 20	7	1.14 (0.49, 2.65)	1.13 (0.49, 2.63)	174	0.94 (0.76, 1.15)	1.07 (0.86, 1.34)
>20	5	0.85 (0.33, 2.24)	0.74 (0.26, 2.13)	189	0.97 (0.80, 1.19)	1.16 (0.92, 1.48)
p (interaction)‡=0.852						

* RR adjusted for average individual income and the education distribution
† One case missing in adjusted analysis
‡ P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 3 degrees of freedom

Table 5.32. Rate ratios for ALL according to percent change in population, by number of paved highways for children aged 0 to 14

% Δ IN POPULATION	≤2 PAVED HIGHWAYS			>2 PAVED HIGHWAYS		
	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR*† (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
≤0	36	1.00 (reference)	1.00 (reference)	187	1.00 (reference)	1.00 (reference)
>0 to 10	41	1.11 (0.71, 1.74)	1.14 (0.73, 1.78)	521	0.97 (0.82, 1.15)	1.01 (0.85, 1.20)
>10 to 20	19	1.33 (0.76, 2.32)	1.37 (0.79, 2.40)	162	0.91 (0.74, 1.12)	1.02 (0.81, 1.29)
>20	10	1.22 (0.61, 2.47)	1.20 (0.57, 2.49)	184	0.95 (0.77, 1.16)	1.11 (0.88, 1.42)
p (interaction)‡=0.810						

* RR adjusted for average individual income and the education distribution
† One case missing in adjusted analysis
‡ P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 3 degrees of freedom

Results of the relationship between concentration of migrant population and ALL in children aged 0 to 14 is presented in tables 5.33 to 5.35. There was no evidence of effect modification by any of the three indicators of geographic isolation. In general, the level of migrant population did not appear to influence the risk of ALL in either isolated or non-isolated areas, with respect to all three indicators of geographic isolation. In areas classified as urban where migrants constituted greater than 30% of the total population, a slightly increased risk of ALL was observed (RR=1.22, 95% CI 0.97-1.53). In areas with two or less paved highways, the rate ratio for areas with greater than 30% of a migrant population was slightly elevated (RR=1.27, 95% CI 0.74-2.17).

Table 5.33. Rate ratios for ALL according to percent of the population that are migrants, by urban-rural status for children aged 0 to 14

% MIGRANT POPULATION	RURAL			URBAN		
	# OF CASES*	CRUDE RR (95% CI)	ADJUSTED RR†‡ (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR† (95% CI)
0 to 20	134	1.00 (reference)	1.00 (reference)	362	1.00 (reference)	1.00 (reference)
>20 to 30	122	0.84 (0.66, 1.08)	0.91 (0.70, 1.17)	321	0.95 (0.82, 1.11)	1.02 (0.87, 1.20)
>30	92	0.96 (0.74, 1.25)	1.11 (0.83, 1.48)	127	1.04 (0.85, 1.27)	1.22 (0.97, 1.53)
p (interaction)§=0.699						

* Two cases missing from analyses of urban-rural status
† RR adjusted for average individual income and the education distribution
‡ One additional case missing in adjusted analysis
§ P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 2 degrees of freedom

Table 5.34. Rate ratios for ALL according to percent of the population that are migrants, by distance from nearest urban centre for children aged 0 to 14

% MIGRANT POPULATION	>50 KILOMETRES			≤50 KILOMETRES		
	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR*† (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
0 to 20	32	1.00 (reference)	1.00 (reference)	465	1.00 (reference)	1.00 (reference)
>20 to 30	18	0.72 (0.40, 1.28)	0.78 (0.44, 1.40)	426	0.93 (0.82, 1.06)	0.95 (0.83, 1.10)
>30	8	0.96 (0.44, 2.08)	0.94 (0.41, 2.15)	211	0.97 (0.83, 1.15)	1.08 (0.90, 1.31)
p (interaction)‡=0.797						

* RR adjusted for average individual income and the education distribution
† One case missing in adjusted analysis
‡ P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 2 degrees of freedom

Table 5.35. Rate ratios for ALL according to percent of the population that are migrants, by number of paved highways for children aged 0 to 14

% MIGRANT POPULATION	≤2 PAVED HIGHWAYS			>2 PAVED HIGHWAYS		
	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR*† (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
0 to 20	42	1.00 (reference)	1.00 (reference)	455	1.00 (reference)	1.00 (reference)
>20 to 30	42	1.04 (0.68, 1.59)	1.10 (0.71, 1.69)	402	0.91 (0.79, 1.04)	0.92 (0.80, 1.07)
>30	22	1.24 (0.74, 2.07)	1.27 (0.74, 2.17)	197	0.95 (0.80, 1.12)	1.04 (0.86, 1.27)
p (interaction)‡ =0.686						

* RR adjusted for average individual income and the education distribution
† One case missing in adjusted analysis
‡ P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 2 degrees of freedom

5.4.4. Results for ALL Among Specific Age Subgroups

In children aged 0 to 4, ALL risk was influenced by population change (tables 5.36 to 5.38). Statistically significant effect modification by urban-rural status was observed (p=0.002). In areas classified as urban, a slightly decreased risk was observed. In rural areas, elevated risks were observed for all levels of population change. Rate ratios were statistically significant for rural areas experiencing between 0% and 10% (RR=1.96, 95% CI 1.27-3.01) and greater than 20% (RR=1.98, 95% CI 1.17-3.33) population change. The overall association between percent population change and ALL in rural areas was statistically significant (p=0.016). Some indication of a gradient of effect was observed (p-value for the trend test=0.058). According to number of paved highways and distance from an urban centre, risk estimates for isolated and non-isolated areas were generally near the null value. There was no significant overall association or trend between percent population change and ALL in isolated areas according to number of paved highways and distance from an urban centre (p-values not presented).

Table 5.36. Rate ratios for ALL according to percent change in population, by urban-rural status for children aged 0 to 4

% Δ IN POPULATION	RURAL			URBAN		
	# OF CASES*	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
≤0	27	1.00 (reference)	1.00 (reference)	94	1.00 (reference)	1.00 (reference)
>0 to 10	96	1.89 (1.23, 2.90)	1.96 (1.26, 3.01)	231	0.76 (0.60, 0.96)	0.78 (0.61, 1.00)
>10 to 20	31	1.51 (0.90, 2.53)	1.63 (0.96, 2.78)	68	0.74 (0.54, 1.01)	0.80 (0.57, 1.12)
>20	42	1.75 (1.08, 2.83)	1.98 (1.17, 3.33)	67	0.71 (0.52, 0.97)	0.83 (0.57, 1.19)
p (interaction)†=0.002						

* RR adjusted for average individual income and the education distribution

† P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 3 degrees of freedom

Table 5.37. Rate ratios for ALL according to percent change in population, by distance from nearest urban centre for children aged 0 to 4

% Δ IN POPULATION	>50 KILOMETRES			≤50 KILOMETRES		
	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
≤0	12	1.00 (reference)	1.00 (reference)	109	1.00 (reference)	1.00 (reference)
>0 to 10	11	0.95 (0.42, 2.16)	0.95 (0.42, 2.15)	316	1.01 (0.81, 1.25)	1.02 (0.82, 1.28)
>10 to 20	4	1.26 (0.41, 3.90)	1.21 (0.39, 3.76)	95	0.90 (0.69, 1.19)	0.95 (0.71, 1.29)
>20	2	0.61 (0.14, 2.74)	0.64 (0.14, 2.88)	107	0.94 (0.72, 1.23)	1.04 (0.76, 1.44)
p (interaction)†=0.868						

* RR adjusted for average individual income and the education distribution
† P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 3 degrees of freedom

Table 5.38. Rate ratios for ALL according to percent change in population, by number of paved highways for children aged 0 to 4

% Δ IN POPULATION	≤2 PAVED HIGHWAYS			>2 PAVED HIGHWAYS		
	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
≤0	17	1.00 (reference)	1.00 (reference)	104	1.00 (reference)	1.00 (reference)
>0 to 10	26	1.45 (0.79, 2.67)	1.44 (0.78, 2.66)	301	0.97 (0.77, 1.21)	0.98 (0.78, 1.23)
>10 to 20	8	1.17 (0.51, 2.71)	1.15 (0.49, 2.67)	91	0.89 (0.67, 1.17)	0.93 (0.68, 1.26)
>20	4	0.94 (0.32, 2.80)	0.99 (0.33, 2.96)	105	0.91 (0.70, 1.20)	1.00 (0.72, 1.38)
p (interaction)†=0.657						

* RR adjusted for average individual income and the education distribution
† P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 3 degrees of freedom

ALL risk in children aged 0 to 4 associated with the concentration of migrant population is presented in tables 5.39 to 5.41. There was no indication of effect modification by any of the three isolation variables. Risk estimates associated with migrant population were generally all near the null value according to all levels of each isolation indicator. In the isolated areas, according to all three indicators, there was no statistically significant overall association or trend (p-values not presented).

Table 5.39. Rate ratios for ALL according to percent of the population that are migrants, by urban-rural status for children aged 0 to 4

% MIGRANT POPULATION	RURAL			URBAN		
	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
0 to 20	73	1.00 (reference)	1.00 (reference)	207	1.00 (reference)	1.00 (reference)
>20 to 30	71	0.89 (0.64, 1.24)	0.92 (0.66, 1.29)	181	0.91 (0.75, 1.11)	0.93 (0.75, 1.16)
>30	52	0.93 (0.65, 1.32)	1.01 (0.68, 1.48)	72	1.02 (0.78, 1.33)	1.15 (0.85, 1.55)
p (interaction)† =0.828						

* RR adjusted for average individual income and the education distribution
† P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 2 degrees of freedom

Table 5.40. Rate ratios for ALL according to percent of the population that are migrants, by distance from nearest urban centre for children aged 0 to 4

% MIGRANT POPULATION	>50 KILOMETRES			≤50 KILOMETRES		
	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
0 to 20	16	1.00 (reference)	1.00 (reference)	264	1.00 (reference)	1.00 (reference)
>20 to 30	8	0.63 (0.27, 1.48)	0.65 (0.28, 1.53)	244	0.92 (0.77, 1.09)	0.91 (0.76, 1.10)
>30	5	1.14 (0.42, 3.11)	1.21 (0.44, 3.34)	119	0.94 (0.76, 1.17)	1.00 (0.77, 1.29)
p (interaction)†=0.613						

* RR adjusted for average individual income and the education distribution
† P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 2 degrees of freedom

Table 5.41. Rate ratios for ALL according to percent of the population that are migrants, by number of paved highways for children aged 0 to 4

% MIGRANT POPULATION	≤2 PAVED HIGHWAYS			>2 PAVED HIGHWAYS		
	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR*† (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
0 to 20	23	1.00 (reference)	1.00 (reference)	257	1.00 (reference)	1.00 (reference)
>20 to 30	22	0.97 (0.54, 1.75)	0.98 (0.54, 1.77)	230	0.90 (0.75, 1.08)	0.89 (0.73, 1.08)
>30	10	0.95 (0.45, 2.00)	0.96 (0.45, 2.04)	114	0.95 (0.76, 1.18)	0.99 (0.76, 1.29)
p (interaction)†=0.923						

* RR adjusted for average individual income and the education distribution
† P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 2 degrees of freedom

Tables 5.42 to 5.44 present the risk for ALL in children aged 5 to 9 according to percent change in population. No significant effect modification was observed for any of the geographic isolation variables. Generally, percent change in population did not appear to influence ALL risk in children aged 5 to 9. Decreased risks were observed in areas with two or less paved highways experiencing between 0% and 10% change in population (RR=0.57, 95% CI 0.23-1.43) and in areas greater than 50 kilometres from an urban centre experiencing greater than 20% change in population (RR=0.50, 95% CI 0.06-4.04), however, both risk estimates were associated with wide confidence intervals.

Table 5.42. Rate ratios for ALL according to percent change in population, by urban-rural status for children aged 5 to 9

% Δ IN POPULATION	RURAL			URBAN		
	# OF CASES*	CRUDE RR (95% CI)	ADJUSTED RR†‡ (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
≤0	24	1.00 (reference)	1.00 (reference)	45	1.00 (reference)	1.00 (reference)
>0 to 10	32	0.73 (0.43, 1.23)	0.81 (0.47, 1.37)	116	0.82 (0.58, 1.16)	0.96 (0.68, 1.37)
>10 to 20	17	0.96 (0.52, 1.79)	1.28 (0.67, 2.43)	37	0.83 (0.54, 1.28)	1.16 (0.73, 1.86)
>20	18	0.92 (0.50, 1.79)	1.25 (0.64, 2.45)	34	0.75 (0.48, 1.17)	1.07 (0.64, 1.78)
p (interaction)§=0.773						

* One case missing for analysis of urban-rural status
† RR adjusted for average individual income and the education distribution
‡ One additional case missing in adjusted analysis
§ P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 3 degrees of freedom

Table 5.43. Rate ratios for ALL according to percent change in population, by distance from nearest urban centre for children aged 5 to 9

% Δ IN POPULATION	>50 KILOMETRES			≤50 KILOMETRES		
	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR*† (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
≤0	8	1.00 (reference)	1.00 (reference)	61	1.00 (reference)	1.00 (reference)
>0 to 10	7	0.94 (0.34, 2.60)	1.04 (0.37, 2.88)	141	0.83 (0.62, 1.13)	0.92 (0.68, 1.25)
>10 to 20	2	0.95 (0.20, 4.49)	0.88 (0.19, 4.21)	53	0.91 (0.63, 1.31)	1.18 (0.79, 1.77)
>20	2	1.02 (0.22, 4.79)	0.50 (0.06, 4.04)	50	0.81 (0.56, 1.18)	1.08 (0.69, 1.68)
p (interaction)‡=0.844						

* RR adjusted for average individual income and the education distribution
† One case missing in adjusted analysis
‡ P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 3 degrees of freedom

Table 5.44. Rate ratios for ALL according to percent change in population, by number of paved highways for children aged 5 to 9

% Δ IN POPULATION	≤2 PAVED HIGHWAYS			>2 PAVED HIGHWAYS		
	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR*† (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
≤0	13	1.00 (reference)	1.00 (reference)	56	1.00 (reference)	1.00 (reference)
>0 to 10	7	0.52 (0.21, 1.31)	0.57 (0.23, 1.43)	141	0.87 (0.64, 1.19)	0.96 (0.70, 1.31)
>10 to 20	6	1.14 (0.43, 3.01)	1.23 (0.47, 3.26)	49	0.89 (0.61, 1.31)	1.15 (0.76, 1.74)
>20	4	1.34 (0.44, 4.11)	1.04 (0.29, 3.67)	48	0.80 (0.54, 1.17)	1.04 (0.67, 1.63)
p (interaction)‡ =0.692						

* RR adjusted for average individual income and the education distribution
† One case missing in adjusted analysis
‡ P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 3 degrees of freedom

Results for the effect of migrant population in children aged 5 to 9 are presented in tables 5.45 to 5.47. Tests for interaction indicated no significant effect modification by geographic isolation. None of the rate ratios were significantly different than the null. In areas with two or less paved roads, elevated risk estimates were observed for increasing concentrations of migrant population, however, confidence intervals were wide. As well, rate ratios were slightly elevated in rural areas.

Table 5.45. Rate ratios for ALL according to percent of the population that are migrants, by urban-rural status for children aged 5 to 9

% MIGRANT POPULATION	RURAL			URBAN		
	# OF CASES*	CRUDE RR (95% CI)	ADJUSTED RR†‡ (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
0 to 20	33	1.00 (reference)	1.00 (reference)	110	1.00 (reference)	1.00 (reference)
>20 to 30	34	0.95 (0.59, 1.53)	1.15 (0.70, 1.87)	88	0.85 (0.64, 1.12)	1.00 (0.74, 1.35)
>30	24	1.00 (0.59, 1.69)	1.36 (0.76, 2.42)	34	0.88 (0.60, 1.29)	1.12 (0.32, 1.72)
p (interaction)§ =0.821						

* One case missing from analyses of urban-rural status
† RR adjusted for average individual income and the education distribution
‡ One additional case missing in adjusted analysis
§ P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 2 degrees of freedom

Table 5.46. Rate ratios for ALL according to percent of the population that are migrants, by distance from nearest urban centre for children aged 5 to 9

% MIGRANT POPULATION	>50 KILOMETRES			≤50 KILOMETRES		
	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR*† (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
0 to 20	10	1.00 (reference)	1.00 (reference)	134	1.00 (reference)	1.00 (reference)
>20 to 30	7	0.89 (0.34, 2.33)	1.14 (0.43, 3.05)	115	0.86 (0.67, 1.10)	0.94 (0.72, 1.23)
>30	2	0.75 (0.16, 3.43)	0.50 (0.06, 3.97)	56	0.87 (0.63, 1.18)	1.04 (0.73, 1.49)
p (interaction)‡=0.677						

* RR adjusted for average individual income and the education distribution
† One case missing in adjusted analysis
‡ P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 2 degrees of freedom

Table 5.47. Rate ratios for ALL according to percent of the population that are migrants, by number of paved highways for children aged 5 to 9

% MIGRANT POPULATION	≤2 PAVED HIGHWAYS			>2 PAVED HIGHWAYS		
	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR*† (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
0 to 20	11	1.00 (reference)	1.00 (reference)	133	1.00 (reference)	1.00 (reference)
>20 to 30	12	1.12 (0.50, 2.54)	1.36 (0.59, 3.14)	110	0.84 (0.65, 1.08)	0.93 (0.71, 1.22)
>30	7	1.49 (0.58, 3.84)	1.69 (0.61, 4.68)	51	0.81 (0.59, 1.12)	0.98 (0.68, 1.42)
p (interaction)‡=0.527						

* RR adjusted for average individual income and the education distribution
† One case missing in adjusted analysis
‡ P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 2 degrees of freedom

ALL risk associated with percent population change in children aged 10 to 14 is presented in tables 5.48 to 5.50. There was no significant interaction between percent change in population and any of the three isolation indicators. In areas with two or less paved highways, an increased risk was observed, particularly in the two highest categories of population change, however these risk estimates were associated with wide confidence intervals. In the non-isolated areas, according to all three indicators of geographic isolation, a percent change in population greater than 20% was associated with a slightly elevated risk of ALL.

Table 5.48. Rate ratios for ALL according to percent change in population, by urban-rural status for children aged 10 to 14

% Δ IN POPULATION	RURAL			URBAN		
	# OF CASES*	CRUDE RR (95% CI)	ADJUSTED RR† (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR† (95% CI)
≤0	15	1.00 (reference)	1.00 (reference)	18	1.00 (reference)	1.00 (reference)
>0 to 10	26	0.97 (0.51, 1.83)	1.01 (0.53, 1.93)	61	1.15 (0.68, 1.95)	1.14 (0.67, 1.96)
>10 to 20	10	0.98 (0.44, 2.18)	1.06 (0.46, 2.46)	16	0.97 (0.50, 1.91)	1.00 (0.49, 2.06)
>20	10	0.93 (0.42, 2.07)	1.10 (0.45, 2.64)	23	1.43 (0.77, 2.65)	1.68 (0.82, 3.42)
p (interaction)‡ =0.810						

* One case missing for analysis of urban-rural status
† RR adjusted for average individual income and the education distribution
‡ P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 3 degrees of freedom

Table 5.49. Rate ratios for ALL according to percent change in population, by distance from nearest urban centre for children aged 10 to 14

% Δ IN POPULATION	>50 KILOMETRES			≤50 KILOMETRES		
	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
≤0	4	1.00 (reference)	1.00 (reference)	29	1.00 (reference)	1.00 (reference)
>0 to 10	4	1.11 (0.28, 4.43)	1.13 (0.28, 4.53)	83	1.09 (0.72, 1.67)	1.13 (0.73, 1.74)
>10 to 20	1	1.04 (0.12, 9.26)	1.04 (0.12, 9.34)	26	1.02 (0.60, 1.73)	1.12 (0.63, 2.00)
>20	1	1.15 (0.13,10.32)	1.29 (0.14,11.60)	32	1.24 (0.75, 2.05)	1.53 (0.84, 2.81)
p (interaction)†=0.999						

* RR adjusted for average individual income and the education distribution
† P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 3 degrees of freedom

Table 5.50. Rate ratios for ALL according to percent change in population, by number of paved highways for children aged 10 to 14

% Δ IN POPULATION	≤2 PAVED HIGHWAYS			>2 PAVED HIGHWAYS		
	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
≤0	6	1.00 (reference)	1.00 (reference)	27	1.00 (reference)	1.00 (reference)
>0 to 10	8	1.34 (0.46, 3.86)	1.32 (0.46, 3.83)	79	1.07 (0.69, 1.66)	1.08 (0.69, 1.70)
>10 to 20	5	2.17 (0.66, 7.11)	2.20 (0.66, 7.29)	22	0.91 (0.52, 1.59)	0.96 (0.52, 1.76)
>20	2	1.64 (0.33, 8.10)	1.83 (0.37, 9.17)	31	1.22 (0.73, 2.04)	1.41 (0.77, 2.60)
p (interaction)†=0.636						

* RR adjusted for average individual income and the education distribution
† P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 3 degrees of freedom

Similar results were observed for ALL risk in children aged 10 to 14 associated with migrant population (tables 5.51 to 5.53). There was some indication of effect modification by urban-rural status (p=0.055). Rate ratios did not differ significantly from the null. Non-significant elevated risks were observed in the non-isolated areas (according to all three isolation indicators) where greater than 30% of the population were migrants.

Table 5.51. Rate ratios for ALL according to percent of the population that are migrants, by urban-rural status for children aged 10 to 14

% MIGRANT POPULATION	RURAL			URBAN		
	# OF CASES*	CRUDE RR (95% CI)	ADJUSTED RR† (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR† (95% CI)
0 to 20	28	1.00 (reference)	1.00 (reference)	45	1.00 (reference)	1.00 (reference)
>20 to 30	17	0.57 (0.31, 1.04)	0.57 (0.31, 1.06)	52	1.30 (0.87, 1.94)	1.34 (0.87, 2.05)
>30	16	0.87 (0.47, 1.61)	0.94 (0.48, 1.85)	21	1.46 (0.87, 2.44)	1.67 (0.94, 2.97)
p (interaction)‡ =0.055						

* One case missing from analyses of urban-rural status
† RR adjusted for average individual income and the education distribution
‡ P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 2 degrees of freedom

Table 5.52. Rate ratios for ALL according to percent of the population that are migrants, by distance from nearest urban centre for children aged 10 to 14

% MIGRANT POPULATION	>50 KILOMETRES			≤50 KILOMETRES		
	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
0 to 20	6	1.00 (reference)	1.00 (reference)	67	1.00 (reference)	1.00 (reference)
>20 to 30	3	0.65 (0.16, 2.59)	0.68 (0.17, 2.74)	67	1.05 (0.75, 1.48)	1.05 (0.73, 1.51)
>30	1	0.68 (0.08, 5.65)	0.71 (0.08, 5.93)	36	1.23 (0.82, 1.84)	1.36 (0.85, 2.17)
p (interaction)† =0.744						

* RR adjusted for average individual income and the education distribution
† P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 2 degrees of freedom

Table 5.53. Rate ratios for ALL according to percent of the population that are migrants, by number of paved highways for children aged 10 to 14

% MIGRANT POPULATION	≤2 PAVED HIGHWAYS			>2 PAVED HIGHWAYS		
	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)	# OF CASES	CRUDE RR (95% CI)	ADJUSTED RR* (95% CI)
0 to 20	8	1.00 (reference)	1.00 (reference)	65	1.00 (reference)	1.00 (reference)
>20 to 30	8	1.07 (0.40, 2.85)	1.09 (0.41, 2.94)	62	1.01 (0.72, 1.44)	0.99 (0.68, 1.44)
>30	5	1.60 (0.52, 4.88)	1.58 (0.50, 4.93)	32	1.15 (0.75, 1.75)	1.22 (0.75, 2.00)
p (interaction)† =0.918						

* RR adjusted for average individual income and the education distribution
† P-value from test of mixing*isolation, adjusted for SES variables, using χ^2 with 2 degrees of freedom

5.5. Results from Secondary Analysis

Following is a presentation of results from the analyses conducted to address the issue of results being influenced by the Toronto Regional Municipality or by possible over-reporting in the Ontario Cancer Registry, and to examine a period effect. The analyses focused on the geographic isolation indicator of urban-rural status, since this was the main variable used to represent isolation in Kinlen's studies and was the only indicator found to interact with population mixing in this study. Also, the analyses focused on the whole age group and also on the 0 to 4 year age group, where elevated risks were observed in the main analysis. Because of the increased rate ratios observed in children aged 10 to 14 in mainly non-isolated areas, the analysis that excluded the Toronto Regional Municipality also was conducted on this age group.

5.5.1. Influence of the Toronto Regional Municipality

In the analysis that excluded the six CSDs that comprise the Toronto Regional Municipality, childhood leukaemia (all subtypes) and ALL in the whole age group and in the specific age groups of 0 to 4 and 10 to 14 were examined. Because the exclusion of the Toronto area affects only the observations in the non-isolated areas, only the rate ratios for these areas are presented. Tables 5.54

to 5.56 present results for urban areas.

Risk of childhood leukaemia in the whole age group did not substantially change when Toronto was excluded. When examining the 0 to 4 year age group, rate ratios changed only slightly. In the 10 to 14 year age group, however, the observed risk increased when Toronto was excluded. In the analysis of ALL (results not shown), little change occurred in the whole age group and in the 0 to 4 year age group. In the 10 to 14 year age group, a slightly increased risk was observed upon exclusion of the Toronto Regional Municipality. Similar results were seen with the other isolation variables and with the variable 'proportion of the population that are migrants' (results not shown).

Table 5.54. Comparison of rate ratios for childhood leukaemia in children aged 0 to 14 in urban areas using the whole data set and excluding Toronto

PERCENT CHANGE IN POPULATION	WHOLE DATA SET			EXCLUDING TORONTO REGIONAL MUNICIPALITY		
	# Cases	Person-years	RR* (95% CI)	# Cases	Person-years	RR* (95% CI)
≤0	181	3,068,975	1.00 (reference)	112	1,826,200	1.00 (reference)
>0 to 10	492	9,543,950	0.92 (0.77, 1.10)	294	5,554,825	0.89 (0.71,1.11)
>10 to 20	141	2,948,175	0.92 (0.73, 1.17)	120	2,466,575	0.88 (0.67,1.15)
>20	149	2,969,075	1.04 (0.81, 1.33)	149	2,969,075	0.96 (0.73,1.27)

* RR adjusted for average individual income and the education distribution

Table 5.55. Comparison of rate ratios for childhood leukaemia in children aged 0 to 4 in urban areas using the whole data set and excluding Toronto

PERCENT CHANGE IN POPULATION	WHOLE DATA SET			EXCLUDING TORONTO REGIONAL MUNICIPALITY		
	# Cases	Person-years	RR* (95% CI)	# Cases	Person-years	RR* (95% CI)
≤0	107	1,009,350	1.00 (reference)	67	588,150	1.00 (reference)
>0 to 10	273	3,274,775	0.80 (0.64,1.01)	175	1,874,050	0.82 (0.62,1.09)
>10 to 20	77	987,800	0.78 (0.56,1.07)	63	834,650	0.66 (0.46,0.96)
>20	77	1,016,025	0.81 (0.58,1.14)	77	1,016,025	0.69 (0.48,1.01)

* RR adjusted for average individual income and the education distribution

Table 5.56. Comparison of rate ratios for childhood leukaemia in children aged 10 to 14 in urban areas using the whole data set and excluding Toronto

PERCENT CHANGE IN POPULATION	WHOLE DATA SET			EXCLUDING TORONTO REGIONAL MUNICIPALITY		
	# Cases	Person-years	RR* (95% CI)	# Cases	Person-years	RR* (95% CI)
≤0	25	1,075,275	1.00 (reference)	13	646,550	1.00 (reference)
>0 to 10	84	3,169,125	1.17 (0.74,1.84)	47	1,856,150	1.26 (0.68, 2.35)
>10 to 20	26	982,075	1.25 (0.69,2.25)	24	808,425	1.77 (0.87, 3.59)
>20	33	960,275	1.84 (1.01,3.34)	33	960,275	2.35 (1.15, 4.78)

*RR adjusted for average individual income and the education distribution

5.5.2. Influence of Hospital-only Registered Cases

Possible over-reporting of leukaemia cases in the OCR was addressed by analyzing data that excluded cases that were registered based on hospital records only. Results of these analyses were not different from results that included these cases. Rate ratios for childhood leukaemia and for ALL in children aged 0 to 4, by urban-rural status, are presented in tables 5.57 and 5.58. P-values for interaction between urban-rural status and percent change in population did not substantially differ between the two data sets (whole data set: p-value for interaction=0.016 for childhood leukaemia and 0.002 for ALL; excluding hospital only diagnosed cases: p-value for interaction=0.010 for childhood leukaemia and 0.002 for ALL). The exclusion of hospital-only registered cases did not affect rate ratios for childhood leukaemia or ALL substantially. Similarly, rate ratios were not affected according to the other geographic isolation indicators or with the 'proportion of the population that are migrants' (results not shown).

Table 5.57. Comparison of rate ratios for childhood leukaemia in children aged 0 to 4 in urban and rural areas using the whole data set and excluding hospital-only registered cases

PERCENT CHANGE IN POPULATION	RR* (95% CI): RURAL		RR* (95% CI): URBAN	
	WHOLE DATA SET	EXCLUDING HOSPITAL-ONLY CASES	WHOLE DATA SET	EXCLUDING HOSPITAL-ONLY CASES
≤0	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
>0 to 10	1.60 (1.10,2.33)	1.62 (1.11,2,38)	0.80 (0.64,1.01)	0.79 (0.62,0.99)
>10 to 20	1.38 (0.86,2.22)	1.32 (0.81, 2.15)	0.78 (0.56,1.07)	0.77 (0.56,1.06)
>20	1.58 (0.99,2.51)	1.62 (1.01,2,60)	0.81 (0.58,1.14)	0.79 (0.56,1.12)

* RR adjusted for average individual income and the education distribution

Table 5.58. Comparison of rate ratios for ALL in children aged 0 to 4 in urban and rural areas using the whole data set and excluding hospital-only registered cases

PERCENT CHANGE IN POPULATION	RR* (95% CI): RURAL		RR* (95% CI): URBAN	
	WHOLE DATA SET	EXCLUDING HOSPITAL-ONLY CASES	WHOLE DATA SET	EXCLUDING HOSPITAL-ONLY CASES
≤0	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
>0 to 10	1.96 (1.27,3.01)	1.95 (1.26,3.00)	0.78 (0.61,1.00)	0.77 (0.60,0.99)
>10 to 20	1.63 (0.96,2.78)	1.50 (0.87, 2.58)	0.80 (0.57,1.12)	0.80 (0.57,1.12)
>20	1.98 (1.17,3.33)	1.97 (1.17,3.32)	0.83 (0.57,1.19)	0.82 (0.57,1.18)

* RR adjusted for average individual income and the education distribution

5.5.3. Examination of a Period Effect

Analyses of percent change in population by urban-rural status for children aged 0 to 4 were conducted on each 5-year period separately (tables 5.59 and 5.60). In all three 5-year periods, there was statistically significant effect modification by urban-rural status (p-values for 1978-82, 1983-87 and 1988-92 were 0.024, 0.032, and 0.021, respectively). In rural areas, the risk due to percent population change appeared to decline over time. The observed trend and risk estimates in rural areas held when hospital-only registered cases were removed (results not shown). In urban areas, there was no apparent change in risk estimates according to 5-year period.

Table 5.59. Comparison of rate ratios for ALL in children aged 0 to 4 in rural areas among the three 5-year periods

PERCENT CHANGE IN POPULATION	RATE RATIOS (95% CI)*		
	1978-82	1983-87	1988-92
≤0	1.00 (reference)	1.00 (reference)	1.00 (reference)
>0 to 10	2.45 (1.06, 5.67)	1.23 (0.68, 2.23)	3.09 (0.74, 12.87)
>10 to 20	2.63 (0.87, 7.91)	0.70 (0.25, 1.93)	2.18 (0.50, 9.49)
>20	3.68 (1.38, 9.81)	2.52 (1.17, 5.43)	1.45 (0.32, 6.71)

*RR adjusted for average individual income and the education distribution

Table 5.60. Comparison of rate ratios for ALL in children aged 0 to 4 in urban areas among the three 5-year periods

PERCENT CHANGE IN POPULATION	RATE RATIOS (95% CI)*		
	1978-82	1983-87	1988-92
≤0	1.00 (reference)	1.00 (reference)	1.00 (reference)
>0 to 10	0.85 (0.56, 1.27)	0.75 (0.49, 1.15)	0.69 (0.34, 1.40)
>10 to 20	1.20 (0.67, 2.12)	1.06 (0.60, 1.90)	0.48 (0.22, 1.08)
>20	0.69 (0.33, 1.45)	0.79 (0.39, 1.60)	0.85 (0.38, 1.88)

*RR adjusted for average individual income and the education distribution

6. DISCUSSION

6.1. Summary of Objectives and Main Analysis

An association between population mixing and childhood leukaemia has been observed in several ecologic studies conducted in the UK.[40,47,49,55] In each, study areas which were relatively extreme in their combined attributes of population mixing and geographic isolation were contrasted. Only one of the three studies conducted outside of the UK was consistent with UK results,[18] and likewise, this was observed in a situation of extreme population growth in a very isolated setting. In this study, the objective was to examine the effect of variations in the level of population mixing on the risk of childhood leukaemia in a study area that experienced relatively moderate levels of population mixing. A large study area (i.e. the province of Ontario), which also varied in levels of geographic isolation, was investigated. Since isolation is a key component of the population mixing hypothesis and is a factor upon which study areas have been restricted to in past research, differences in risk according to level of isolation were determined.

Leukaemia risk was investigated among the entire childhood population (ages 0 to 14) and specific age groups. The inclusion of interaction terms between population mixing and geographic isolation provided risk estimates for increasing levels of population mixing according to two levels of geographic isolation. Based on past studies which implicated an association between SES and childhood leukaemia, and the fact that SES is conceivably associated with population mixing, measures of SES that were found to be associated with childhood leukaemia incidence were controlled for in the analysis.

The consistent increased risk due to population mixing found in past studies has mainly been a reflection of risk in specific subgroups, particularly children aged 0 to 4 and specifically related to the ALL subtype.[18,40] It has therefore been postulated that the etiology of ALL in the youngest age group may differ from that for older children.[46] In regard to this hypothesis, leukaemia risk was

examined in subgroups of the study population defined by age (0-4, 5-9, 10-14) and leukaemia subtype (all subtypes combined and ALL separately).

6.2. Summary of Key Findings

Kinlen's original hypothesis postulates that a population influx into a geographically isolated (rural) area results in an increased risk of leukaemia.[13] Past studies point to this association being most consistent for ALL among young children.[18,49] In areas classified as rural in this study, results according to percentage change in population were suggestive of an increase in risk for childhood leukaemia among the whole age group. When this relationship was examined among specific groups of the study population, it was found that this relationship was strongest among the group for which the hypothesis is most relevant based on past studies, i.e. ALL in young children (adjusted RR=1.98, 95% CI 1.17-3.33 for greater than 20% population change in rural areas).

However, when isolation was represented by number of paved highways (2 or less) and distance from an urban centre (>50 kilometres), there was no increased risk of leukaemia (all subtypes or ALL) related to percentage change in population among 0 to 4 year old children. Also, when population mixing was represented as the percent of the population that are migrants, no association was observed for ALL or all leukaemias for children aged 0 to 4.

In contrast to what was observed among 0 to 4 year old children, increased risks of leukaemia (all subtypes or ALL) due to percentage change in population were not observed for children aged 5 to 9 or 10 to 14 in rural areas. Similarly, no association was observed according to percent of the population that are migrants. As well, population mixing was not found to be associated with leukaemia risk according to other measures of isolation for these age groups.

In non-isolated areas, there was no significant association observed between leukaemia (all subtypes and ALL separately) and either of the population mixing indicators, among children aged 0

to 4. Similarly, there was no association observed in children aged 5 to 9. In children aged 10 to 14, a significantly increased risk of leukaemia of all subtypes was observed in the highest categories of both population mixing variables in areas considered to be non-isolated according to all three geographic isolation indicators (adjusted RR=1.84, 95% CI 1.01-3.34, for greater than 20% change in population in urban areas). When the same analysis was restricted to ALL cases, risk estimates associated with high levels of population mixing were diminished and no longer statistically significant.

Examination of ALL among 0 to 4 year old children by period showed that the effect due to percent population change observed in rural areas was found to be strongest in the first of the three 5-year periods (1978-82). In the following two 5-year periods, rate ratios progressively declined. Removal of cases that were registered exclusively on hospital records did not affect the risk estimates by 5-year period.

6.3. Main Indicator Variables of the Kinlen Hypothesis

6.3.1. Geographic Isolation

Three different indicators of geographic isolation were analyzed for their effect in modifying the relationship between population mixing and the incidence of childhood leukaemia, yet only one (urban-rural status) was found to interact with the effect of population mixing. The distribution of the population at risk by categories of each isolation indicator was different for each variable (see table 5.2, chapter 5) and resulted in very small populations in the isolated categories of the two variables 'number of paved roads' and 'proximity to an urban centre'. Although a lack of statistical power was a factor in the isolated areas, based on these two variables, risk estimates were rarely similar to that observed in rural areas and sometimes rate ratios were in the opposite direction.

In past studies of childhood leukaemia which have considered geographic isolation,[18,46,52,53,58,59,91,92] urban-rural status has most often been employed as an indicator of isolation.

Therefore, it is informative to compare the two additional measures of isolation used in this study with this more traditionally used indicator. According to the variable 'number of paved highways', it was found that 44% of the CSDs that were classified as urban were categorized as isolated (and 51% of rural CSDs were categorized as non-isolated). With respect to the variable 'distance from an urban centre', 10% of urban CSDs were classified as isolated (and 68% of rural CSDs were considered non-isolated). Figures 5.1 to 5.6 (chapter 5) illustrate that isolated and non-isolated areas basically correspond to northern and southern Ontario, respectively, according to 'distance from an urban centre' and 'number of paved roads', while according to urban-rural status, many isolated areas also located in southern Ontario. Of note is that many of the very small CSDs considered urban in southern Ontario are classified as isolated, particularly with the variable 'number of paved highways' (see figures 5.1, 5.3, and 5.5).

Regarding the population mixing hypothesis and the conditions for transmission of infections, it is likely that the initial population density of an area is an important factor in establishing herd immunity. The definition of urban-rural status was based on population density and therefore accounted for this factor. On the other hand, 'number of paved highways' and 'distance from an urban centre' did not directly account for initial population density. Thus, the absence of effect modification by 'number of paved highways' and the lack of similarity between risk estimates from rural areas and from areas with two or less highways is probably due to misclassification of the study population according to initial population density. The effect of misclassification is probably not as high with the variable 'distance from an urban centre', and it is more likely a lack of statistical power that is affecting results according to this variable. Nonetheless, assuming that initial population density is of importance in establishing herd immunity, urban-rural status provided the best measure of geographic isolation.

6.3.2. Population Mixing

Population mixing was measured on the basis of two variables, which have been used in a few previous studies.[18,47,49,55] The percentage change in population provided an indication of whether a CSD was experiencing decreasing, stable or increasing populations. Any net increase in population would contribute to an increase in population density thereby increasing the potential for contact between people (susceptible and infected), and thus facilitating the transmission of infections.[14] The proportion of the population that are migrants allowed for the examination of population mixing where a high population growth was balanced by a high level of out-migration, a situation that would not be captured in the analysis of percentage population change. The rationale for employing this measure of population mixing was that although an area may not experience a net increase in total population (or population density), a large proportion of migrants could nonetheless contribute to exposure (by the original population) to new infectious agents to which the population do not have an established herd immunity.

In this study, rural areas with a large proportion of migrants did not experience an increased risk of leukaemia. If the incomers were from other rural areas and therefore not previously exposed to the relevant infectious agents, then a high concentration of migrants in rural areas would not be expected to result in an increased risk of childhood leukaemia. Unfortunately, information is not available on the migrants' specific origins (relating to urban versus rural areas). However, given that areas that experienced a high change in population would presumably also have a high proportion of migrants, it is probably not the case that migrants to rural areas were predominantly individuals that were previously unexposed (and therefore not infected or carriers of the relevant infection).

In rural areas, it was found that the correlation between the two measures of population mixing (at the CSD level) was not very high ($r=0.32$) thus indicating that these two indicators were not measuring similar aspects of population mixing. Although it is likely that areas that experienced

a net positive change in population would also have a high proportion of migrants, it does not similarly follow that areas with a high proportion of migrants have experienced a high change in population, since in-migration has the potential to be balanced by out-migration. It is even possible that areas with a relatively high proportion of migrants experienced negative population growth. An increase in population density can be expected to lead to an increased potential for contact between individuals in an area. It therefore seems reasonable that the percentage change in population, which also captures changes in population density, provides a better measure of population mixing than the concentration of the migrant population.

In contrast, the elevated risks observed for 10 to 14 year old children in urban areas were apparent when examining both percentage population change and proportion of the population that are migrants. In fact, results for the two population mixing indicators in urban areas among the other age groups were also similar. In urban areas, it was found that the correlation between the two population mixing indicators was 0.65. Thus, in contrast to rural areas, the concentration of migrant population appears to provide a measure of population mixing that is similar to the percentage change in population in urban areas.

6.4. Methodological Considerations

6.4.1. Ecologic Fallacy

The major methodological problem affecting ecologic studies is the potential for ecologic bias, also referred to as the 'ecologic fallacy',[83] which is the failure of estimates of effect at the ecologic level to reflect the effect at the individual level.[87] The ecologic fallacy is an issue when testing etiological hypotheses by using aggregate measures of exposure to predict rates of disease, with the intent of extrapolating effect estimates to the individual.[76] The failure of ecologic estimates of effect to represent individual level effects arises from three sources: within-group bias, confounding by

group, and effect modification by group.[87]

Within-group bias refers to biases, confounding and misclassification that occur at the individual level within each ecologic unit. Since there is no possibility of obtaining information on confounders at the individual level, it is not possible to account for this effect. Confounding by group, also referred to as specification bias, occurs when a factor that is associated with the outcome at the individual level is differentially distributed across ecologic units.[83] In other words, specification bias occurs if this factor is an ecologic confounder, regardless of whether it is a confounder at the individual level. In this study, the potential risk factor of SES was accounted for in order to minimize this bias. Also, exposure to nuclear facilities was considered for its possible confounding effect. Unfortunately, because little is known about the etiology of childhood leukaemia, it was not possible to control for other unknown potential ecologic confounders. Effect modification by group can occur when risk estimates due to the exposure of interest varies across certain groups. In this study, herd immunity measured by geographic isolation, was suspected of modifying risk due to population mixing. Therefore, risk due to population mixing was examined among CSDs grouped according to isolation status. Although SES and geographic isolation were accounted for, the possibility exists that rate ratios are biased due to unknown risk factors or effect modifiers that were differentially distributed across CSDs.

6.4.2. Confounding

Confounder control is also problematic in ecologic studies, since ecologic measures do not necessarily reflect individual measures.[87] To minimize biases due to this misspecification of confounders,[76] the covariate distribution of highest level of education was analyzed. Ideally, income would have been analyzed similarly, however, adequate data were not available. In any case, inadequate confounder control is most likely due to the lack of knowledge regarding risk factors for

childhood leukaemia, and therefore, an inability to control for all possible confounding effects.

6.4.3. Population Denominators

Person-years at risk were calculated by multiplying the population counts in the census of which each 5-year period was centered on by five. If a CSD experienced a relatively high net increase in population during a 5-year period, then denominators may have been overestimated by using a population estimate following a period of population growth to calculate person-years at risk. The effect of this would be to underestimate the incidence rate. Likewise, in areas where population growth was negative, denominators may have been underestimated through the use of a population estimate following a period of population decline. In this case, incidence rates in the referent category may have been overestimated. The overall effect of this would be to decrease observed rate ratios. Thus, assuming that population counts from the census were adequately estimated, elevated risks observed in this study are not likely a result of bias due to inaccurate calculations of population at risk. On the other hand, null risk estimates may have subsequently resulted due to this effect.

6.5. Comparison with Previous Studies

A summary of the literature that is discussed in this section is presented in table 6.1. Similar to many previous studies, this study showed an increased risk of childhood leukaemia, specifically of the ALL subtype in children aged 0 to 4, due to population mixing in rural areas. Indicators of population mixing and methods to estimate relative risks have varied among past studies. In Kinlen's studies, standardized mortality ratios (SMR) were calculated to estimate risk due to population mixing.[40] Rate ratios for the highest categories of population mixing in this study were within the range of the SMRs associated with the highest categories of exposure to population mixing in each of the 5 original Kinlen studies (1.47 to 2.79 for the ages 0 to 14). However, a similar level of risk was

also observed in the lowest category of population mixing in this study. Because different measures of population mixing were used in Kinlen's 5 studies, it is difficult to make comparisons across studies. However, in the study of New Towns,[52] percent population change was measured and it was found that the increased leukaemia mortality among children aged 0 to 4 was observed in areas that experienced greater than 150% change in population. Rural areas experiencing less than this level of population change were not studied.

The results of this study were consistent with the three UK studies which, similar to this study, covered a broad region with more variation in population mixing and geographic isolation compared to Kinlen's studies.[47,49,55] Rate ratios from this study, however, were slightly higher, particularly for ALL. In the study by Langford,[55] relative risks were not determined for different levels of isolation, therefore, it is possible that risk estimates were obscured by differences in risk by isolation. In that study, it was also found that the effect of population mixing on childhood leukaemia mortality among local authority districts of England and Wales occurred at a threshold value between 30% and 50% population change. In this study, leukaemia risk was elevated even between 0% and 10% change. Stiller and Boyle[47] attempted to examine threshold effects by comparing the top decile of ecologic units with the bottom decile. Similar to the results of this study, it was concluded that the association between childhood ALL and population mixing is not restricted to extreme levels of mixing. In another study,[49] an association between childhood leukaemia and population mixing was found only when examining the risk of lymphocytic leukaemia in rural areas of high SES. Results of this study are consistent in that the strongest associations were found for risk of ALL in rural areas.

The results of this study are also consistent with the one study conducted outside of the UK which supported Kinlen's hypothesis.[18] In that study, the most elevated (although non-significant) risks were seen in the youngest age groups. As well, risk estimates were highest when the analysis was restricted to ALL and common ALL.

Table 6.1. Summary of previous literature discussed in section 6.5

Title, Author, Year, Reference no.	Description	Leukaemia Outcome	Population Mixing Indicator	Main Results
Epidemiological evidence for an infective basis in childhood leukaemia Kinlen, LJ, 1995, 40	Review article including Kinlen's five original studies	Generally, childhood leukaemia, all subtypes among children aged 0 to 14	Formation of New Towns, military servicemen, migrant construction workers, situations of commuting, wartime evacuations	SMRs varied from 1.47 to 2.79 in the highest categories of population mixing for the ages 0 to 14
Childhood leukaemia mortality and population change in England and Wales 1969-73 Langford, I, 1993, 55	Study examining all 1365 local authority areas of England and Wales	Childhood leukaemia mortality among children aged 0 to 14 (from 1969-73)	Percent population change between 1961 and 1971	RR=1.408 for childhood leukaemia among children aged 0 to 14, for areas experiencing >50% population increase relative to all other areas
Socio-economic factors in relation to childhood leukaemia and non-Hodgkin lymphomas: an analysis based on small area statistics for census tracts Rodrigues, L et al., 1991, 49	Study examining all census tracts in England and Wales	Childhood leukaemia incidence among children aged 0 to 14 (1966-76 and 1977-83)	Percentage population ratio (ratio of child population at 1981 census to that at 1971 census)	No apparent association between childhood leukaemia and population ratio, except for census tracts with a population ratio of 140% or more in rural areas of the 3 highest SES quintiles (RR=1.4 for ALL)
Effect of population mixing and socioeconomic status in England and Wales, 1979-85, on lymphoblastic leukaemia in children Stiller, CA, et al., 1996, 47	Study examining all counties of England and Wales	ALL in children aged 0 to 14 (from 1979-85)	Proportion of total migrants and child migrants	Significant trends in ALL incidence at ages 0 to 4 and 5 to 9 with child migrants
Clustering of childhood leukaemia in Hong Kong: association with the childhood peak and common acute lymphoblastic leukaemia and with population mixing Alexander FE et al., 1997, 18	Study examining new towns in New Territories in Hong Kong	Childhood leukaemia, ALL, and common ALL in children aged 0 to 14 (1984-90)	Population growth in previously uninhabited areas	Non-significant increased risk of ALL in young age groups associated with population growth

6.6. Interpretation of Results

The excess risk of leukaemia in children was observed in areas considered rural and that experienced some amount of population growth. Transmission of infections and the progress of an epidemic require a relatively large population of susceptibles and a high rate of contact between infected and susceptible people.[14] A large proportion of susceptibles, or in other words a low herd immunity, is expected to be found in isolated areas, therefore it is not surprising that increased risks were observed primarily in rural areas. These areas were also characterized by having a net increase in population. Thus, not only did the population grow but the population density would have increased thereby resulting in an increased likelihood of contact between individuals. Assuming the incomers were either infected or were carriers, population growth in these isolated areas provided an environment conducive to the spread of infection. The fact that the associations were not replicated in urban areas further supports the Kinlen hypothesis since the initially high population density in these areas would have facilitated contact between infected and susceptible people, rendering a high level of herd immunity. The lack of an association observed in rural areas according to the population mixing indicator of migrant population is likely due to the fact that increases in population (and population density) are not strongly correlated with percent of the population that are migrants.

The increased leukaemia risk in rural areas was found to be strongest for the ALL subtype. This result provides support to the suggestion that the etiology of ALL differs from that of other subtypes.[18,46] Small numbers preclude subanalyses by other leukaemia subtypes, however, the observed weakening of risk estimates when all leukaemia subtypes were analyzed is consistent with other subtypes being unrelated (or inversely related) to population mixing in rural areas. Differences in the descriptive epidemiology of ALL and AML[16,17] as well as suggestive evidence that risk factors for ALL and AML may differ[37] also support the fact that the etiology of ALL may differ from that of AML and other leukaemia subtypes.

Accordingly, an immunological model specific to common ALL and related to Kinlen's population mixing hypothesis has been proposed by Greaves.[93] It is postulated that common ALL requires two sequential genetic events (i.e. spontaneous mutations) for its clinical manifestation.[94] Because B cell precursors have a high proliferation rate during early developmental stages and at infancy, they are at a high risk of spontaneous mutation.[95] Thus, the first spontaneous mutation likely occurs *in utero* and results in a pre-leukaemic subclinical state of acute leukaemia. A second event in a cell belonging to the existing pre-leukaemic clone is required to precipitate overt leukaemia. This second event is thought to be induced by an abnormal pattern of infection experienced by the individual due to immunological isolation (i.e. delayed exposure). Normal or early exposure to common infectious agents is expected to occur around the time of birth through mothers or postnatally through breast feeding and contacts with others, such as older siblings.[19] This first exposure establishes immune protection (derived from transplacental maternal antibodies or through breast milk) and will modulate the infant's immunological system. When subsequent exposure occurs, immunocompetent B-lymphocytes will be released. However, when first exposure is delayed, the immune system is unmodulated. When first exposure occurs at an older age, stress is put on B cell precursors. The proliferative stress thus puts the pre-leukaemic clone at risk for the second spontaneous mutation which results in overt leukaemia.

According to this hypothesis, delayed exposure occurs due to a lack of exposure in infancy. Upon later exposure, on rare instances, rather than responding in the usual manner, leukaemia (i.e. common ALL) results. Since young children would have had the least opportunity to have been infected, the proportion of susceptibles would be higher in this age group.[40] Thus, the observed excess risk of ALL among children in the youngest age group is expected. At older ages, more children would have had the opportunity for exposure and subsequently would have been accorded some form of immunity (if leukaemia did not occur). The relative proportion of susceptible children would

therefore be smaller in older age groups compared to younger children and the risk associated with population mixing would be expected to be lower. In fact, a lower non-significant risk of ALL due to population mixing in children aged 5 to 9 was observed. Likewise, there was little risk of ALL associated with population mixing in rural areas for children aged 10 to 14.

In children aged 0 to 4, the risk of ALL due to population mixing in rural areas was found to be strongest in the first of the three 5-year periods included in the study period. Rate ratios declined in the next two 5-year periods. This observation is also consistent with the population mixing hypothesis and was observed by Kinlen.[13,52] In the way that age offers individual immunity (due to increased opportunities for exposure), time following a period of population mixing can offer herd immunity to a geographic area. Assuming that the increase in the incidence of leukaemia was due to infectious transmission through population mixing, then for those individuals that did not respond by developing leukaemia, immunization was established. One would expect that newly born children should be at risk since they have not been directly exposed, however, according to Greaves' hypothesis, a newborn's immunity will reflect that of their mothers.[19] Since herd immunity is a characteristic of the group and is dependent upon individual immunity, as more individuals become immune, the herd immunity will increase. Thus, over time, the herd immunity of Ontario should be increasing. Population growth has occurred in both urban and rural areas during all three time periods. In fact, a larger proportion of CSDs were categorized in the highest levels of population change in both urban and rural areas in the last 5-year period, compared to the first two 5-year periods. On this observation alone one would expect a higher risk of leukaemia in rural areas in this period. However, given that population mixing and an associated increased incidence of leukaemia occurred in the earlier period, it is possible that the individuals that did not get leukaemia, either responded in the usual manner to the infection and developed immunity, or went through a process of symptomless immunization.[14] In effect, the concentration of susceptibles would have decreased in rural areas during the study period.

The increased risk of leukaemia in children aged 10 to 14 associated with high levels of population mixing in urban areas was not expected according to Kinlen's hypothesis. In urban areas, the herd immunity to common infectious agents is expected to be high. As well, the majority of children in this age group are expected to be immune by virtue of their age and having had many opportunities for exposure to common infections. If the cause of leukaemia in rural areas is due to transmission of a common infection during population mixing, it seems unlikely that a common infection is causing leukaemia in 10 to 14 year old children in urban areas.

Whereas increased risks of leukaemia in rural areas was seen at all levels of population change, relative to zero or negative population change, the effect in urban areas among older children seemed to be concentrated in the highest categories of population mixing. One hypothesis for this unexpected finding is that some factor associated with high levels of population mixing in urban areas is causing this observed increased risk. The adjustment of rate ratios by income and education was done to exclude SES as an alternative explanation.

A number of environmental factors are suspected in the etiology of childhood leukaemia,[25,29] however, since the etiology of childhood leukaemia is largely unknown, it is difficult to measure and control for unknown confounding factors. If any leukaemogenic factor is associated with population mixing in urban areas, then it is possible that this factor is resulting in leukaemia in the 10 to 14 year old age group. In one study, population density in urban cities was found to be associated with leukaemia and non-Hodgkin's lymphoma in children aged 0 to 14.[44] The effect was found to be stronger in the youngest age group (0 to 4), therefore, the association was attributed to the infectious hypothesis. Since a change in population would also change the population density of an area, an environmental (leukaemogenic) factor associated with density, other than a common infectious agent, could account for the effect. In a study conducted in Denmark,[96] ambient air levels of volatile organic compounds associated with traffic exhaust fumes, including benzene which is known to be

leukaemogenic in adults,[97] was found to be higher in urban areas (where population density is high) than in rural areas. Outdoor measurements were, however, not found to be comparable to personal exposures. Nonetheless, if high levels of population change in urban areas, and the resulting increase in population density, contributes to an increase in the levels of certain environmental leukaemogenic factors, then it is possible that the increased concentration of these factors are related to an increased risk of leukaemia in children aged 10 to 14. If cumulative exposures are relevant, then this would account for the lack of relationship observed in the younger age groups.

In evaluating the influence of the Toronto Regional Municipality, the CSDs that comprise this area were excluded in a secondary analysis. Rate ratios in urban areas were found to increase. The exclusion of this region had the effect of excluding cases and population at risk primarily from the two lowest categories of population change. The population densities of the CSDs that comprise the Toronto Regional Municipality are among the ten highest in Ontario in all three censuses covered in the study period. If an environmental factor related to population density is associated with leukaemia in children, then it is possible that in areas with a very high population density, the levels are invariably high and consistently contributing to cumulative exposure. Thus, the rates of leukaemia are high. The exclusion of the areas with very high population density that experienced negative to zero population change, decreases the rates in the referent category, thereby increase rate ratios in the highest categories.

If such an environmental exposure associated with high population density is in fact resulting in leukaemia among older children in urban areas, then the absence of a relationship in rural areas (i.e. less population dense) can be accounted for by the fact that concentrations of these environmental pollutants do not achieve as high concentrations. High levels of population mixing, even if it does increase environmental levels in rural areas, does not contribute to a large cumulative exposure among children.

Given that there was no a priori expectation that an increased risk of leukaemia would be observed among children in this age group, this interpretation is presented only as a hypothesis. In any case, the evidence regarding childhood leukaemia in 10 to 14 year old children supports the suggestion that the etiology of childhood leukaemia not only differs by subtype, but also by age.[46]

6.7. Strengths of the Study

Approximately 29,000,000 person-years were included as part of the study population at risk. This gave rise to approximately 1400 cases during the 15 year study period. The relatively large population size included in the study allowed for analyses to have adequate statistical power. The use of existing databases (i.e. the OCR) to obtain information on leukaemia cases allowed for the examination of this large population in a timely manner. As well, the methods used by the OCR ensured high quality data were analyzed in this study.[88]

Population mixing can be considered an environmental factor, representing contact between susceptible individuals and infected people or carriers. To measure an individual's exposure to population mixing would require measuring a susceptible individual's actual exposure to infected individuals. However, it is unlikely that one would be able to identify all the individuals they were in contact with and, in addition, identify who were infected. In fact, this exposure measurement would be impossible since the actual infectious agent that is potentially responsible for causing childhood leukaemia is unknown. Thus, in light of the limitations of exposure data at the individual level, an ecologic study offered a feasible approach to measurement of exposure to this unknown pathogenic agent through the ecologic measure of population mixing.

6.8. Suggestions for Further Research

The results of this study are consistent with findings from many UK studies and one study

conducted in Hong Kong about the relationship between population mixing and childhood leukaemia in isolated areas. All together, results from these studies provide support to the hypothesis regarding an infectious component to the etiology of childhood leukaemia.

All these studies, however, were ecologic studies, and therefore subject to the biases discussed in section 6.4.1. Also due to the ecologic nature of the study, it is unknown whether it was the original population of a rural area or rather the incoming population that experienced the increased risk of leukaemia following population mixing. Thus, no clear conclusion about an infectious etiology of childhood leukaemia can be made.

Analyses by time period in this study suggested that the risk due to population mixing was declining over time, thereby implying a possible increase in herd immunity over time. However, leukaemia rates in Ontario have remained relatively stable throughout the study period and in recent times. Herd immunity and population mixing are both ecologic measures, and it is possible that the risk due to population mixing is decreasing over time. However, given that rates are not concurrently decreasing, if the etiology of leukaemia involves an infectious component, then individual factors are probably maintaining the stable rates. Because of individual variation in immunity as well as in exposure to the relevant infectious agent or agents, in order to adequately examine the infectious etiology hypothesis, it is necessary to conduct a study at the individual level.

Based on this study as well as past studies, it appears that an infectious etiology is most pertinent to the ALL subtype. Greaves' hypothesis which is specific to ALL, also refers to timing of exposures. Specifically, exposure to infections in infancy seems to be relevant in protection from leukaemia risk. Further examination of the potential infectious etiology of leukaemia in children requires the restriction of cases to the ALL subtype and the assessment of exposure should account for timing with respect to age, specifically whether exposure occurred in infancy. Because the causal pathogenic agent is unknown, a detailed assessment of occurrence of common infections in infancy

may provide clues.[98] Since an infectious etiology of ALL may be most likely in the youngest age group, it is important that attention be given to age at diagnosis in future studies.

6.9. Conclusions

In this study, the objective was to examine the ecologic effect of population mixing on the rate of childhood leukaemia. The intention of this study was to provide information which, in conjunction with results of other similar studies, could generate hypotheses with regard to a viral component to the etiology of childhood leukaemia. In fact, results from this study were supportive of previous research, and furthermore, the association was demonstrated in a new setting. In accounting for different levels of geographic isolation, it was found that the relationship between population mixing and childhood leukaemia, with respect to the infectious hypothesis, was specific to rural areas. Analyses by age and subtype emphasized the importance of accounting for possible etiological differences by these factors. In total, the results from ecologic studies of the Kinlen hypothesis provide information suggestive of an infectious etiology of childhood ALL, and requires a more in-depth examination.

7. REFERENCES

1. Holowaty EJ, Marrett LD, Fehringer G. Cancer Incidence in Ontario: Trends and Regional Variations in the 1980s. 1995; Toronto, Ont: The Ontario Cancer Treatment and Research Foundation.

2. McKenna SM, Baehner RL. Diagnosis and treatment of childhood acute lymphocytic leukemia. In: Wiernik PH, Canellos GP, Dutcher JP, Kyle RA, editors. Neoplastic disease of the blood. 3rd ed. New York: Churchill Livingston, 1996:271-294.

3. Goasguen JE, Bennett JM, Henderson ES. Biologic diagnosis of leukemias. In: Henderson ES, Lister TA, Greaves MF, editors. Leukemia. 6th ed. Philadelphia: Saunders, 1996:8-33.

4. Bain BJ. Leukaemia Diagnosis A Guide to the FAB Classification. London: Gower Medical Publishing, 1990.

5. Wang PP, Haines CS. Childhood and adolescent leukaemia in a North American population. Int J Epidemiol 1995; 24:1100-1109.

6. National Cancer Institute of Canada: Canadian Cancer Statistics 1998. Toronto, Canada, 1998.

7. Linet MS, Cartwright RA. The leukaemias. In: Schottenfield D, Fraumeni JF, editors. Cancer epidemiology and prevention. 2nd ed. New York: Oxford University Press, 1996.

8. Silverman LB, Weinstein HJ. Treatment of childhood leukemia. Curr Opin Oncol 1997; 9:26-33.

9. Petridou E, Revinthi K, Alexander FE, Haidas S, Koliouskas D, Kosmidis H, et al. Space-time clustering of childhood leukaemia in Greece: evidence supporting a viral aetiology. Br J Cancer 1996; 73:1278-1283.

10. Reynolds T. Causes of childhood leukemia beginning to emerge. J Natl Cancer Inst 1998; 90:8-10.

11. Reynolds T. Researchers hunt for elusive environmental causes of leukemia. J Natl Cancer Inst 1998; 90:90-92.

12. Alexander FE. Viruses, clusters and clustering of childhood leukaemia: a new perspective? Eur J Cancer 1993; 29A:1424-1443.

13. Kinlen LJ. Evidence for an infective cause of childhood leukaemia: comparison of a Scottish New Town with nuclear reprocessing sites in Britain. Lancet 1988; 2:1323-1327.

14. Miles A. Herd infection and herd immunity. In: Wilson G, Miles A, Parker MT, editors. Topley & Wilson's Principles of bacteriology, virology and immunity. 7th ed. Baltimore: Williams & Wilkins, 1983:413-428.

15. Ross JA, Potter JD, Robison LL. Infant leukemia, topoisomerase II inhibitors, and the MLL gene. J Natl Cancer Inst 1994; 86:1678-1680.

16. Sandler DP, Ross JA. Epidemiology of acute leukemia in children and adults. Semin Oncol 1997; 24:3-16.

17. MacMahon B. Is acute lymphoblastic leukemia in children virus-related? Am J Epidemiol 1992; 136:916-924.

18. Alexander FE, Chan LC, Lam TH, Yuen P, Leung NK, Ha SY, et al. Clustering of childhood leukaemia in Hong Kong: association with the childhood peak and common acute lymphoblastic leukaemia and with population mixing. Br J Cancer 1997; 75:457-463.

19. Greaves MF. Aetiology of acute leukaemia. Lancet 1997; 349:344-349.

20. Boice JD, Inskip PD. Radiation-induced leukemia. In: Henderson ES, Lister TA, Greaves MF, editors. Leukemia. 6th ed. Philadelphia: Saunders, 1996:195-209.

21. Miller RW. Special susceptibility of the child to certain radiation-induced cancers. Environ Health Perspect 1995; 103:Suppl 6:41-4.

22. Sandler DP. Recent studies in leukemia epidemiology. Curr Opin Oncol 1995; 7:12-18.

23. Nelson NJ. Solutions bring new problems: secondary cancers are a risk for some survivors. J Natl Cancer Inst 1997; 89:1483-1485.

24. Smith MA, McCaffrey RP, Karp JE. The secondary leukemias: challenges and research directions. J Natl Cancer Inst 1996; 88:407-418.

25. Ross JA, Davies SM, Potter JD, Robison LL. Epidemiology of childhood leukemia, with a focus on infants. Epidemiol Rev 1994; 16:243-272.

26. Martinez-Climent JA. Molecular cytogenetics of childhood hematological malignancies. Leukemia 1997; 11:1999-2021.

27. Ochs J. Childhood acute lymphoblastic leukemia. In: Henderson ES, Lister TA, Greaves MF, editors. Leukemia. 6th ed. Philadelphia: Saunders, 1996:419-445.

28. Taylor GM, Birch JM. The hereditary basis of human leukemia. In: Henderson ES, Lister TA, Greaves MF, editors. Leukemia. 6th ed. Philadelphia: Saunders, 1996:210-245.

29. Roman E, Beral V. Possible aetiological factors in childhood leukaemia. Arch Dis Child 1991; 66:179-180.

30. Gilman EA, Knox EG. Geographical distribution of birth places of children with cancer in the UK. Br J Cancer 1998; 77:842-849.

31. Zahm SH, Devesa SS. Childhood cancer: overview of incidence trends and environmental carcinogens. Environ Health Perspect 1995; 103:Suppl 6:177-84.

32. Daniels JL, Olshan AF, Savitz DA. Pesticides and childhood cancers. Environ Health Perspect 1997; 105:1068-1077.

33. John EM, Savitz DA, Sandler DP. Prenatal exposure to parents' smoking and childhood cancer. Am J Epidemiol 1991; 133:123-132.

34. Klebanoff MA, Clemens JD, Read JS. Maternal smoking during pregnancy and childhood cancer. Am J Epidemiol 1996; 144:1028-1033.

35. Ji BT, Shu XO, Linet MS, Zheng W, Wacholder S, Gao YT, et al. Paternal cigarette smoking and the risk of childhood cancer among offspring of nonsmoking mothers. J Natl Cancer Inst 1997; 89:238-244.

36. Kaye SA, Robison LL, Smithson WA, Gunderson P, King FL, Neglia JP. Maternal reproductive history and birth characteristics in childhood acute lymphoblastic leukemia. Cancer 1991; 68:1351-1355.

37. Westergaard T, Andersen PK, Pedersen JB, Olsen JH, Frisch M, Sorensen HT, et al. Birth characteristics, sibling patterns, and acute leukemia risk in childhood: a population-based cohort study. J Natl Cancer Inst 1997; 89:939-947.

38. Petridou E, Trichopoulos D, Kalapothaki V, Pourtsidis A, Kogevinas M, Kalmanti M, et al. The risk profile of childhood leukaemia in Greece: a nationwide case-control study. Br J Cancer 1997; 76:1241-1247.

39. Heath CW, Hasterlik RJ. Leukemia among children in a suburban community. Am J Med 1963; 34:796-812.

40. Kinlen LJ. Epidemiological evidence for an infective basis in childhood leukaemia. Br J Cancer 1995; 71:1-5.

41. Gilman EA, Knox EG. Childhood cancers: space-time distribution in Britain. J Epidemiol Community Health 1995; 49:158-163.

42. Petridou E, Alexander FE, Trichopoulos D, Revinthi K, Dessypris N, Wray N, et al. Aggregation of childhood leukemia in geographic areas of Greece. Cancer Causes Control 1997; 8:239-245.

43. Knox EG, Gilman EA. Spatial clustering of childhood cancers in Great Britain. J Epidemiol Community Health 1996; 50:313-319.

44. Muirhead CR. Childhood leukemia in metropolitan regions in the United States: a possible relation to population density? Cancer Causes Control 1995; 6:383-388.

45. Hjalmars U, Kulldorff M, Gustafsson G, Nagarwalla N. Childhood leukaemia in Sweden: using GIS and a spatial scan statistic for cluster detection. Stat Med 1996; 15:707-715.

46. Alexander FE, Ricketts TJ, McKinney PA, Cartwright RA. Community lifestyle characteristics and risk of acute lymphoblastic leukaemia in children. Lancet 1990; 336:1461-1465.

47. Stiller CA, Boyle PJ. Effect of population mixing and socioeconomic status in England and Wales, 1979-85, on lymphoblastic leukaemia in children. BMJ 1996; 313:1297-1300.

48. Alexander FE, Cartwright RA, McKinney PA, Ricketts TJ. Leukaemia incidence, social class and estuaries: an ecological analysis. Journal of Public Health Medicine 1990; 12:109-117.

49. Rodrigues L, Hills M, McGale P, Elliott P. Socio-economic factors in relation to childhood leukaemia and non-Hodgkin lymphomas: an analysis based on small area statistics for census tracts. In: Draper GJ, editor. The geographical epidemiology of childhood leukaemia and non-Hodgkin's lymphomas in Great Britain, 1966-83. London: HMSO, 1991:47-56.

50. Petridou E, Kassimos D, Kalmanti M, Kosmidis H, Haidas S, Flytzani V, et al. Age of exposure to infections and risk of childhood leukaemia. BMJ 1993; 307:774

51. Schulz TF, Neil JC. Viruses and leukemia. In: Henderson ES, Lister TA, Greaves MF, editors. Leukemia. 6th ed. Philadelphia: Saunders, 1996:160-178.

52. Kinlen LJ, Clarke K, Hudson C. Evidence from population mixing in British New Towns 1946-85 of an infective basis for childhood leukaemia. Lancet 1990; 336:577-582.

53. Kinlen LJ, Petridou E. Childhood leukemia and rural population movements: Greece, Italy, and other countries. Cancer Causes & Control 1995; 6:445-450.

54. Kinlen LJ. High-contact paternal occupations, infection and childhood leukaemia: five studies of unusual population-mixing of adults. Br J Cancer 1997; 76:1539-1545.

55. Langford I. Childhood leukaemia mortality and population change in England and Wales 1969-73. Soc Sci Med 1991; 33:435-440.

56. Draper GJ. The geographical epidemiology of childhood leukaemia and non-Hodgkin lymphomas in Great Britain, 1966-83. London: HMSO, 1991.

57. Petridou E, Hsieh CC, Kotsifakis G, Skalkidis Y, Trichopoulos D. Absence of leukaemia clustering on Greek islands. Lancet 1991; 338:1204-1205.

58. Dockerty JD, Cox B, Borman B, Sharples K. Population mixing and the incidence of childhood leukaemias: retrospective comparison in rural areas of New Zealand. BMJ 1996; 312:1203-1204.

59. Kinlen LJ, O'Brien F, Clarke K, Balkwill A, Matthews F. Rural population mixing and childhood leukaemia: effects of the north sea oil industry in Scotland, including the area near Dounreay nuclear site. BMJ 1993; 306:743-748.

60. Statistics Canada. Geography guide book. Ottawa: Industry, Science and Technology Canada, 1993. 1991 Census of Canada. Catalogue number 92-310E.

61. Statistics Canada. Standard Geographical Classification 1981, Volume 1. Ottawa: Minister of Supply and Services, 1982.

62. Clarke EA, Marrett LD, Kreiger N. Cancer registration in Ontario: a computer approach. IARC Scientific Publications 1991; 246-257.

63. Statistics Canada. 1986 Postal Code Conversion File, 1988.

64. Statistics Canada. 1991 Postal Code Conversion File, 1993.

65. Statistics Canada. Census Divisions and Subdivisions - Population, Occupied Private Dwelling, Private Households, Census Families in Private Households Selected Characteristics. Ottawa: Supply and Services Canada, 1982. 1981 Census of Canada. Catalogue number 93-X-918 (E-564).

66. Statistics Canada. Population and Dwelling Characteristics - Census Divisions and Subdivisions Ontario: Part 1. Profiles. Ottawa: Supply and Services Canada, 1987. 1986 Census of Canada. Catalogue number 94-111.

67. Statistics Canada. Profile of Census Divisions and Subdivisions in Ontario, Part A. Ottawa: Industry, Science and Technology Canada, 1992. 1991 Census of Canada. Catalogue number 95-337.

68. TETRAD Computer Applications Ltd. PCensus Version 4.12 for Windows. 1998.

69. Statistics Canada. 1981 Census Dictionary. Ottawa: Supply and Services Canada, 1982. 1981 Census of Canada. Catalogue number 99-101.

70. Statistics Canada. 1986 Census Dictionary. Ottawa: Supply and Services Canada, 1987. 1986 Census of Canada. Catalogue number 99-101E.

71. Statistics Canada. 1991 Census Dictionary. Ottawa: Supply and Services Canada, 1992. 1991 Census of Canada. Catalogue number 92-301E.

72. Jones CB. Coordinate systems, transformations and map projections. In: Jones CB. Geographical Information Systems and Computer Cartography. Essex: Addison Wesley Longman Limited, 1997:61-69.

73. Ontario/Canada Official Road Map. 1:800,000. Downsview, ON: Ministry of Transportation and Communications, 1977.

74. Ontario Yours to Discover Official Road Map 1980/81. 1:800,000. Downsview, ON: Ministry of Transportation and Communications, 1981.

75. Ontario, Canada 1986/87 Official Road Map. 1:700,000. Downsview, ON: Ministry of Transportation and Communications, 1987.

76. Greenland S, Robins J. Invited commentary: Ecologic studies--biases, misconceptions, and counterexamples. Am J Epidemiol 1994; 139:747-760.

77. Statistics Canada. Consumer Price Indexes for Ontario, Annual, 1992 Classification. CANSIM (Canadian Socio-economic Information Management System) Database 1997.

78. Sharp L, Black RJ, Harkness EF, McKinney PA. Incidence of childhood leukaemia and non-Hodgkin's lymphoma in the vicinity of nuclear sites in Scotland, 1968-93. Occupational & Environmental Medicine 1996; 53:823-831.

79. Bithell JF, Dutton SJ, Draper GJ, Neary NM. Distribution of childhood leukaemias and non-Hodgkin's lymphomas near nuclear installations in England and Wales. BMJ 1994; 309:501-505.

80. Clarke EA, McLaughlin J, Anderson TW. Childhood leukaemia around Canadian nuclear facilities - Phase II. Ottawa: Atomic Energy Control Board, 1991.

81. SAS Institute Inc. The SAS System for Windows, Release 6.12. Cary, NC.

82. Statistics and Epidemiology Research Corporation (SERC). EGRET Reference Manual. Seattle, WA. 1993.

83. Walter SD. The ecologic method in the study of environmental health. II. Methodologic issues and feasibility. Environ Health Perspect 1991; 94:67-73.

84. Rothman KJ, Greenland S. Modern Epidemiology. 2nd ed. Philadelphia: Lippincott-Raven, 1998.

85. Frome EL, Checkoway H. Epidemiologic Programs for Computers and Calculators. Use of Poisson regression models in estimating incidence rates and ratios. Am J Epidemiol 1985; 121:309-323.

86. Kuhn L, Davidson LL, Durkin MS. Use of Poisson regression and time series analysis for detecting changes over time in rates of child injury following a prevention program. Am J Epidemiol 1994; 140:943-955.

87. Morgenstern H. Ecologic studies in epidemiology: concepts, principles, and methods. Ann Rev Public Health 1995; 16:61-81.

88. Clarke EA, Marrett LD, Kreiger N. Twenty Years of Cancer Incidence, 1964-1983: The Ontario Cancer Registry. Toronto: The Ontario Cancer Treatment and Research Foundation, 1987.

89. Statistics Canada. 1991 Census Handbook. Ottawa: Supply and Services Canada, 1992. 1991 Census of Canada. Catalogue number 92-305E.

90. Statistics Canada. Immigration and Citizenship. Ottawa: Supply and Services Canada, 1992. 1991 Census of Canada. Catalogue number 95-316.

91. Kinlen LJ, Hudson C. Childhood leukaemia and poliomyelitis in relation to military encampments in England and Wales in the period of national military service, 1950-63. BMJ 1991; 303:1357-1362.

92. Kinlen LJ, Dickson M, Stiller CA. Childhood leukaemia and non-Hodgkin's lymphoma near large rural construction sites, with a comparison with Sellafield nuclear site. BMJ 1995; 310:763-768.

93. Greaves MF, Alexander FE. An infectious etiology for common acute lymphoblastic leukemia in childhood? Leukemia 1993; 7:349-360.

94. Greaves MF. Speculations on the cause of childhood acute lymphoblastic leukemia. Leukemia 1988; 2:120-125.

95. Greaves MF, Chan LC. Is spontaneous mutation the major 'cause' of childhood acute lymphoblastic leukaemia? Br J Haematol 1986; 64:1-13.

96. Raaschou-Nielsen O, Lohse C, Thomsen BL, Skov H, Olsen JH. Ambient air levels and the exposure of children to benzene, toluene, and xylenes in Denmark. Environmental Research 1998; 75:149-159.

97. Schnatter AR, Armstrong TW, Nicolich MJ, Thompson FS, Katz AM, Huebner WW, et al. Lymphohaematopoietic malignancies and quantitative estimates of exposure to benzene in Canadian petroleum distribution workers. Occupational & Environmental Medicine 1998; 53:773-781.

98. van Steensel-Moll HA, Valkenburg HA, van Zanen GE. Childhood leukemia and infectious diseases in the first year of life: a register-based case-control study. Am J Epidemiol 1986; 124:590-594.