Lifestyle variables in relation to disease patterns a statistical analysis of the Windsor data.

Michael Arthur. Desmarais

University of Windsor

Follow this and additional works at: https://scholar.uwindsor.ca/etd

Recommended Citation
https://scholar.uwindsor.ca/etd/2607
The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us a poor photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30. Please read the authorization forms which accompany this thesis.

THIS DISSERTATION HAS BEEN MICROFILMED EXACTLY AS RECEIVED.

Ottawa, Canada
K1A 0N4
LIFESTYLE VARIABLES IN RELATION TO DISEASE PATTERNS:
A STATISTICAL ANALYSIS OF THE WINDSOR DATA

by

Michael Arthur Desmarais

A Thesis
Submitted to the Faculty of Graduate Studies
through the Department of Geography
in Partial Fulfillment of the
Requirements for the Degree
of Master of Arts at the
University of Windsor

Windsor, Ontario, Canada
1981
ABSTRACT

An investigation into the incidences of ischaemic heart disease and lung cancer mortalities, as represented by standardized mortality rates (S.M.R.'s) computed separately for both sexes, 1970-1977, was carried out in order to determine if inter-metropolitan mortality variations exist within the city of Windsor, Ontario. Also, an investigation into various lifestyle patterns was carried out in order to determine if an one-lifestyle, or combined lifestyles, were causal in nature for the occurrence of these mortalities in Windsor. The mortality data were made available by Dr. J. Jones, Medical Officer of the Essex County Metropolitan Windsor Health Unit, and analyzed by the manipulation and choropleth mapping of standardized mortality rates, significant levels, Analytic Model analyses, multiple factor analyses, and multiple regression analyses. The lifestyle data were collected by means of a questionnaire survey, administered randomly throughout the city of Windsor, Ontario, during the summer of 1978. These data were then choropleth mapped and statistically analyzed as variables entered into the regression model. This paper established, by the census-tract delineation of standardized mortality rates, the existence of distinct areal variations of ischaemic heart disease and lung cancer mortalities within the confines of metropolitan Windsor. Results indicated that for both disease types, and both sexes, higher than expected S.M.R. values occurred in census-tracts situated near and/or along the downriver southwest portion of the city. Lower than expected rates for both disease types and sexes were found to be located in middle
to upper income suburban newer housing areas of the southern and eastern portions of Windsor. In the city-wide context, female mortality incidences for both disease types were approximately 18% higher than the male rates and both values were higher than expected when calculated by the Provincial Derived Norm methodology. Census-tract based multiple regression analyses were undertaken and the results, at best, weakly related certain ethnic types, occupation, income groups, smoking, and the birth control pill, to the spatial pattern of these diseases. Future research will have to concentrate on obtaining more specific medical histories of individuals in order to gain more understanding of the factors related to both heart disease and lung cancer.
ACKNOWLEDGEMENTS

I would like to express my gratitude to all those who assisted in the completion of this project. I would like to thank my family for their patience and much needed support. Also, I would like to extend special thanks to Dr. F. C. Innes, my advisor, for his support and helpful ideas. As well, I would like to thank Dr. P. D. LaValle for his guidance on the statistical procedures used. Also, special thanks to Ron Welch, who freely offered his cartographic expertise.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Nature and Scope of the Problem</td>
<td>4</td>
</tr>
<tr>
<td>II</td>
<td>REVIEW OF THE LITERATURE</td>
<td>7</td>
</tr>
<tr>
<td>III</td>
<td>MODEL DESIGN</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>Hypotheses,</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>Data Collection and Methodology</td>
<td>25</td>
</tr>
<tr>
<td>IV</td>
<td>DATA ANALYSIS</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>Interpretation of Survey Evidence</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>Male Ischaemic Heart Disease Mortalities:</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td>Findings</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Male Lung Cancer Mortalities:找到了</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>Findings</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lifestyle Characteristics of the Respondents in Aggregate: Male</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td>Ischaemic Heart Factor Variables Findings: Male</td>
<td>66</td>
</tr>
<tr>
<td></td>
<td>Males</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Regression Analysis: Findings</td>
<td>69</td>
</tr>
<tr>
<td></td>
<td>Lung Cancer Factor Variables Findings: Male</td>
<td>74</td>
</tr>
<tr>
<td></td>
<td>Findings</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Regression Analysis Male Lung Cancer: Findings</td>
<td>79</td>
</tr>
<tr>
<td></td>
<td>Findings</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Female Ischaemic Heart Mortality: Findings</td>
<td>81</td>
</tr>
<tr>
<td></td>
<td>Female Lung Cancer Mortalities: Findings</td>
<td>85</td>
</tr>
<tr>
<td></td>
<td>Lifestyle Characteristics of the Respondents in Aggregate: Female</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td>Ischaemic Heart Factor Variables Findings: Female</td>
<td>101</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Regression Analysis Female Heart Disease: Findings</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td>Findings</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lung Cancer Factor Variables Findings: Female</td>
<td>108</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td></td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table                                    | Page  
---                                       | ------
1  Multi-Etiologic Model                  | 24    
2  Summary Table for Windsor Ischaemic Heart Disease and Lung Cancer Mortality Rates | 40    
3  Male Ischaemic Heart Disease Mortalities: 1970-1977 | 43    
4  Mortalities of Bronchogenic Lung Cancer - Male (1970-1977) | 49    
5  Eigenvalues for the Lifestyle Factors | 67    
6  Quartimax Rotated Factor Matrix        | 68    
7a Eigenvalues for the Ethnic Factors    | 70    
7b Quartimax Rotated Factor Matrix        | 70    
8a Eigenvalues for the Socio-Economic Factors | 71    
8b Quartimax Rotated Factor Matrix        | 71    
9  The Relationship between Male Ischaemic Heart Mortalities and Selected Lifestyle, Ethnic and Socio-Economic Variables by Census Tract | 72    
10 Eigenvalues for the Lifestyle Factors  | 75    
   Male Lung Cancer                       |       
11 Eigenvalues for the Ethnic Variables   | 77    
   Male Lung Cancer                       |       
12 Eigenvalues for the Socio-Economic Variables | 78    
   Male Lung Cancer                       |       
13 Relationship between Male Lung Cancer Mortalities and Selected Lifestyle, Ethnic and Socio-Economic Variables by Census Tract | 80    
14 Mortalities of Ischaemic Heart Disease - Female 1970-1977 | 82    
15 Mortalities of Bronchogenic Lung Cancer - Female 1970-1977 | 86    
16 Eigenvalues for the Lifestyle Factors  | 103   
   Female Heart Disease                   |       

VII
<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td>Eigenvalues for the Ethnic Variables Female Heart Disease</td>
</tr>
<tr>
<td>18</td>
<td>Eigenvalues for the Socio-Economic Factors Female Heart Disease</td>
</tr>
<tr>
<td>19</td>
<td>Relationship between Female Ischaemic Heart Disease Mortalities: 1970-1977 and Selected Lifestyle, Ethnic and Socio-Economic Variables by Census Tract</td>
</tr>
<tr>
<td>20</td>
<td>Eigenvalues for the Lifestyle Factors Female Lung Cancer</td>
</tr>
<tr>
<td>21</td>
<td>Eigenvalues for the Ethnic Factors Female Lung Cancer</td>
</tr>
<tr>
<td>22</td>
<td>Eigenvalues for the Socio-Economic Variables Female Lung Cancer</td>
</tr>
<tr>
<td>23</td>
<td>Relationship between Female Lung Cancer Disease Mortalities: 1970-1977 and Selected Lifestyle, Ethnic and Socio-Economic Variables by Census Tract</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>1</td>
<td>City of Windsor Census Tracts</td>
</tr>
<tr>
<td>2</td>
<td>City of Windsor Census Tracts: [isodemographic Base] (2)</td>
</tr>
<tr>
<td>3</td>
<td>Mortalities of Ischaemic Heart Disease - Male 1970-1977</td>
</tr>
<tr>
<td>4</td>
<td>Mortalities of Ischaemic Heart Disease - Male 1970-1977</td>
</tr>
<tr>
<td>5</td>
<td>Mortalities of Lung Cancer - Male 1970-1977</td>
</tr>
<tr>
<td>7</td>
<td>Percentage of Respondents with High Blood Pressure - Male</td>
</tr>
<tr>
<td>8</td>
<td>Percentage of Heavy Smokers - Male</td>
</tr>
<tr>
<td>9</td>
<td>Percentage of Respondents Overweight - Male</td>
</tr>
<tr>
<td>10</td>
<td>Percentage of Respondents Underweight - Male</td>
</tr>
<tr>
<td>11</td>
<td>Per Capita Alcohol Consumption Per Week in Ounces</td>
</tr>
<tr>
<td>12</td>
<td>Prevalence of Diagnosed Heart Disease - Male</td>
</tr>
<tr>
<td>13</td>
<td>Mortalities of Ischaemic Heart Disease - Female 1970-1977</td>
</tr>
<tr>
<td>14</td>
<td>Mortalities of Ischaemic Heart Disease - Female 1970-1977</td>
</tr>
<tr>
<td>15</td>
<td>Mortalities of Lung Cancer - Female 1970-1977</td>
</tr>
<tr>
<td>16</td>
<td>Mortalities of Lung Cancer - Female 1970-1977</td>
</tr>
<tr>
<td>17</td>
<td>Percentage of Respondents with High Blood Pressure - Female</td>
</tr>
<tr>
<td>18</td>
<td>Percentage of Heavy Smokers - Female</td>
</tr>
<tr>
<td>19</td>
<td>Percentage of Respondents Overweight - Female</td>
</tr>
<tr>
<td>20</td>
<td>Percentage of Respondents Underweight - Female</td>
</tr>
</tbody>
</table>
Figure 21  Percentage of Females on the Birth Control Pill Greater than Four Years who Smoke

Figure 22  Prevalence of Diagnosed Heart Disease - Female
CHAPTER 1

INTRODUCTION

The World Health Organization defines health as a state of physical, psychological, and social well-being, and not merely the absence of disease and infirmity (Susser, 1973, 4). Ill-health, on the other hand, has been described as the consequence of the maladjustment of man to his environment (May, 1974, 33). Disease, therefore, can be seen to be the result of an interaction between man and his environment; and the risk of developing a disease will increase with the lengthening of exposure to a causal factor or factors in the environment (Grist, 1972, 213).

The spatial observation of disease patterns represents one approach to the identification of causal factors associated with abnormal occurrence of disease. An investigative sub-branch of geography, termed "Medical Geography," concerns itself with such analytic activities. Medical geography may be defined as the comparative study of the incidence of disease and the distribution of physiological traits in people belonging to different communities throughout the world, and the correlation of these data with factors of the environment. Different ethnic, national, or social groups may be involved. These groups may live in close proximity or may be remote from one another. True association with place is suggested when age-adjusted risk of disease increases for immigrants and decreases for
emigrants, when all ethnic groups present are at a similar high risk. The subject of medical geography is closely related to epidemiology, except that in medical geography the emphasis is on patterns of distributions and relationships (Howe, 1972).

Historically, epidemiological and medical geographic studies have emphasized the study of great epidemic diseases (especially cholera, smallpox, typhus and yellow fever) which, until the twentieth century, were the most important threats to human life and health. Thus, the term epidemic is defined as "an unusually frequent occurrence of a disease. To understand unusual disease occurrence, it soon was realized that the usual, or endemic, occurrence must also be studied. The methods developed for such study were readily applicable to the study of infectious disease; whether epidemic or not. In relation to chronic diseases, the importance of cultural patterns and environmental factors, depending on social and economic status, rather than infectious agents became apparent.

Epidemiological, medical and geographical studies have led investigators to conclude that the inhabitants of certain areas exhibit health, others disease, and that such occurrences must be responses to the social and physical environment. If areas of health and disease can be identified, researchers can attempt to associate environmental factors identified with particular disease types. Such knowledge, gathered from interdisciplinary efforts, may enable the determination of those precursors of agents that affect states of health or ill-health in man and thus resolve the existence of geographical differentiation in disease incidence.

Today, influenza outbreaks excepted, major epidemics no longer seri-
ously threaten the more developed countries and most of the more important infectious diseases are reasonably under control.

Meanwhile, diseases of other types have assumed importance as the current major health problems of Western civilization. The importance of these diseases stems in part, from major changes in environment and way of life imposed by industrialization of the economy and related migration of people to the cities. These patterns also relate to the great numerical and proportional increase in older age groups in the population which has resulted from the withdrawal of infection as a common cause of early death. Cancer of nearly all types, high blood pressure, coronary artery disease, diabetes, and arthritis are among the chronic diseases associated with older age. Cancer of several types are related, in some part, to changes in the environment. With little or no modification, the principles and methods developed to explore causes for the occurrence of infectious diseases are proving applicable to the study of these increasingly important chronic diseases. Investigation of causal factors represents the new epidemiological and medical geographic frontier.

Some may argue that the forced withdrawal of infection as a major cause of death has simply left us to face the ultimate and sometimes less merciful alternate causes of death from which no escape is possible (Fox, 1971). While death itself is inevitable, certain important causes of death can be eliminated or greatly reduced when their causation is fully understood; and in other cases, death may be delayed. Lung cancer is one example where the increasing activity of some potent environmental influence (i.e., cigarette smoking) may be effectively neutralized. This, and other instances in which an environmental cause of cancer has been clearly identified (i.e., radiation in leukemia) afford a
reasonable basis for expectation that environmental influences susceptible to control may be important causes of death for many forms of cancer.

Coronary heart disease is another important example. Although sex and genetic factors are partial determinants, this disease has become increasingly frequent in middle aged men in the United States and Canada. The business of the medical geographer, therefore, is to undertake a study designed to identify the basic spatial structure of disease mortalities and attempt to derive a relationship between a number of complex environmental and personal lifestyle factors, which might be responsible for the generation of disease mortality patterns.

Nature and Scope of the Problem

Earlier work on the patterns of disease causing death in Windsor, Ontario, has shown that for both lung cancer and heart disease there are, when standardized mortality rates are employed, marked differences in experience between city census tracts. In this Windsor is no different from Chicago (Pyle, 1971) or the British cities of London and Glasgow (Howe, 1972, 1979).

It is initially important, however, to notice the basis of information used. As Howe, et al. have said, "Various indices are used to measure disease frequency. The incidence rate, based on the number of new cases developing each year, is the most informative but is generally not available on a global basis. Neither is prevalence rate which provides a measure of the total number of people affected at a given point in time. Actual mortality is the most readily available statistic, since certified deaths are collected as a matter of routine by many governments." It is this latter data, which were used in previous studies of Windsor, that
were tested and mapped after standardization and compared to expected values provincially and nationally. Moreover, the data used were for the age cohorts 15 thru 64 and hence excluded older age groups where diagnosis may be more difficult.

It is the intent of this writer to examine the basic risk factors involved in the possible causation of heart disease and respiratory disease within the city of Windsor, Ontario. All relevant mortality statistics have been compiled by the Essex County Metropolitan Windsor Health Unit. These data, disaggregated by sex and ten year age groupings, 15 thru 64, were transformed to standardized mortality rates, for the time period of 1970 to 1977.

All relevant morbidity statistics have been compiled by a questionnaire survey administered randomly throughout the city of Windsor, Ontario, during the summer of 1978. The data was then used to try to identify patterns of lifestyle differences that might indicate a correlation of coincidence. Then these correlations were compared to the mortality information to try and identify areas where there are both lifestyle differences and high mortalities of heart disease and respiratory disease. (The various types of lifestyle data will be discussed in detail later on). These data will be disaggregated spatially for each of Windsor's 45 census tracts.

Owing to the scope of such research, the reader must be made aware of the limitations that this study entails. The firmest measure of non-health of a population lies in mortality statistics and yet the recording of cause of death is often of a biased or misrepresented nature. For example, persons who have received a diagnosis of carcinoma are more likely to have the diagnosis appear on their death certificate even though the carcinoma may not be the underlying or contributory cause of death for that indivi-
dual (Morgan, 1977).

For a complete assessment of the associations of the various lifestyle habits with health or disease conditions, it is necessary to have both morbidity and mortality statistics. However, like mortality statistics, morbidity rates can be misleading especially when these rates are based on a sample population. When using a questionnaire survey to gather these statistics, one must rely on the honesty and reliability of the individuals answering such a questionnaire. It must be recognized, therefore, that no one study can assess all factors which may act independently or in conjunction with risk factors to affect the health of the individual.
CHAPTER II

REVIEW OF THE LITERATURE

Over the years, we have seen a relative increase of Medical Geographic papers. Many books have been written concerning the scope, aims and nature of this discipline. Examples of such works are Gelyakova (1967), McGlashan (1965, 1969), Khan (1971), Kratochvil (1971), Learmonth (1972), and Hunter (1974).

In the aforementioned references, a central theme supports the idea that disease may be regarded as an interaction between agent, host and environment. Since the field of medical geography is a discipline that focuses on the analysis of man-environment interaction, it meshes very finely with and carries analytical implications for the better understanding of the concept.

Literature pertaining to the application and design of medical geographic analytic techniques include Armstrong's Computer and Mapping in Medical Geography (1972), McGlashan and Bond's Computer Assistance with Geographic Correlations (1970), and McGlashan and Chick's Methods for Assessing Spatial Variations of Mortality (1974).

Spatially oriented studies of dietary patterns are best illustrated by May's Studies in Medical Geography (1953-1973), which include twelve volumes of country by country surveys of malnutrition in South Africa.

Stress related diseases, including alcoholism, drug addiction and
various types of mental illness, are represented by Pyle's *Heart Disease, Cancer and Stroke in Chicago* (1971) and Hunter and Brum's *The Geography of Stress* (1974). These types of investigations assist in the design of strategies in social and environmental engineering that will be preventative rather than causative of ill-health.

Disease associated with various types of environmental contaminants include Grist's *Simple Chronic Bronchitis and Urban Ecological Structure* (1976), Crofton's *A Study of Lung Cancer Mortality in Relation to Coal Mining in Scotland* (1969), and Hoover and Fraumeni's *Cancer Mortality in U.S. Counties with Chemical Industries* (1975).

The aspects of spatial allocation and planning for health care facilities are exemplified by Pyle's *Heart Disease, Cancer and Stroke in Chicago* (1971) and *The Geography of Health Care* (1974), and Shannon's *Urban Ecological Containers, Environmental Risk Cells, and the Use of Medical Services* (1976).

Work on the investigation of regional differences in chronic diseases, which include malignant neoplasms and coronary heart diseases, have mainly been confined to the United Kingdom, South Africa, North America and Australia.

The most common method of studying regional differences in these diseases involves examining the observed mortality rates for consistent trends. The types of consistency sought may be between the sexes, or adjacent geographical regions. Thus, if a region has unusually high (or low) mortality rates for both sexes in a number of age specific groups, and if rates are consistently high (or low) over a number of time periods and throughout a number of contiguous regions with high (or low) rates,
then it is probable that the region has abnormal mortality. Conversely, if rates are high in some age groups but low in others, and there is no consistent trend over time, then it is likely that the abnormal rates represent random fluctuations (El-Shaarwi, 1976).

One large collection of regional mortality data is G. M. Howes' *A National Atlas of Disease Mortality in the United Kingdom* (1970), which used standardized mortality rates to produce choropleth maps displaying the distributions of selected mortalities and diseases.

Maps that illustrate mortality rates by county, for various types of cancer, have recently been prepared by the National Cancer Institute in the United States (Mason, 1975). Also, in the United States, Sauer (1974) systematically studied regional differences in cardiovascular disease mortalities using States and State Economic Areas as the geographical units of study. Enterline (1960) and Chase (1963) also showed that marked differences in coronary death rates exist within American States. McGlashan and Chick (1974) examined the distribution patterns of ischaemic heart disease mortalities in Tasmania, Australia.

Stavraky (1976) investigated the distribution patterns of intestinal cancer in Ontario in the 1960's. Howe (1971) and Burbank (1971) undertook similar chronic disease studies in different parts of the world over varying time spans of from three to fifteen years. These studies, and other related works, use a standard statistical and cartographic methodology. Morbidity and mortality data for at least a concurrent three year period, based on some type of areal unit, is collected and categorized (in respect to disease site and type). For each category, data disaggregation by age and sex is made, and standardized mortality rates (S.M.R.'s) are calculated.

Standard deviation measures, Chi Square tests, the Geary Contiguity
Index or Significant Poisson Distribution values are then employed for these measurements to determine if significantly higher or lower than normal mortality or morbidity zones exist in the study area. A final series of maps, usually of a choropleth type, are made to combine the more extreme levels of S.M.R.'s and statistically significant mortality or morbidity variations (Murray, 1974).

McGlashan (1972) has stated that the geographers' contribution to medical knowledge can properly be expected to lie chiefly in the field of environmental studies and relation of diseases to other geographic variables. Such cause and effect related studies, utilizing regression analysis techniques, are few in number and have only recently been undertaken by geographers. Many biological and social environmental factors have been hypothesized as possible explanations of the geographic variations in respiratory mortality rates. These variables include aspects of urbanization, population density, socio-economic status, ethnicity, air pollution, and climatic factors. Schroeder (1960) showed that a marked association existed between the hardness of water supplies and mortality rates of cardiovascular disease. Pyle (1971), in Chicago, correlated demographic and socio-economic variables with the incidence of heart disease, cancer and stroke in an attempt to predict the concentrations of these diseases in 1980 for the purpose of health planning. Christou (1977) studied the spatial variations of cancer and heart disease in Ontario by county. He found that the variables which were significant for cancer were of two types: ethnic variables and occupational variables based on data supplied by the Census of Canada, 1971. Ethnic differences have also been investigated in Japan, and it was found by McMahon and Pugh (1970) that there was a very marked increase in the death rates due
to heart disease and lung cancer. Reasons for this tendency were either due to genetic background or life-styles of the Japanese, but this was not statistically determined.

Present day epidemiological research has concluded that respiratory diseases are of an etiological nature. That is, they are diseases of multiple causation, which appear to have so many determinants that they tend to defy orderly classification under the titles of agent, host and environment. It is certain, however, that diseases of this nature arise within the host as a result of a complex interaction between host and environment.

In a number of respiratory diseases, the evidence leading towards causal factors are indirect. The initial step is to investigate a possible association between the suspected factor, or factors and the disease. Any apparent association must be tested before a causal relationship can be inferred. Lung cancer greatly increased in Britain after the introduction of cigarette smoking, but it would be highly dangerous on this evidence alone to assume a causal relationship; the consumption of bananas and the manufacture of gramophones may have increased impressively during the same period (Crofton, 1975). However, when it is found that cigarette consumption has been very much higher in patients with lung cancer than in those of the same age, sex and social background without the disease, and this pattern is found in many countries, then this is evidence that there is an association between the suspected factor and the disease. In addition, throughout many countries, when individuals record their smoking habits when they are healthy, and in following up there appears to be a close relationship between the number of cigarettes smoked per day and the chances of developing lung cancer, then
this cumulative evidence leads to an overwhelming probability that there is not only an association but that it is causal.

Factors involved in the pathogenesis of respiratory disease may be classified as (1) Anaplastic, and (2) Adenocarcinomatous (Crofton, 1975). Anaplastic diseases arise mostly from the bronchial epithelium. It is these types which are related to tobacco smoking. The adenocarcinomata are derived from the mucous glands of the bronchial wall (Crofton, 1975). The approximate frequency of each type of carcinoma related to respiratory disease is as follows:

1. Anaplastic carcinoma: including sub-types 96%
2. Adenocarcinoma 3%
3. Very rare sub-types 1%

100%

Factors intrinsic to the host that may also modify agent factors (i.e., anaplastic and adenocarcinomata) leading to respiratory disease include genetic constitution, age, sex, and psychic state.

The environmental factors suspected to influence one or more of these conditions within the host include such variables as cigarette smoking, culture, ethnicity, atmospheric pollution, industrial and occupational factors and petrol fumes. It has also been suggested that carcinomas tend to arise at sites of previous scarring on lung tissue from tuberculosis, bronchitis, and other such types of respiratory disease (Spencer, 1978).

Current epidemiological research has also concluded that cardiovascular disease is a disease of multiple causation, which like respiratory diseases, has orderly classification under agent, host and environment (Olson, 1968).
Factors involved in the pathogenesis of cardiovascular disease may be grouped into humoral and local factors (Olson, 1957). The humoral factors include the lipoproteins of plasma, the heparinoid substances involved in clearing lipid from the blood and other plasma proteins concerned with coagulation. Under local factors, that is, those promoting atherogenesis as a result of influencing the condition and function of the artery itself are: (1) factors relating to blood flow, turbulence and pressure, (2) metabolic factors intrinsic to the arterial tissue, (3) structural factors relating to differentiation of intima, internal elastic membrane, and other components of the arterial wall, and (4) traumatic factors relating to injury, ulceration and repair (Olson, 1968). Environmental factors known to influence one or more of these conditions within the host include such variables as diet, drugs, cigarette smoking, exercise, occupation, water hardness, culture and climate.

Factors intrinsic to the host that may also modify the humoral and local factors leading to cardiovascular disease include genetic constitution, age, sex, and endocrine balance. The wide range of these diverse risk factors have led, in recent years, to a large number of cause and effect related disease studies being undertaken by medical practitioners, epidemiologists and geographers.

The following paragraphs will review those major risk factors currently held by epidemiologists to be primarily suspect for the occurrence and, in turn, geographic variability of cardiovascular disease and respiratory disease. Although treated separately for descriptive purposes, it must be remembered that these variables interact with or are associated with other risk factors in the occurrence of heart disease and respiratory disease.
The major risk factors for both heart disease and respiratory disease fall under the broad heading of lifestyle which includes dietary patterns, drug consumption (whether prescribed or not), cigarette consumption, alcohol consumption, and the degree of physical activity. Unlike any other time in modern history, a large mass of the population in affluent countries have been able to enjoy a rich diet high in animal products and have not been restricted by economic conditions to the consumption of cheap, starchy foods such as bread, potatoes and pasta. This modern diet, which is excessive in caloric intake in relation to energy expenditure, high in fatty substances, cholesterol, sugar, and salt, leads to high prevalence rates of hyperlipidemia in the adult population. When hyperlipidemia is sustained, it markedly increases the risk of cardiovascular disease (Stamler, 1973).

The International Atherosclerosis Project (1968) quantified the degree of atherosclerotic of the aorta and coronary arteries at autopsy in over 31,000 persons ages 10 thru 69 who died during 1960 thru 1965 in 15 cities throughout the world. Significant positive correlations were found among intake of fat, cholesterol levels and severity of coronary heart disease, at autopsy, for the populations from the 15 cities. Raine (1958) revealed a high mortality and morbidity rate from coronary heart disease in Finland, that positively correlated with a high mean cholesterol level and a diet with 35% of calories from fatty foods. Lehr (1972) found that a subject's income, occupation, and cigarette smoking habit were key factors in coronary heart disease incidence. Morris (1953) showed that lower frequencies of coronary heart disease occur in active rather than sedentary occupations. Stone (1978) found
that women under the age of 50 who smoke heavily (over 35 cigarettes per day) were 20 times more likely to have heart attacks than non-smokers. Dick (1977) found that the incidence of fatal heart attack, for women who took the birth control pill and consumed over 15 cigarettes per day, was 11 times greater than for non-smoking women not taking the pill. A study done by Health and Welfare Canada (1965) found that death rates among men who smoked cigarettes only increased 43% when 10 cigarettes or less were consumed per day. If more than 20 cigarettes per day were consumed, then the death rates among men increased 63% for all mortalities. Men who consumed more than 20 cigarettes per day had a lung cancer death rate almost 1400% higher than that of non-smokers. For other circulatory diseases, such as bronchitis and emphysema, the death rates for men who smoked were 87% higher than the death rates of non-smokers. Also, no evidence was found in this study of clear-cut associations between cause of death and occupation. Further, occupation did not appear to modify the established association of cigarette smokers with death rates in excess of those of non-smokers. The data concerning females indicated that a parallelism existed between the females and males. In both cases, excess mortality was associated with cigarette smokers as compared to non-smokers and the difference in mortality rates increased as the quantities of cigarettes smoked increased.

Other risk factors that are suspect in the etiology of heart disease include heredity, obesity, physical activity, and stress. The impact of these variables is briefly outlined in the following paragraphs.

Obesity adversely affects the function of many organs in the body including overloading of the circulation, and is likely to be an un-
desirable factor in a cardiac patient. Studies by Marks (1960) and Kannel (1967) have demonstrated that the obese individual has an increased risk in many respects. The incidence of heart disease in general is said to be 50% higher in overweight individuals than in those with normal weight, as measured by standardized mortality statistics (Seleer, 1966). Obesity is associated with an increased risk of coronary heart disease, particularly angina pectoris and sudden death. However, specific information linking obesity with acceleration of the atherosclerotic process is not available in statistical terms.

There is some evidence that physical inactivity increases the risk of coronary heart disease, but this relationship has not been firmly established. Mann (1955), Wong (1975), and Kannel (1967), in similarly conducted patient studies, found that exercise resulted in reduced plasma cholesterol and lower incidence and severity of aortic and coronary atherosclerosis when compared to non-exercised sedentary groups.

There appears to be some evidence that individuals living under constant stress and tension are subject to more frequent heart attacks than those with well regulated lives (Selzer, 1966), and Rosenman (1963). However, since stress can only be estimated, the relationship between atherosclerosis and stress could be expressed only in very general terms.

Hereditary factors also appear to influence heart disease incidence rates. Certain families show a high incidence of heart disease early in life, regardless of the environmental conditions. Other families seem to be immune to this disorder (Selzer, 1966).

It has been widely accepted that the epidemiologic search for any one specific cause of heart disease and respiratory disease is illus-
ory because heart disease and respiratory disease are of an etiologic nature and the effect of most risk factors start to operate early in life. Most epidemiological factors are either directly or indirectly linked with behaviour, environment and society. If these factors are to be changed, there must be sociological and lifestyle changes made also if we are to reduce the incidence of heart disease and respiratory disease.

The conclusions one must draw from the literature clearly indicate that changes in diet and life-style mode of life are essential to reduce the current rates of coronary heart disease and respiratory disease mortalities. These environmentally conditioned risk factors are all avoidable and their changes are vital prerequisites for rearing upcoming generations with new and better habits to assure the disappearance of now widely prevalent overeating, sedentary living, cigarette smoking and resulting respiratory disease and heart disease.
CHAPTER III

MODEL DESIGN

The principal conclusion from the literature is that many epidemiologists have demonstrated a link between life-style risk factors and heart disease and respiratory disease mortality rates.

As noted earlier, the most readily available measure of health status is the mortality rate for a geographically defined group. A large number of factors are hypothesized to affect the index of mortality. These factors can be grouped into four basic categories: 1) The physical characteristics, 2) The socioeconomic characteristics, 3) The environmental characteristics, and 4) The personal life-style characteristics.

The physical factors hypothesized to affect mortality are age cohorts, sex, and weight groupings. The socioeconomic factors hypothesized to affect mortality are income distribution, occupational mix, ethnic distribution, housing density and differential migration factors. The environmental factors hypothesized to affect mortality are atmospheric conditions, air pollution characteristics, and climate factors. The personal life-style characteristics hypothesized to affect mortality are tobacco consumption, exercise habits, alcohol consumption, nutritional habits, and drug consumption.

To prove causation, an investigation must control each of these factors statistically. Controlling all factors which might influence
mortality is, of course, impossible; but even though the analysis may be incomplete, because these factors cannot all be controlled, it can still be powerful enough to gain an association of these causal affects on mortality.

The variables to be utilized in this analysis will fall under three of the four categories previously mentioned. The risk factors which will not be utilized are those which fall under the category of environmental characteristics. This category will not be used because air pollution data for the city of Windsor is not complete. Also, it would be impossible to disaggregate the air pollution data so as to conform to the areal units utilized in this study. Climatic factors, which have been used in similar studies, have not produced positive correlations associated with the incidence of mortality (Bondy, 1968).

Under the physical characteristics, the age factor will be broken up into six sub-groupings: 1) less than 25 years, 2) 26 to 35 years, 3) 36 to 45 years, 4) 46 to 55 years, 5) 56 to 65 years, and 6) greater than 65 years of age.

Also, these groupings will be sub-divided by sex. The justification for breaking up the age groupings in this manner is in part because mortalities under the age of 25 may be small. The 10 year age groupings were chosen because of the significant number of mortalities in each cohort grouping. The maximum limit of 65 years of age was chosen because after this age other factors, such as deterioration of the internal organs, will affect the mortality, thus limiting the significance of the mortality rate. The final variable under this category, weight, will be broken down so as to represent the percentage of the sample population which is over-
weight and underweight. This was done by calculating a weight/height ratio for the sample population of each census tract and then, with the aid of a standard weight/height chart, the average weights for each sex were calculated. Once this was completed, the percentage found to be overweight and underweight, based on height, were calculated for each tract by sex groupings.

The weight variable was chosen because obesity has been cited as one risk factor which may cause the incidence of heart disease mortality. Also, allegedly associated with obesity is the incidence of high blood pressure and consumption of fatty foods.

The socio-economic variables include income distribution, occupational mix, and ethnic distribution. Incomes were classified into three sub-groups which are as follows: 1) less than $12,000, 2) $12,001 to $30,000, and 3) greater than $30,000. These groupings were chosen because the lowest represents the poverty level and each subsequent level represents the average income based on each different major occupational type. The occupational characteristics will be divided into six sub-groups including: 1) managerial, 2) clerical and sales, 3) farming, 4) chemical and mining, 5) assembly and construction, and 6) transportation and related industries. These groups were chosen because they represent the major occupational types as set forth and used by the census of Canada. The ethnic variables will be divided into four sub-groups including: 1) British and French, 2) German, 3) East European, and 4) North European. These groups were chosen because they are representative of the ethnic breakup, spatially, throughout the city of Windsor.

The personal life-style variables include tobacco consumption,
alcohol consumption, nutritional habits, and drug consumption. Tobacco consumption will be divided into three major categories based on type of smoking, including: 1) cigarettes, 2) cigars, and 3) pipe. These three major categories will then be divided into three sub-categories based on the intensity of smoking. These include: 1) light smokers (less than 10 cigarettes per day), 2) medium smokers (between 10 and 20 cigarettes per day), and 3) heavy smokers (greater than 20 cigarettes per day). These categories were chosen so as to disaggregate the data according to type and intensity. Intensities were based on categories used in a previous study conducted by Health and Welfare Canada (1965).

Alcohol consumption will be broken down into three major categories and six sub-categories, based on type and intensity of consumption per week. The three major categories are: 1) beer, 2) wine, and 3) spirits. The six sub-categories for each major category are based on intensity of consumption per week. They are as follows:

1) for beer consumption:
   a. none
   b. less than six bottles per week
   c. six to 18 bottles per week
   d. 18 to 50 bottles per week
   e. 50 to 48 bottles per week
   f. more than 48 bottles per week

2) for wine consumption:
   a. none
   b. less than one bottle per week
   c. one to two bottles per week
   d. two to three bottles per week
e. four to six bottles per week
f. seven bottles or more per week

3) for spirits consumption:
   a. none
   b. less than 12 ounces per week
   c. 13 to 26 ounces per week
   d. 27 to 52 ounces per week
   e. 53 to 78 ounces per week
   f. greater than 78 ounces per week

These categories were chosen so as to give a significant representation of types of alcohol consumed and the intensity of consumption for the sample population.

Once these categories were tabulated, they were then disaggregated into one single variable representing total pure alcohol consumption per week. This was calculated by assigning a fixed percentage of alcohol content to each of the three categories, beer, wine, and spirits, and then calculating the percentage of pure alcohol consumption for all categories combined for each census tract. The variable represents percentage consumption per week, per head, for the sample group in each tract.

The final variable, drug consumption, includes the number of women who are currently taking the birth control pill and for how long. This was chosen because previous studies have indicated that women who smoke, drink, and take the birth control pill have a higher incidence risk of heart disease mortality than those who are non-smokers, non-drinkers, and do not take the pill. Other drugs include those prescribed for a particular illness and those which have not been prescribed, including illicit drugs.
A more helpful way to look at the variables to be utilized in this study, as well as the associations with mortality, can be found in the form of a model (Table 1).

**Hypotheses**

Based on the information derived from the literature review and the proposed model, the following hypotheses concerning the behaviour of risk factors and mortality rates are proposed.

1. In the city of Windsor, there is a significant correlation of coincidence between various risk factors.

2. In the city of Windsor, there is a significant relationship between the rates of death caused by cardiovascular disease and the following sets of risk factors:

   A) **Socioeconomic Factors**
      1. Income
      2. Occupational mix

   B) **Personal Lifestyle Factors**
      1. Tobacco consumption
      2. Alcohol consumption
      3. Drug consumption

   C) **Personal Physiological Factors**
      1. Age by cohort group
      2. Sex by cohort group
      3. Weight
      4. Ethnic distribution

3. In the city of Windsor, there is a significant relationship between the rates of death caused by respiratory disease and the following
sets of risk factors:

A) **Socioeconomic Factors**
   1. Income
   2. Occupational mix

B) **Personal Lifestyle Factors**
   1. Tobacco consumption
   2. Drug consumption
   3. Alcohol consumption

C) **Personal Physiological Factors**
   1. Age by cohort group
   2. Sex by cohort group
   3. Ethnic distribution

4. In the city of Windsor, there is a similarity in the spatial pattern of respiratory disease mortality and cardiovascular disease mortality.

**Data Collection and Methodology**

The data source used to derive the mortality figures for this study was death certificate information, compiled and made available by the Essex County Metropolitan Windsor Health Unit.

All relevant morbidity statistics have been compiled by a questionnaire survey, administered randomly throughout the city of Windsor, during the summer of 1978.

This study, investigating the risk factors involved with the mortalities of respiratory disease and heart disease in the city of Windsor, consisted of five stages.

**Stage 1.** The first stage of this research was to formulate a ques-
tionnaire survey. A preliminary questionnaire was set up and a pilot study carried out in May of 1978. For this trial run, 50 forms out of a total of 1,845 were administered by going door to door and dropping off the questionnaire for pick-up one week later. Upon completion and return, these were examined to determine whether or not the data from the questionnaires could be usefully tabulated. Based on the return of the 50 forms, changes were made to improve reporting of the data.

In the third week of May, 1978, the questionnaire was in its final form. The forms were administered by calculating 5% of the total households for each census tract. From this, the total per tract were divided by the number of enumeration areas in each census division. For example, if 5% of the total households in a census tract was equal to 56, and there were eight enumeration districts in that census tract, then in each enumeration area in that tract, seven questionnaires were dropped off for one week, and then collected. The addresses where each questionnaire was dropped were determined by selecting the first two numbers from a random numbers table, and using these numbers for the last two digits of that household address.

The completed and returned questionnaires were checked, edited and subsequently coded. Once coding was completed and checked again for errors, the data were punched on cards and subsequently stored on computer tape.

Using the Statistical Package for the Social Sciences (S.P.S.S.), a Frequencies sub-program was utilized to examine the frequency distribution of each response by Census Tract and Enumeration District. The data tabulated at the Enumeration District was not utilized for this study because the Standardized Mortality Rates (S.M.R.s) have been calculated
for census tracts. The reason for this is that the actual number of observed deaths was not sufficient for an accurate S.M.R. to be calculated at the smaller enumeration district level. Therefore, to keep the spatial units uniform for analytical purposes, the census tract areal units will be the basis for the study. Selected risk factors will subsequently be mapped in choropleth fashion and a detailed description of these maps will follow in the Data Analysis section. The questionnaire, in its final form, can be found in Appendix A.

**Stage 2.** The second stage of this research was to statistically determine, and cartographically depict, the spatial distribution patterns of the standardized death rates of cardiovascular disease and lung cancer during an eight year time period, 1970-1977, in the city of Windsor, from the death certificate data. For each mortality, the age, place of residence and sex in a particular census tract was recorded.

In order to allow direct comparisons to be made between populations diverse in both age structure and total number, standardized mortality rates (S.M.R.s) were used. For each census tract in Windsor, the total number of observed deaths were divided by the calculated expected deaths derived as explained below, and multiplied by 100 to yield, for each census tract, a specific disease S.M.R. value. Thus S.M.R.s above the norm (100) occurred where more than the expected number of deaths was recorded. An S.M.R. value which fell below the Provincial Norm (i.e., 100) indicated that observed deaths were fewer than had been expected.

Standardized mortality rates, of an interval scale of measurement, were calculated in the following manner:
1. Six divisions of each sex by age cohorts; 0-35, 36-45, 46-55, 66-75, and greater than 75 were made for each census tract.

2. From the death certificates, the total number of observed deaths occurring during the study period for each cohort group were recorded.

3. The mean annual number of deaths for each cohort group were then determined.

4. These mean values were then expressed as age standardized death rates (death rates/1,000) by dividing the mean values for each cohort group by its cohort group population. These values were then expressed as decimal fraction rates per 1,000.

5. The age standardized death rates/1,000 were then multiplied by the study units cohort group population to yield an expected death rate for each study unit.

6. The expected death rates for each unit cohort groupings were then summated, yielding an overall expected death rate for each study unit.

7. Standardized mortality rates (S.M.R.s) were then calculated for each study unit by means of the following formula:

   \[ S.M.R. = \frac{\text{Actual Deaths}}{\text{Expected Deaths}} \times 100 \]

   It should be noted that S.M.R.s based on less than 20 observed deaths may prove unreliable as a measure of the actual mortality rate (Howe, 1972).

   In this study, mortality data in the form of standardized mortality rates for each sex were obtained for each of Windsor's 45 census tracts for lung cancer and heart disease.
Scatter diagrams of these S.M.R.s were constructed in order to detect any discontinuities of the S.M.R. spread. These discontinuities served as determinants for the S.M.R. class boundaries in a series of choropleth maps which were constructed. In order to determine, for each tract's S.M.R. values, the occurrence of statistically significant deviations in the actual number of deaths observed compared to the expected number of deaths calculated, a test of standard deviation was used to determine the levels of significance at the 0.05 level. The null hypothesis (H₀) for this test was that the number of observed deaths did not differ significantly from the expected number of deaths. Each tract which exceeded the 0.05 level of significance, was mapped in choropleth fashion.

Stage 3. The third stage of this research was to statistically determine whether or not commonalities existed between the selected risk variables. This was done by means of a Multiple Factor Analysis utilizing the Statistical Package for the Social Sciences (S.P.S.S.). Factor analysis focuses on N variables and results are derived from an N x N matrix of relationships between all pairs of variables. It is important for classifying health information, because it emphasizes the relationships among the variables.

The first step in factor analysis is to produce a correlation matrix between sets of variables. The second step is to extract initial factors, that is, new variables based on the relationships within the data. The third step is to define these factors, carried out by transforming the variables into a new set of composite variables that are uncorrelated with each other. The final step is to rotate the initial factors to
terminal factors which are easier to identify. There may be several ways of defining the underlying structure of the data set which can be considered statistically significant (Kim, 1975).

Results obtained from factor analysis are only as good as the data. Non-linear relations between variables used in factor analysis result in factors that do not reflect the non-linearity. A transformation therefore is sometimes required which must be used when data normality is required for inferential factor analysis. When testing a theory or model, it is desirable to leave the data untransformed, where possible, to obtain a true picture of the actual situation observed (Taylor, 1977).

The health variables were subjected to a factor analysis to determine those health factors which could be useful for further analysis. The socio-economic variables were then added to determine their effect upon the model. The ethnic and demographic variables were then added to determine their overall effect upon the health and socio-economic factors.

A complete list of the variables entered into the factor analysis can be found in Appendix B.

Stage 4. The fourth stage of this research dealt with the etiologic risk factors, derived from the factor analysis, associated with heart disease.

Heart disease is hypothesized as a disease of multiple causes. The hypothesized risk factors associated with heart disease include occupational status, socio-economic status, ethnic and health factors.

By means of a stepwise Multiple Regression and Correlation Analysis, using the Statistical Package for the Social Sciences (S.P.S.S.), 14 in-
dependent variables for the male case and 16 independent variables for
the female case were separately tested for their degree of areal associa-
tion on the occurrence of heart disease mortalities in the city of Windsor
for the period 1970-1977. The independent census tract based variables,
generated by factor analysis included only certain demo-
graphic, ethnic, socio-economic, and lifestyle factors. Other relevant
environmental and physical factors were unobtainable or irrelevant for
a metropolitan level study. It was hoped that the results from these
analyses would indicate strong relationships between disease occurrence,
lifestyle and socio-economic risk factors. The independent variables
selected for this analysis, based on the results of the factor analysis,
are as follows:

<table>
<thead>
<tr>
<th>X Variables (male) by Census Tract</th>
<th>X Variables (female) by Census Tract</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1 Pure alcohol consumption per</td>
<td>X1 % Managerial</td>
</tr>
<tr>
<td>Capita</td>
<td></td>
</tr>
<tr>
<td>X2 % Smokers over 25 years of</td>
<td>X2 % Clerical</td>
</tr>
<tr>
<td>smoking</td>
<td></td>
</tr>
<tr>
<td>X3 % Underweight</td>
<td>X3 % Factory labour</td>
</tr>
<tr>
<td>X4 % High blood pressure</td>
<td>X4 % Income less than $12,000</td>
</tr>
<tr>
<td>X5 % Young smokers</td>
<td>X5 % Income $12,000 - $30,000</td>
</tr>
<tr>
<td>X6 % Managerial</td>
<td>X6 % Income greater than $30,000</td>
</tr>
<tr>
<td>X7 % Clerical</td>
<td>X7 % British and French</td>
</tr>
<tr>
<td>X8 % Farming</td>
<td>X8 % North European</td>
</tr>
<tr>
<td>X9 % Income $12,000 - $30,000</td>
<td>X9 % German</td>
</tr>
<tr>
<td>X10 % Income greater than $30,000</td>
<td>X10 % East European</td>
</tr>
<tr>
<td>X11 % British and French</td>
<td>X11 % Smokers less than 35 years of age</td>
</tr>
</tbody>
</table>
X12 % North European
X13 % East European
X14 % Male S.M.R.
X15 % On pill over 4 years who smoke
X16 Female S.M.R. value

At each step in the analysis, F-observed values for the partial regression coefficients were obtained and their associated F-critical values were calculated at the 0.05 level of significance where df1 = 1 and df2 = N-k-1. The null hypothesis tested was that no significant correlation at each step existed between the independent variables and the dependent variables. The multiple correlation coefficient (Ry 1, 2, 3, ... 14) for males and (Ry 1, 2, 3, ... 16) for females, and the multiple coefficient of determination statistics were extracted from the programme's summary tables. F-tests at the 0.05 level of significance with the null hypothesis Ry 1, 2, 3, ... 14 = 0 for males, and Ry 1, 2, 3, ... 16 = 0 for females, were run. The best regression equation was then determined and formally expressed for each case.

Stage 5. The fifth stage of this research dealt with the etiologic risk factors associated with lung cancer.

Cancer of the lung is hypothesized as a disease of a multi-etiologic nature. The hypothesized risk factors for lung cancer include occupational, as well as economic, ethnic and lifestyle factors.

By means of a stepwise multiple regression correlation analysis, utilizing the Statistical Package for the Social Sciences (S.P.S.S.), 17 independent variables for the male case and 17 independent variables for the female case were separately tested for their degree of areal associa-
tion for the occurrence of male and female lung cancer mortalities in the city of Windsor for the period 1970-1977. The independent variables, generated by a multiple factor analysis, regarded selected demographic, socio-economic, and lifestyle variables. As with the heart disease analysis, relevant environmental and physical factors were not considered for this study. The independent variables selected for these analyses were as follows:

<table>
<thead>
<tr>
<th>X Variables (male) by Census Tract</th>
<th>X Variables (female) by Census Tract</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1 Per capita alcohol consumption/ oz.</td>
<td>X1 % Managerial</td>
</tr>
<tr>
<td>X2 % Smokers greater than 25 years</td>
<td>X2 % Clerical</td>
</tr>
<tr>
<td>X3 % Overweight</td>
<td>X3 % Farming</td>
</tr>
<tr>
<td>X4 % Underweight</td>
<td>X4 % Mining</td>
</tr>
<tr>
<td>X5 % Smokers less that 35 years age</td>
<td>X5 % Factory labour</td>
</tr>
<tr>
<td>X6 % Smokers ages 36-55</td>
<td>X6 % Income less than $12,000</td>
</tr>
<tr>
<td>X7 % Managerial</td>
<td>X7 % Income greater than $30,000</td>
</tr>
<tr>
<td>X8 % Clerical</td>
<td>X8 % British</td>
</tr>
<tr>
<td>X9 % Farming</td>
<td>X9 % North European</td>
</tr>
<tr>
<td>X10 % Mining</td>
<td>X10 % German</td>
</tr>
<tr>
<td>X11 % Factory labour</td>
<td>X11 % East European</td>
</tr>
<tr>
<td>X12 % Income $12,000-$30,000</td>
<td>X12 % Smokers age less than 35</td>
</tr>
<tr>
<td>X13 % Income greater than $30,000</td>
<td>X13 % Heavy smokers</td>
</tr>
<tr>
<td>X14 % North European</td>
<td>X14 % Respiratory diseases</td>
</tr>
<tr>
<td>X15 % German</td>
<td>X15 % Women on pill over 4 years and smoking</td>
</tr>
<tr>
<td>X16 % East European</td>
<td>X16 SMR female</td>
</tr>
<tr>
<td>X17 SMR male cancer</td>
<td>X17 % Underweight</td>
</tr>
</tbody>
</table>
At each step in the analysis, F-observed values for the partial regression coefficients were obtained and their associated F-critical values were calculated at the 0.05 level of significance and df1 = 1, and df2 = N-k-1. The null hypothesis tested was that no significant correlation, at each step, existed between the independent variables and the dependent variables. The multiple correlation coefficient (RY 1, 2, 3, ... 17) for the males and (RY 1, 2, 3, ... 17) for the females, and the multiple coefficient of determination statistics were extracted from the programmes summary table. F-tests at the 0.05 level of significance with the null hypothesis Ry 1, 2, 3, ... 17 = 0 for males, and Ry 1, 2, 3, ... 17 = 0 for the females, were run. The best regression equation was then determined and formally expressed.
CHAPTER IV

DATA ANALYSIS

Interpretation of Survey Evidence

Introduction. A fundamental objective of this study was to identify possible causative factors associated with heart disease and lung cancer mortalities in the city of Windsor for the period 1970-1977.

The literature cites a wide variety of possible causative factors for high mortality experience due to cardiovascular disease and lung cancer. Those mentioned most often for ischaemic heart disease, types 410-414 inclusive of the W.H.O. International Statistical Clarification of Diseases, Injuries and Causes (I.C.D.), where there is a major reduction in the supply of blood to all or part of the heart, include age, altitude, blood type, cholesterol levels, climate, diabetes and glucose tolerance, E.C.G. abnormalities, living standards, obesity, physical activity, sex, smoking, and for females, the use of the oral birth control pill (Christou, 1978, quoting Strausser, 1972).

The W.H.O. defines cancer as a generic term for over 100 clinical entities manifesting malignant, cellular neoplasia in body tissue. The W.H.O. estimates that up to 85% of all cancer mortalities are the direct result of exposure to environmental factors. In many instances, self-inflicted habits such as overeating, smoking, excessive alcohol intake,
overexposure to sunlight and dangerous chemicals can be possible causative factors associated with lung cancer mortalities.

At this point it should be noted that in the case of heart disease, factors such as blood type, cholesterol levels, glucose tolerance, and the internal physiological characteristics can be very important genetic factors associated with heart disease mortality. This study, however, concerns itself with only those factors which fall under the broad heading of lifestyles. Without a detailed medical history of each of the respondents, the genetic characteristics cannot be utilized with any degree of accuracy. Also, it was not the intent of this writer to study the genetic factors but only to assess which, if any, lifestyle characteristics could possibly be associated with disease mortality.

In light of the foregoing, it was deemed imperative that new and much more specific data be generated. This was undertaken by means of a questionnaire survey, administered randomly throughout the city of Windsor by census tract, to try and identify patterns essentially of lifestyle differences that might indicate a correlation of coincidence that could then be compared to the mortality information.

Ideally, lifestyle characteristics could be derived from a detailed medical history of each of the respondents. However, since the medical histories were unobtainable, the lifestyle characteristics were obtained by the questionnaire survey. The questionnaire, while subjective in nature, still provides a reliable source of information to be utilized for the study of possible causative factors associated with disease mortality.

Some of the specific questions asked regarded the use of the oral
birth control pill, tobacco and alcohol consumption and nutritional practices. Also included were questions regarding the prevalence of disease, that is, asking what diseases respondents had, and what they were currently being treated for in terms of conditions of ill health. Additional questions included those dealing with socio-economic and personal physiological factors possibly associated with the prevalence of disease, that is, income, occupation, age and sex by cohort groups.

Since prevalence of disease is included in the data collected, a linkage may be derived between the questionnaire derived data with the standardized mortality data derived from the death certificates, for each census tract. One could also assume, with some degree of reliability, that variations in the patterns of lifestyle do exist by census tract and these variations are characteristic of individual neighbourhoods within each tract. This is a reasonable assumption as a census tract is defined as an area homogeneous in socio-economic status.

Wigle and Mao (1980) studied the associations between income level and mortality rates using census metropolitan areas as the areal units. They found that income was associated with factors such as education, occupation, and lifestyle, and each can independently contribute to the risk of disease mortalities. They caution that, although income level has been associated with mortality rates, chronic disease and disability may cause a decline in income due to loss of employment or occupational status. Thus, low income at the time of death can be a result of disease rather than a cause. Therefore, one must never interpret that income per se directly influences mortality rates.

On the other hand, lifestyles such as smoking and alcohol addiction,
have been directly implicated in disease etiology. The data base used here, therefore, can throw light on the connection between income, lifestyles and mortalities.

In order to depict the mortality incidence as represented by the standardized mortality rates, choropleth maps, using census tracts in Windsor as the study units, were constructed for each sex. Inferences drawn from the resultant map patterns must be viewed with caution because mortality data over a short period of time does not necessarily infer that these areas with high rates are areas where heart disease and lung cancer may be contracted. The standardized mortality rates (SMR's) have been compared to provincially derived norms and are summarized, for Windsor as a whole, in Table 2.

<table>
<thead>
<tr>
<th>Study Period</th>
<th>Heart Disease</th>
<th>Lung Cancer S.M.R.'s</th>
<th>Sex</th>
<th>Provincially Derived Norm</th>
</tr>
</thead>
<tbody>
<tr>
<td>1970-1977</td>
<td>X</td>
<td>112.48</td>
<td>Male</td>
<td>X</td>
</tr>
<tr>
<td>1970-1977</td>
<td>X</td>
<td>79.54</td>
<td>Male</td>
<td>X</td>
</tr>
<tr>
<td>1970-1977</td>
<td>X</td>
<td>135.05</td>
<td>Female</td>
<td>X</td>
</tr>
<tr>
<td>1970-1977</td>
<td>X</td>
<td>119.49</td>
<td>Female</td>
<td>X</td>
</tr>
</tbody>
</table>
For each sex elevated city-wide SMR's are evident, except for the male lung cancer rates, the calculations being based on the Provincially derived norm. It is also readily apparent that, in all city-wide cases, female standardized mortality rates exceed male SMR's by an approximate factor of 17% for heart disease and 34% for lung cancer mortalities.

In both the male and female instances for both disease types, consistently higher than expected SMR's are found in those census tracts within and around the older city core area, as well as in tracts situated near the downriver and south-west portion of the city. The higher female values, although based on a sample size only one-third as large as the male rates, was an unexpected result of this study, and its explanation is not attempted here, although the differential may relate to the error factor in a smaller sample, rather than a real difference. It can be stated that these tracts are all of a relatively low to middle income character composed of older housing with poverty pockets as identified by Oliver (1977). Those census tracts reporting lower than expected mortality rates are found, for the most part, in areas of upper income, newer suburban housing zones in the Southern and Eastern portions of the city, and therefore in areas where the population trends, in any case, to be younger providing fewer cases as a result.

The following sections of this chapter will deal with cartographic and verbal descriptions of the metropolitan SMR variations, as well as cartographic and verbal descriptions of the lifestyle variables collected from the questionnaire data. Results of, and subsequent inferences from the factor and regression analyses (Chapter 3: Methodology) will also be discussed.
It should be noted that all subsequent maps have been drawn in proportion to the population of each census tract, that is, they are in isodemographic form. The advantage of this mode of presentation is that reliability can be immediately related to the size of the shaded areas as these are proportional both to the total population and the sample size.

**Male Ischaemic Heart Disease Mortalities: Findings**

Male ischaemic heart disease mortalities for all tracts, expressed as standardized mortality rates for the eight year period 1970-1977, are summarized in Table 3.

Figure 3 shows very high S.M.R. values (exceeding +2 standard deviations) within two census divisions; census tract 31, on the near West side of Windsor with an S.M.R. value of 242.05, and census tract 8, on the West side of the city, with an S.M.R. value of 220.74. Other high areas of incidence (exceeding +1 standard deviation) include census tract 35, with an S.M.R. value of 172.21; tract 29, with an S.M.R. value of 171.82; tract 7, with an S.M.R. value of 167.56; and census tract 20, with an S.M.R. value of 164.60. These areas, with the exception of census tract 20, are all basically composed of older housing neighbourhoods situated near or along the Detroit river and are on the West side of the city. Census tracts reporting statistically significant higher than expected rates (Figure 4) corroborate the map patterns depicting tracts at or above +1 standard deviation with the exceptions of the exclusion of tracts 7 and 8, and the inclusion of tract 21. The resultant map patterns regarding the significance levels of higher than expected S.M.R. values indicate a clustered map pattern.
\textbf{TABLE 3}

\textbf{Male Ischaemic Heart Disease Mortalities: 1970-1977}

<table>
<thead>
<tr>
<th>Census Tract</th>
<th>Observed Deaths</th>
<th>Expected Deaths</th>
<th>Standardized Mortality Ratio</th>
<th>Standard Deviation = 44.46</th>
<th>Scatter Class Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>31</td>
<td>23</td>
<td>1.05</td>
<td>242.05</td>
<td>248.88 (+3)</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>7</td>
<td>0.34</td>
<td>220.74</td>
<td>202.41 (+2)</td>
<td>200</td>
</tr>
<tr>
<td>35</td>
<td>34</td>
<td>2.19</td>
<td>172.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>28</td>
<td>1.81</td>
<td>171.82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>6</td>
<td>0.39</td>
<td>167.56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>33</td>
<td>2.22</td>
<td>164.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>53</td>
<td>3.97</td>
<td>148.09</td>
<td>157.94 (+1)</td>
<td>150</td>
</tr>
<tr>
<td>32</td>
<td>40</td>
<td>3.00</td>
<td>147.79</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>4</td>
<td>0.30</td>
<td>145.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>26</td>
<td>2.04</td>
<td>141.57</td>
<td></td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>22</td>
<td>1.75</td>
<td>139.30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>15</td>
<td>1.21</td>
<td>137.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>20</td>
<td>1.62</td>
<td>136.48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>30</td>
<td>2.53</td>
<td>131.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>23</td>
<td>1.96</td>
<td>129.98</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>24</td>
<td>2.06</td>
<td>128.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>11</td>
<td>0.97</td>
<td>125.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>28</td>
<td>2.65</td>
<td>117.39</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>21</td>
<td>2.04</td>
<td>114.27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>21</td>
<td>3.91</td>
<td>113.54</td>
<td>113.48 mean</td>
<td></td>
</tr>
<tr>
<td>Census Tract #</td>
<td>Observed Deaths</td>
<td>Expected Deaths</td>
<td>Standardized Mortality Ratio</td>
<td>Standard Deviation = 44.46</td>
<td>Scatter Class Limits</td>
</tr>
<tr>
<td>---------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>------------------------------</td>
<td>-----------------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>10</td>
<td>21</td>
<td>3.91</td>
<td>113.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>32</td>
<td>3.14</td>
<td>113.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>6</td>
<td>0.58</td>
<td>113.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>26</td>
<td>2.56</td>
<td>112.48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>25</td>
<td>2.51</td>
<td>110.49</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>22</td>
<td>2.26</td>
<td>107.89</td>
<td></td>
<td></td>
</tr>
<tr>
<td>43</td>
<td>27</td>
<td>3.81</td>
<td>107.77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>16</td>
<td>1.70</td>
<td>103.63</td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>25</td>
<td>2.68</td>
<td>103.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>22</td>
<td>2.44</td>
<td>96.93</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>21</td>
<td>2.44</td>
<td>95.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>14</td>
<td>1.63</td>
<td>95.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>20</td>
<td>2.34</td>
<td>94.77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>16</td>
<td>1.91</td>
<td>92.45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>29</td>
<td>3.59</td>
<td>89.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>19</td>
<td>2.38</td>
<td>88.38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>35</td>
<td>4.47</td>
<td>86.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>16</td>
<td>2.08</td>
<td>85.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>12</td>
<td>1.69</td>
<td>78.66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>16</td>
<td>2.47</td>
<td>71.56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>11</td>
<td>1.93</td>
<td>62.94</td>
<td></td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>15</td>
<td>3.95</td>
<td>43.18</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

100% Provincial Norm

69.00 (-1)
<table>
<thead>
<tr>
<th>Census Tract #</th>
<th>Expected Deaths</th>
<th>Standardized Mortality Ratio</th>
<th>Scatter Class Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>101</td>
<td>1</td>
<td>1.53</td>
<td>38.44</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2.31</td>
<td>19.90</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.88</td>
<td>18.51</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

Source: Series A and Series B Census of Canada 1971, Windsor and Essex County Health Unit, Author
MORTALITIES OF ISCHAEMIC HEART DISEASE-MALE (1970-1977)

FIGURE 3
Areas delineating lower than expected values (at or below -1) standard deviation) include census tracts 11, with an S.M.R. value of 62.94; tract 42 with an S.M.R. value of 45.18; tract 101 with an S.M.R. value of 38.44; tract 6 with an S.M.R. value of 19.90, and tract 2 with an S.M.R. value of 18.51. These areas, with the exception of census tract 42, are situated in the Southern section of the city of Windsor. Tract 42 is located in the far Eastern portion of the city, along the Detroit river. All of these areas are composed of white collar workers, newer housing and are in the middle to upper income brackets. No census tracts, however, were described as having statistically significant lower than expected mortality rates.

Male Lung Cancer Mortalities: Findings

The male lung cancer mortalities for all tracts, expressed as standardized mortality rates for the period 1970-1977, in which 357 cases were reported, are summarized in Table 4.

Figure 5 depicts very high standardized mortality rates (exceeding +2 standard deviations) within two census divisions; tract 18, on the East side of the city with an S.M.R. value of 202.90, and census tract 8, on the West side of the city, with an S.M.R. value of 277.70. Other high areas of incidence (exceeding +1 standard deviation) include tracts 38 which has an S.M.R. value of 183.80, and census tract 27 which has an S.M.R. value of 178.50. These areas, with the exception of census tract 18, are all basically composed of neighbourhoods with older housing and are in the lower to middle income groups. They are also located on or near the Detroit river. Census tracts reporting statistically significant higher than expected rates of mortality (Figure 6) corroborate the map
<table>
<thead>
<tr>
<th>Census Tracts</th>
<th>Observed Mortalities</th>
<th>Expected Mortalities</th>
<th>Standardized Mortality Ratio</th>
<th>Windsor Standard Deviations = 53.36</th>
<th>Scatter Class Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>4</td>
<td>0.18</td>
<td>277.7</td>
<td>239.61 (+3)</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>19</td>
<td>1.17</td>
<td>202.9</td>
<td>186.25 (+2)</td>
<td>200</td>
</tr>
<tr>
<td>38</td>
<td>15</td>
<td>1.02</td>
<td>183.82</td>
<td>132.89 (+1)</td>
<td>150</td>
</tr>
<tr>
<td>27</td>
<td>7</td>
<td>0.49</td>
<td>178.57</td>
<td>114.95</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>12</td>
<td>1.23</td>
<td>121.95</td>
<td>115.03</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>15</td>
<td>1.63</td>
<td>111.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>17</td>
<td>1.88</td>
<td>111.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>9</td>
<td>1.00</td>
<td>112.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>18</td>
<td>2.01</td>
<td>111.92</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>7</td>
<td>0.79</td>
<td>110.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>43</td>
<td>13</td>
<td>1.55</td>
<td>104.84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>14</td>
<td>1.73</td>
<td>101.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>2</td>
<td>0.25</td>
<td>100.00</td>
<td>Province Norm = 100</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>2</td>
<td>0.25</td>
<td>100.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>11</td>
<td>1.47</td>
<td>93.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>17</td>
<td>2.37</td>
<td>89.66</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>6</td>
<td>0.87</td>
<td>86.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>15</td>
<td>2.27</td>
<td>82.59</td>
<td></td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>11</td>
<td>1.77</td>
<td>77.68</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>10</td>
<td>1.62</td>
<td>77.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>8</td>
<td>1.30</td>
<td>76.92</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>1.37</td>
<td>72.99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>10</td>
<td>1.74</td>
<td>71.84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Census Tracts</td>
<td>Observed Mortalities</td>
<td>Expected Mortalities</td>
<td>Standardized Mortality Ratio</td>
<td>Windsor Standard Deviation = 53.36</td>
<td>Scatter Class Limits</td>
</tr>
<tr>
<td>--------------</td>
<td>----------------------</td>
<td>---------------------</td>
<td>-----------------------------</td>
<td>-----------------------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>39</td>
<td>8</td>
<td>1.48</td>
<td>67.57</td>
<td></td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>9</td>
<td>1.67</td>
<td>67.36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>1.01</td>
<td>61.88</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>7</td>
<td>1.45</td>
<td>60.34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>10</td>
<td>2.12</td>
<td>58.96</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>5</td>
<td>1.08</td>
<td>57.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>6</td>
<td>1.37</td>
<td>54.74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>6</td>
<td>1.41</td>
<td>53.19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>0.74</td>
<td>50.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>10</td>
<td>2.49</td>
<td>50.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>7</td>
<td>1.76</td>
<td>49.71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>7</td>
<td>1.94</td>
<td>45.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>7</td>
<td>1.95</td>
<td>44.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>3</td>
<td>0.85</td>
<td>44.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>6</td>
<td>1.74</td>
<td>43.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>3</td>
<td>1.37</td>
<td>27.37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>2</td>
<td>1.06</td>
<td>23.58</td>
<td>26.17 (-1)</td>
<td>25.0</td>
</tr>
<tr>
<td>101</td>
<td>1</td>
<td>0.53</td>
<td>23.58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>1</td>
<td>0.73</td>
<td>17.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>0.83</td>
<td>15.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0.65</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>0.17</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Series A and Series B
Census of Canada, 1976
Windsor and Essex County Health Unit
Author
patterns depicting tracts at or above -1 standard deviation with the exceptions of the exclusion of tracts 27 and 38. The map patterns regarding the significance levels of higher than expected S.M.R. values indicate a random map pattern.

Areas delineating lower than expected values (at or below -1 standard deviation) include census tracts 20, which has an S.M.R. value of 23.58; census tract 101 which has an S.M.R. value of 17.12; census tract 6 which has an S.M.R. value of 15.06, and census tracts 2 and 7 which both have S.M.R. values of 0.00. These areas, with the exception of census tract 20, are situated in the Southern section of the city of Windsor. Tract 20 is located in the East side of Windsor, along the southern boundary of the city. All of these areas are basically composed of white collar workers, newer housing and are in the middle to upper income brackets. Census tracts reporting statistically significant lower than expected rates (Figure 6) corroborate the map patterns at or below -1 standard deviation. The map patterns regarding the significance levels of lower than expected S.M.R. values indicate a clustered pattern, with the exception of census tract 20.

**Lifestyle Characteristics of the Respondents in Aggregate: Male**

This section will examine the lifestyle characteristics of the male respondents in aggregate in order to identify patterns which may indicate a correlation of coincidence that could then be compared and tested against the mortality information. The following set of choropleth maps were generated in order to discern where patterns of high rates of lifestyle variables and patterns of high mortalities coexisted. The data for these maps were derived from the questionnaire survey which was carried
out during the summer of 1978. It must be noted that while these maps may have similar patterns to that of the mortality data, they may not themselves infer a statistical correlation between the patterns of lifestyle and mortality. Summary tables used to produce these maps may be found in the Appendix.

Figure 7, percentage of respondents with high blood pressure, coincides with tracts which have mortalities above the provincial norm for both disease types. The tracts with a high incidence of males with high blood pressure (over 40%) are 27, 33, 37, 20, 15, and 17. The tracts which have ischaemic heart mortalities above the provincial norm are 13, 8, 35, 29, 7, 20, 27, 13, 37, and 17. The tracts coinciding with high blood pressure incidence are 27, 13, 37, 17, and 20. The tracts which have lung cancer mortalities above the provincial norm are 18, 8, 27, 38, 12, 15, and 21. The only tract coinciding with the high blood pressure incidence is tract 27. Census tract 33 is the only tract which has a high incidence of high blood pressure and S.M.R. values below the provincial norm for both ischaemic heart and lung cancer mortalities.

Figure 8, percentage of heavy smoking males (over 30 cigarettes per day) also coincides with those tracts which have mortalities above the provincial norm for both disease types. The tracts with high incidence of males who are heavy smokers (over 60%) are 11, 1, 10, 5, 3, 8, 31, and 23. The census tracts reporting ischaemic heart disease mortalities above the provincial norm are 31, 8, 35, 29, 7, 20, 27, 13, 37, and 17. The areas coinciding with heavy smokers are tracts 31 and 8. Tracts reporting lung cancer mortalities above the provincial norm are 18, 8, 27, 38, 12, 15, 21, 32, 35, and 36. The only tract coinciding with heavy
smoking males and lung cancer mortalities is 8.

Census tracts 11, 1, 10, 5, 3, and 23 all report a high percentage of heavy smoking males but report S.M.R. values below the provincial norm for both cancer and heart disease mortality.

Looking back at the previous map (Figure 7), note that the percentage of respondents with high blood pressure does not correspond with this map (Figure 8). Although patterns of high incidence of high blood pressure are not found in the same areas as those with a high incidence of heavy smoking males, this in itself may not indicate that there is no correlation between heavy smoking and high blood pressure. This anomaly may indicate that there be no spatial correlation between the two but may not indicate that the two, as a whole, could not be otherwise correlated. Although this was an unexpected result of this study, its explanation is unfortunately beyond the scope of this investigation.

The next set of maps pertain to the percentage of male respondents who are overweight and underweight. The data for these maps were derived by collecting, from the questionnaire, weight and height information. These were then cross-tabulated to yield the percentage of respondents in various weight and height categories. A standard weight/height chart, provided by the Metropolitan Life Insurance Company, was then used to determine the average weight for each height category. The chart was broken down by small, medium and large frame build for each sex. Those percentages found to be over and under the average weight by height were then used as the data base for these variables.

Figure 9, percentage of male respondents who are overweight, also coincides with these tracts reporting high S.M.R. values above the pro-
vncial norm for ischaemic heart disease, but does not correspond closely
with the patterns of high S.M.R. values for lung cancer. Census tracts
39, 33, 37, 29, 4, 35, and 6 report fairly high percentages of overweight
males. Also, census tracts reporting ischaemic heart mortalities above
the provincial norm are 31, 8, 35, 29, 7, 20, 27, 15, 37, and 17. Those
tracts corresponding with areas which have reported high percentage of
overweight males are 35, 37, and 29.

Comparing the patterns of overweight males to high cancer mortalities
above the provincial norm, there is no discernable pattern coinciding with
each other. This may indicate that there is no correlation between the
two, but again this is not definite.

Figure 10, percentage of male respondents who are underweight, are
found in tracts reporting high S.M.R. values above the provincial norm for
lung cancer mortalities. Tracts 19, 26, 36, 32, 18, 2, and 6 report high
percentages of underweight males. Also, census tracts 19, 18, 38, 37, 35,
32, and 27 report high cancer mortalities above the provincial norm.
Areas reporting both high rates of cancer and high rates of underweight
males are 18, 19, and 36. The tracts reporting both high rates of heart
disease and percentage of males who are underweight are 19, and 32.

Looking back at the previous maps, note the percentage of the male
respondents who are overweight coincides closely with the patterns of high
rates of high blood pressure. Census tracts 29, 33, 37, and 39 have both
high rates of high blood pressure and ischaemic heart disease mortalities
above the provincial norm.

At this point, a definite pattern becomes evident, in that tracts 37,
31, 27, and 8 continuously show up in the figures coinciding variables.
PERCENTAGE OF RESPONDENTS UNDERWEIGHT-MALE

PERCENTAGE CLASS DISTRIBUTION

40% AND OVER
30-39%
20-29%
10-19%
0-9%
0%

SCALE
km
0
m
0
1
2

Detroit River
Detroit River
with incidence of mortality. These variables alone may not suggest, by any means, that there is a direct correlation between the variables and the disease types. However, these variables, added with other variables, may produce a more definite and reliable pattern, which then may be associated with disease mortality.

The next map (Figure 11) depicts areas where a high amount of alcohol consumption per capita in ounces per week occurred. This data was derived by collecting, from the questionnaire, questions pertaining to the consumption of beer, wine and spirits in bottles per week. These figures were then tabulated and transformed to yield the amount of pure alcohol consumed per capita per week in ounces. The transformations were calculated using the following formula:

\[
\frac{(N \times B \times A \times V)}{N} + \frac{(N \times W \times A \times V)}{N} + \frac{(N \times S \times A \times V)}{N}
\]

where: 
- \(N\) = number of respondents
- \(B\) = number of bottles of beer consumed
- \(W\) = number of litres of wine consumed
- \(S\) = number of ounces of spirits consumed
- \(V\) = volume of consumption
- \(A\) = alcohol content per serving

The values of \(A\), alcohol content, were broken down according to the standard alcohol content by volume for beer, wine, and spirits. The alcohol contents used in the equation were:

- Beer = .05% by volume
- Wine = .11% by volume
- Spirits = .40% by volume

Census tracts reporting a high volume of alcohol consumption per week
PER CAPITA ALCOHOL CONSUMPTION PER WEEK IN OUNCES

FIGURE 11
in ounces are 25, 100, 41, 17, 101, and 23. Also tracts reporting a fairly high volume of alcohol consumption are 35, 12, 24, 36, 16, 20, 1, 18, 26, 11, 22, 38, 34, and 42. Again, looking back at the previous maps (Figures 3 through 10), tracts 23, 35, 20, 36, and 38 report both high volume of alcohol consumption and ischaemic heart mortalities. Also, tracts 35, 36, 38, and 18 report high lung cancer mortalities and high volume of alcohol consumption.

There is also a pattern of coincidence between the incidence of high blood pressure and high alcohol consumption. Tracts 17 and 20 both coincide with one another for both maps (Figures 7 and 11). There also seems to be a fairly high incidence of heavy smokers in the same tracts reporting a high volume of alcohol consumption. Tracts coinciding with one another are 1, 11, 23, and 101. Although the pattern of heavy smoking appears to be clustered, the pattern of heavy drinking is more random in nature. These latter areas are ones which belong to the middle to upper income brackets.

The census tracts indicating a high incidence of overweight males do not coincide with those which have a high incidence of alcohol consumption. Unlike instances previously mentioned, there seems to be no pattern of coincidence between the overweight respondents and alcohol consumption.

The tracts depicting underweight respondents tend to coincide with high alcohol consumption. Tracts with similar patterns are 19, 17, 23, 24, 34, and 36.

At this point, a definite pattern became evident in that tracts 34, 23, 36, and 19 continuously show up as areas coinciding with the incidence
of disease mortality.

The next map (Figure 12) depicts areas with a high prevalence of diagnosed heart disease. The census tracts reporting high rates are 10, 5, 33, 37, 11, 12, and 30. Also, tracts reporting a high incidence of heart mortality include 10, 30, and 37. While these tracts depict both high mortality and high prevalence of diagnosed heart disease, the results were somewhat unexpected in that more areas were expected to have coincided with each other.

The areas depicting high prevalence rates of heart disease also correspond with tracts 33 and 37 on Figure 7, depicting the areas which have a high incidence of high blood pressure. Also tracts 11, 12, 10, and 6 coincide with the same tracts on Figure 8, depicting percent of heavy smoking males. Again tracts 33 and 37 correspond with the high tracts in Figure 9, depicting percent of overweight males.

The areas depicting high volume of consumption of alcohol do not coincide with the areas which have a high prevalence rate of diagnosed heart disease. With this fact in mind, one might assume, albeit with much caution, that these do not coincide because those with a diagnosed heart condition probably do not consume as much, if any, alcohol as those who do not have the condition.

In summary, it should be stressed that although these variables may suggest some distinct patterns, they do not suggest that these factors will be statistically significant causal factors for the occurrence of ischaemic heart disease and lung cancer mortalities within the city of Windsor. In order to determine, statistically, if these variables may be causal in nature they were tested using factor and regression analyses. The results
of these analyses are found in the following section.

**Ischaemic Heart Factor Variables Findings: Males**

**Factor Analysis.** This analysis was used to locate a smaller number of factors contained in the initial set of variables collected from the questionnaire data. The variables identified as being part of the underlying structure were then tested using the multiple regression and correlation technique. The findings of the multiple regression will be discussed in the next section.

The factor analysis tested three assumptions: 1) That the 12 lifestyle variables affect the occurrence of heart disease mortality and may be identified as significant factors; 2) That the four ethnic variables affect the mortality variable, and 3) That the socio-economic variables affect the mortality variables and may be identified as significant factors.

To test the first assumption, only the lifestyle variables were considered. The eigenvalues of the resultant factors express their relative importance. A rule of thumb that only those factors with eigenvalues greater than one are significant was employed. Three factors, from a total of five, were established in the initial run (Table 5). The next task was to label these new factors by determining the importance of the factor loadings of the variables in the rotated factor matrix, where a negative loading was as important as a positive one (Table 6). The label should be representative of all the significant variables chosen for that factor. The first factor identified was middle age male smokers, as variable 2, the number of years smoking being greater than 6, but less than 15; variable 4, the number of years of smoking greater than 25 years; variable 16
TABLE 5

Eigenvalues for the Lifestyle Factors

<table>
<thead>
<tr>
<th>Factor</th>
<th>Eigenvalues</th>
<th>% of Variation</th>
<th>Cumulative %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.07644</td>
<td>33.3</td>
<td>33.3</td>
</tr>
<tr>
<td>2</td>
<td>1.41758</td>
<td>22.7</td>
<td>56.0</td>
</tr>
<tr>
<td>3</td>
<td>1.23585</td>
<td>19.8</td>
<td>75.8</td>
</tr>
<tr>
<td>4</td>
<td>0.87205</td>
<td>14.0</td>
<td>89.8</td>
</tr>
<tr>
<td>5</td>
<td>0.63609</td>
<td>10.2</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Source: SPSS Factor Analysis, 1978 Questionnaire Data

being the percentage of heavy smokers and variable 26 being middle aged smokers (36-55 years old) were loaded significantly into this factor.

The second factor identified was the percentage of males with high blood pressure. The variable which loaded significantly into this factor was percentage of males on a diet, the percentage who were overweight, and the percentage of males with high blood pressure. The third factor identified was young smokers where the percentage of males with bronchitis and the percentage of young smokers both loaded significantly into this factor.

The percentage of young and middle age smokers and the percentage of overweight males with high blood pressure represent the most important factors identified as being characteristic of a possible causal relationship for the incidence of ischaemic heart disease mortalities of males in Windsor. Therefore, the first assumption was confirmed and accepted.
### TABLE 6

Quartimax Rotated Factor Matrix

<table>
<thead>
<tr>
<th>Variable</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>% on diet</td>
<td>-0.09117</td>
<td>0.43017</td>
<td>-0.10482</td>
</tr>
<tr>
<td>% respondents smoking 6-15 years</td>
<td>-0.80687</td>
<td>0.09420</td>
<td>-0.11089</td>
</tr>
<tr>
<td>Alcohol consumption</td>
<td>0.03405</td>
<td>-0.00113</td>
<td>0.09741</td>
</tr>
<tr>
<td>Smoking more than 25 years</td>
<td>0.71215</td>
<td>-0.02717</td>
<td>-0.30879</td>
</tr>
<tr>
<td>Overweight</td>
<td>-0.17526</td>
<td>0.55416</td>
<td>-0.09661</td>
</tr>
<tr>
<td>Underweight</td>
<td>0.12197</td>
<td>0.08088</td>
<td>0.14361</td>
</tr>
<tr>
<td>Heavy smoking</td>
<td>0.45110</td>
<td>-0.43046</td>
<td>0.13212</td>
</tr>
<tr>
<td>Bronchitis</td>
<td>0.17768</td>
<td>-0.00302</td>
<td>0.42203</td>
</tr>
<tr>
<td>High blood pressure</td>
<td>0.18200</td>
<td>0.84763</td>
<td>0.01650</td>
</tr>
<tr>
<td>Young smokers</td>
<td>0.02408</td>
<td>-0.20114</td>
<td>0.96328</td>
</tr>
<tr>
<td>Middle age smokers</td>
<td>0.51772</td>
<td>0.04321</td>
<td>0.12511</td>
</tr>
<tr>
<td>Old smokers</td>
<td>-0.19941</td>
<td>0.16881</td>
<td>-0.19911</td>
</tr>
</tbody>
</table>

Source: SPSS Factor Analysis, 1978 Questionnaire Data
The second run tested the second assumption with only the ethnic variables included. The eigenvalues provided two significant factors (Table 7). The first factor identified was percent North European where the percentage of North Europeans and the percentage of Germans loaded significantly into this factor. The second factor identified was percent East European, as this variable was the only one which loaded significantly into this second factor. Therefore, the ethnic variables were also identified as being characteristic as possible causal factors associated with the incidence of ischaemic heart mortality in males in Windsor.

The socio-economic variables were used to test the third assumption. The eigenvalues provided two significant factors (Table 8). The first factor identified was middle to upper income white collar workers where percent clerical workers, income between $12,000 and $30,000 and income greater than $30,000 were loaded significantly into this factor. The second factor identified was white collar workers where percent of managerial and percent clerical workers were loaded significantly into this factor. With these new factors in mind, regression analysis was employed to test the null hypothesis which states that no significant correlation exists between the style, ethnic, and socio-economic variables and the standardized mortality rates of ischaemic heart disease for males in the city of Windsor. The best fit model was used to test the hypothesis and the results of this analysis are summarized in the following paragraphs.

Regression Analysis: Findings

The results of the regression analysis are found in Table 9. The overall regression equation accounted for 32.87% of the variations explained for the incidence of ischaemic heart disease mortalities in males.
### TABLE 7a

Eigenvalues for the Ethnic Factors

<table>
<thead>
<tr>
<th>Factor</th>
<th>Eigenvalues</th>
<th>% of Variation</th>
<th>Cumulative %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.47303</td>
<td>39.5</td>
<td>39.5</td>
</tr>
<tr>
<td>2</td>
<td>1.21950</td>
<td>34.6</td>
<td>74.1</td>
</tr>
<tr>
<td>3</td>
<td>0.83858</td>
<td>22.8</td>
<td>96.9</td>
</tr>
<tr>
<td>4</td>
<td>0.65454</td>
<td>3.1</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Source: SPSS Factor Analysis, 1978 Questionnaire Data

### TABLE 7b

Quartimax Rotated Factor Matrix

<table>
<thead>
<tr>
<th>Variable</th>
<th>Factor 1</th>
<th>Factor 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>British and French</td>
<td>0.10884</td>
<td>-0.53482</td>
</tr>
<tr>
<td>North European</td>
<td>0.28722</td>
<td>0.24791</td>
</tr>
<tr>
<td>German</td>
<td>-0.42417</td>
<td>-0.04104</td>
</tr>
<tr>
<td>East European</td>
<td>0.62330</td>
<td>-0.06298</td>
</tr>
</tbody>
</table>

Source: SPSS Factor Analysis, 1978 Questionnaire Data
### TABLE 8a

**Eigenvalues for the Socio-Economic Factors**

<table>
<thead>
<tr>
<th>Factor</th>
<th>Eigenvalues</th>
<th>% of Variation</th>
<th>Cumulative %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.14666</td>
<td>46.8</td>
<td>46.8</td>
</tr>
<tr>
<td>2</td>
<td>1.40440</td>
<td>30.6</td>
<td>77.5</td>
</tr>
<tr>
<td>3</td>
<td>0.58204</td>
<td>12.7</td>
<td>90.2</td>
</tr>
<tr>
<td>4</td>
<td>0.45128</td>
<td>9.8</td>
<td>100.0</td>
</tr>
</tbody>
</table>

### TABLE 8b

**Quartimax Rotated Factor Matrix**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Factor 1</th>
<th>Factor 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Managerial</td>
<td>-0.20086</td>
<td>0.35766</td>
</tr>
<tr>
<td>Clerical</td>
<td>0.65574</td>
<td>0.20217</td>
</tr>
<tr>
<td>Farming</td>
<td>0.12261</td>
<td>0.59840</td>
</tr>
<tr>
<td>Mining</td>
<td>-0.05978</td>
<td>-0.04932</td>
</tr>
<tr>
<td>Factory Labour</td>
<td>0.09365</td>
<td>0.02312</td>
</tr>
<tr>
<td>Income less than $12,000</td>
<td>0.07658</td>
<td>-0.17522</td>
</tr>
<tr>
<td>Income $12,000 to $30,000</td>
<td>0.64350</td>
<td>-0.68004</td>
</tr>
<tr>
<td>Income greater than $30,000</td>
<td>0.58165</td>
<td>-0.30134</td>
</tr>
</tbody>
</table>

*Source: SPSS Factor Analysis, 1978 Questionnaire Data*
### TABLE 9
The Relationship between Male Ischaemic Heart Mortalities and Selected Lifestyle, Ethnic and Socio-Economic Variables by Census Tract

<table>
<thead>
<tr>
<th>Variable</th>
<th>'r' (Simple Correlation Coefficient)</th>
<th>$B_r$ (Partial Regression Coefficient)</th>
<th>F-Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>X5- Smokers less than 35 years of age</td>
<td>0.28434</td>
<td>0.9384417</td>
<td>2.522*</td>
</tr>
<tr>
<td>X6- Managerial</td>
<td>-0.18570</td>
<td>-0.7937813</td>
<td>2.617*</td>
</tr>
<tr>
<td>X8- Farming</td>
<td>-0.21085</td>
<td>-5.501324</td>
<td>2.569*</td>
</tr>
<tr>
<td>X3- Underweight</td>
<td>0.20692</td>
<td>1.153670</td>
<td>4.015*</td>
</tr>
<tr>
<td>X7- Clerical</td>
<td>0.17700</td>
<td>1.228433</td>
<td>2.795*</td>
</tr>
<tr>
<td>X2- Smoked more than 25 years</td>
<td>-0.18189</td>
<td>-0.53737</td>
<td>2.551*</td>
</tr>
</tbody>
</table>

$r$ = Simple Correlation Coefficient  
$B_r$ = Simple Regression Coefficient  
$F$ = Partial F-Statistic for Significance  
$X^2$ = Coefficient of Determination = 0.2871  
$F$ Critical Value = 2.18  
* = Significant at the 0.05 Significance Level

Regression Equation:  
$$Y_c = 84.86059 + 0.9384X_5 - 0.7937X_6 - 5.5013X_8 + 1.1537X_3 + 1.2284X_7 - 0.5374X_2$$
in Windsor. The first six variables entered into the equation accounted for 28.71% of the variation for the disease occurrence. These independent variables, in order of entry are: \(X_5\) (% of smokers less than 35 years of age); \(X_6\) (% of males employed in a managerial situation); \(X_8\) (% of males employed in farming); \(X_3\) (% of male respondents underweight); \(X_7\) (% employed in a clerical situation), and \(X_2\) (% of males who have smoked for 25 years or more). As a total, these variables were found to be significant at the 0.05 level. However, only the first four variables entered had significant partial F-values. The first variable entered, \(X_5\), had a strong positive B value indicating that young smokers may be very susceptible to heart disease. The next two variables entered, \(X_6\) and \(X_8\) (the occupational variables) had significant negative B values, indicating an inverse relationship between the high stress occupational variables and heart disease. This fact is surprising in that the literature cites cases where stress related occupations have correlated with the occurrence of heart disease. The next variable entered, \(X_3\) (% underweight) has a significant positive B value indicating a possible relationship between weight and heart disease. In this case, however, the underweight males may be an effect rather than a possible cause of heart disease. The final two variables, \(X_7\) and \(X_2\) were by themselves not significant, but were found to be significant in terms of the overall equation. All of the above mentioned variables were tested using the F-test to determine their significance as independent variables. All were found to be above the F-critical value of 2.18 at the 0.05 level. The null hypothesis, as stated previously, was rejected and the alternate hypothesis was accepted because the F-values were above the F-critical value. This, however, does not mean that these
variables alone are immediate causes, but only that they are possible factors associated with the occurrence of heart disease mortality in males in Windsor.

Lung Cancer Factor Variables Findings: Male

Factor Analysis. This analysis, as with the ischaemic heart, was used to locate a smaller number of factors contained in the initial set of variables collected from the questionnaire data. The variables identified as being part of the underlying structure were then tested using the multiple regression and correlation technique. The findings of the multiple regression will be discussed in the next section.

The factor analysis again tested three assumptions: 1) That the 14 variables affect lung cancer mortalities and may be identified as significant factors; 2) That the four ethnic variables affect the mortality variable and may be identified as significant factors, and 3) That nine socio-economic variables may be identified as new significant factors.

To test the first assumption, the lifestyles variables were considered. The eigenvalues of the resultant factors provide six factors of which three were significant (Table 10). The first factor identified was middle age heavy smoking males where variable 4 (3% of respondents who smoke for more than six, but less than 15 years), variable 16 (% of heavy smokers), and variable 26 (middle aged respondents who smoke) were all loaded significantly into this first factor. The second factor identified was underweight respondents with respiratory diseases as these two variables were loaded significantly into this factor. The third factor identified young smokers as these two variables loaded significantly into this factor. Therefore, the smoking variables, despite the age of the respondents,
TABLE 10
Eigenvalues for the Lifestyle Factors
Male Lung Cancer

<table>
<thead>
<tr>
<th>Factor</th>
<th>Eigenvalue</th>
<th>% of Variation</th>
<th>Cumulative %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.10070</td>
<td>28.2</td>
<td>28.2</td>
</tr>
<tr>
<td>2</td>
<td>1.67973</td>
<td>22.5</td>
<td>50.7</td>
</tr>
<tr>
<td>3</td>
<td>1.19618</td>
<td>16.0</td>
<td>66.7</td>
</tr>
<tr>
<td>4</td>
<td>0.92653</td>
<td>12.4</td>
<td>79.2</td>
</tr>
<tr>
<td>5</td>
<td>0.88123</td>
<td>11.8</td>
<td>91.0</td>
</tr>
<tr>
<td>6</td>
<td>0.67257</td>
<td>9.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Source: SPSS Factor Analysis, 1978 Questionnaire Data

Quartimax Rotated Factor Matrix

<table>
<thead>
<tr>
<th>Variable</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diet</td>
<td>-0.21172</td>
<td>-0.23271</td>
<td>-0.12493</td>
</tr>
<tr>
<td>6-15 years of smoking</td>
<td>-0.70544</td>
<td>-0.09416</td>
<td>0.00362</td>
</tr>
<tr>
<td>Alcohol consumption</td>
<td>0.06088</td>
<td>0.00209</td>
<td>0.06047</td>
</tr>
<tr>
<td>&gt; 25' years of smoking</td>
<td>0.51540</td>
<td>0.10603</td>
<td>-0.38147</td>
</tr>
<tr>
<td>Overweight</td>
<td>-0.32208</td>
<td>0.00366</td>
<td>-0.14454</td>
</tr>
<tr>
<td>Underweight</td>
<td>0.01293</td>
<td>0.92283</td>
<td>0.04045</td>
</tr>
<tr>
<td>Heavy smoking</td>
<td>0.64875</td>
<td>-0.28016</td>
<td>0.20857</td>
</tr>
<tr>
<td>Diabetic</td>
<td>0.04004</td>
<td>0.36049</td>
<td>0.03449</td>
</tr>
<tr>
<td>Bronchitis</td>
<td>0.10702</td>
<td>0.16938</td>
<td>0.36611</td>
</tr>
<tr>
<td>Pneumonia</td>
<td>-0.04309</td>
<td>0.16325</td>
<td>-0.04419</td>
</tr>
<tr>
<td>Respiratory Diseases</td>
<td>-0.00155</td>
<td>0.62176</td>
<td>0.05121</td>
</tr>
<tr>
<td>Young smokers</td>
<td>0.13279</td>
<td>0.11541</td>
<td>-0.89780</td>
</tr>
<tr>
<td>Middle age smokers</td>
<td>0.55585</td>
<td>0.23866</td>
<td>0.05931</td>
</tr>
<tr>
<td>Old smokers</td>
<td>-0.24951</td>
<td>0.09977</td>
<td>-0.27575</td>
</tr>
</tbody>
</table>

Source: SPSS Factor Analysis, 1978 Questionnaire Data
established a pattern as being the most important factors characteristic of lung cancer mortalities. Therefore, the first assumption was confirmed.

The second assumption tested used only the ethnic variables. The eigenvalues provided three significant factors (Table 11). The first factor identified percent North European where this variable was the only one which loaded significantly. The second factor identified percent German where this variable is the only one which was loaded significantly. The third factor identified percent East European where this variable was the one which loaded significantly. Therefore, the ethnic variables also appear to be important in the spatial distribution of lung cancer mortalities in Windsor.

The socio-economic variables were used to test the third assumption. The eigenvalues provided four factors, three of which were found to be significant (Table 12). The first factor identified blue collar workers where percent farming and percent factory labour were loaded significantly into this factor. The second factor identified was lower to middle income where variables depicting income were loaded significantly into this factor. The third factor identified white collar workers where variable 30 (% managerial) and variable 31 (% clerical) were loaded significantly into this factor. The middle to upper income variables and blue and white collar variables establish a pattern as being the most important socio-economic factors characteristic of lung cancer mortalities. Therefore, the third assumption was also accepted.

With these new factors in mind, regression analysis was used to establish which of the variables derived from the factor loadings will best fit the model. The null hypothesis tested states that there will be no significant correlation between these selected lifestyles, ethnic and
<table>
<thead>
<tr>
<th>Factor</th>
<th>Eigenvalue</th>
<th>Pct of Var</th>
<th>Cum Pct</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.38001</td>
<td>27.6</td>
<td>27.6</td>
</tr>
<tr>
<td>2</td>
<td>1.21165</td>
<td>24.2</td>
<td>51.8</td>
</tr>
<tr>
<td>3</td>
<td>1.05717</td>
<td>21.1</td>
<td>73.0</td>
</tr>
<tr>
<td>4</td>
<td>0.77430</td>
<td>15.5</td>
<td>88.5</td>
</tr>
<tr>
<td>5</td>
<td>0.57686</td>
<td>11.5</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Quartimax Rotated Factor Matrix

<table>
<thead>
<tr>
<th>Variable</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>British French</td>
<td>-0.18712</td>
<td>-0.23147</td>
<td></td>
</tr>
<tr>
<td>North European</td>
<td>0.63587</td>
<td>0.13449</td>
<td>-0.14635</td>
</tr>
<tr>
<td>German</td>
<td></td>
<td>-0.51510</td>
<td>0.31797</td>
</tr>
<tr>
<td>East European</td>
<td>0.27988</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
TABLE 12

Eigenvalues for the Socio-Economic Variables

Male Lung Cancer

<table>
<thead>
<tr>
<th>Factor</th>
<th>Eigenvalue</th>
<th>Pct of Var</th>
<th>Cum Pct</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.61033</td>
<td>41.7</td>
<td>41.7</td>
</tr>
<tr>
<td>2</td>
<td>1.79846</td>
<td>28.7</td>
<td>70.4</td>
</tr>
<tr>
<td>3</td>
<td>1.12943</td>
<td>18.0</td>
<td>88.5</td>
</tr>
</tbody>
</table>

Source: SPSS Factor Analysis, 1978 Questionnaire Data

Quartimax Rotated Factor Variables

<table>
<thead>
<tr>
<th>Variables</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Managerial</td>
<td>-0.92968</td>
<td>0.04954</td>
<td>-0.34115</td>
</tr>
<tr>
<td>Clerical</td>
<td>-0.04022</td>
<td>0.13396</td>
<td>0.96522</td>
</tr>
<tr>
<td>Farming</td>
<td>0.30502</td>
<td>0.41572</td>
<td>-0.17971</td>
</tr>
<tr>
<td>Mining</td>
<td>0.03802</td>
<td>-0.52194</td>
<td>-0.02169</td>
</tr>
<tr>
<td>Factory Labour</td>
<td>0.92689</td>
<td>-0.12712</td>
<td>-0.18415</td>
</tr>
<tr>
<td>Income &lt; $12,000</td>
<td>0.29399</td>
<td>-0.83848</td>
<td>0.02012</td>
</tr>
<tr>
<td>Income $12,000-$30,000</td>
<td>-0.08967</td>
<td>0.94624</td>
<td>0.13070</td>
</tr>
<tr>
<td>Income &gt; $30,000</td>
<td>-0.32258</td>
<td>-0.06266</td>
<td>-0.22207</td>
</tr>
</tbody>
</table>

Source: SPSS Factor Analysis, 1978 Questionnaire Data
socio-economic variables and the incidence of male lung cancer mortalities in Windsor.

Regression Analysis Male Lung Cancer: Findings

The results of the regression analysis are found in Table 13. The overall regression equation accounted for 47.68% of the variation explained for the incidence of male lung cancer mortalities. The first five variables entered accounted for 40.10% of the variation. These independent variables in order of entry are: X5 (% of smokers less than 35 years of age), X14 (% of respondents who are North European), X7 (% of respondents employed in managerial), X9 (% of respondents farming), and X15 (% of respondents who are German). As a total, these variables were found to be significant at the 0.05 level of significance. However, the first three variables are the only ones which had significant partial F-values. The first variable entered, X5, had a strong positive B value, indicating a significant positive relationship between young smokers and lung cancer mortality. The next variable entered, X14, had a significant inverse B value, indicating an inverse relationship between the ethnic variable and lung cancer mortalities. The third variable entered, X7, also had a significant negative B value, indicating a significant inverse relationship between the white collar variable and male lung cancer mortalities. The next two variables entered, X9 and X15, were not found to be significant when the partial F-values were considered, but were significant when added to the other three variables when contributing to the overall significance of the equation. All of the above mentioned variables were tested with the F-test to determine their significance as independent variables. All were found to be above the F-critical value of 2.45 at the 0.05 level.


TABLE 13

Relationship Between Male Lung Cancer Mortalities and Selected Lifestyle, Ethnic and Socio-Economic Variables by Census Tract

<table>
<thead>
<tr>
<th>Variable</th>
<th>r Simple Correlation Coefficient</th>
<th>B (Partial Regression Coefficient)</th>
<th>F-Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>X5 Young smokers</td>
<td>0.4532</td>
<td>2.4233</td>
<td>14.472*</td>
</tr>
<tr>
<td>X14 North Europeans</td>
<td>-0.2556</td>
<td>-2.5175</td>
<td>8.075*</td>
</tr>
<tr>
<td>X7 Managerial</td>
<td>-0.1545</td>
<td>-1.0213</td>
<td>3.769*</td>
</tr>
<tr>
<td>X9 Farming</td>
<td>-0.0994</td>
<td>-5.4932</td>
<td>6.239*</td>
</tr>
<tr>
<td>X15 German</td>
<td>-0.1834</td>
<td>-3.0917</td>
<td>5.222*</td>
</tr>
</tbody>
</table>

r = simple correlation coefficient
B = partial regression coefficient
F = partial F-statistic for significance
R² = coefficient of determination = .4010
F₀ = critical value = 2.45
* = significant at the 0.05 level of significance

Regression equation: (Yc) = 90.232 + 2.423 (X5) - 2.517 (X14) - 1.021 (X7) - 5.493 (X9) - 3.091 (X15)
The null hypothesis, as stated previously, was rejected and the alternate hypothesis was accepted because the F-values were above the F-critical values. Although these variables were found to be significant contributors to the equation, this result does not suggest that these variables are definite contributing factors but only that they are associated with the occurrence of male lung cancer mortality in the city of Windsor.

**Female Ischaemic Heart Mortality: Findings**

Female ischaemic heart disease mortalities for all tracts, expressed as standardized mortality rates for the eight year period 1970-1977, are summarized in Table 14.

The female S.M.R. values, as has been previously mentioned, surpass their male S.M.R. counterparts by an approximate value of 17%. Figure 13 illustrates higher than expected S.M.R. values (exceeding +2 standard deviations) in census areas 32, 8, 28, and 7. Values exceeding +1 standard deviation are found in tracts 21 and 35. These areas are highly similar to the high male rate zones in geographic situation, lifestyle characteristics and socio-economic status. These resultant patterns of higher than expected S.M.R. areas indicate a clustered map pattern. Census tracts reporting statistically significant higher than expected rates (Figure 14) corroborate the map patterns depicting tracts at or above +1 standard deviation with the exceptions of the exclusion of tracts 7 and 8, and the addition of tract 9.

Areas delineating lower than expected values (at or below -1 standard deviation) include census tracts 2, 4, 15, 5, and 6. All of these areas are found in South Windsor, which is basically comprised of newer housing
<table>
<thead>
<tr>
<th>Census Tract #</th>
<th>Observed Deaths</th>
<th>Expected Deaths</th>
<th>Standardized Mortality Ratio</th>
<th>Standard Deviation = 63.46</th>
<th>Scatter Class Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>32</td>
<td>17</td>
<td>0.66</td>
<td>285.36</td>
<td></td>
<td>327.44 (+3)</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>0.07</td>
<td>279.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>15</td>
<td>0.61</td>
<td>269.84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>0.08</td>
<td>267.66</td>
<td></td>
<td>263.97 (+2)</td>
</tr>
<tr>
<td>21</td>
<td>18</td>
<td>0.92</td>
<td>215.32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>11</td>
<td>0.56</td>
<td>214.84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>6</td>
<td>0.32</td>
<td>203.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>14</td>
<td>0.76</td>
<td>202.20</td>
<td></td>
<td>200.50 (+1)</td>
</tr>
<tr>
<td>24</td>
<td>13</td>
<td>0.80</td>
<td>179.36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>12</td>
<td>0.74</td>
<td>178.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>7</td>
<td>0.44</td>
<td>176.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>0.45</td>
<td>171.30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>9</td>
<td>0.58</td>
<td>170.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>10</td>
<td>0.69</td>
<td>168.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>7</td>
<td>0.49</td>
<td>156.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>15</td>
<td>1.09</td>
<td>151.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>8</td>
<td>0.58</td>
<td>151.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>6</td>
<td>0.45</td>
<td>146.32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>0.07</td>
<td>143.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>12</td>
<td>0.94</td>
<td>140.80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>3</td>
<td>0.23</td>
<td>140.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>43</td>
<td>11</td>
<td>0.89</td>
<td>136.51</td>
<td></td>
<td>( \bar{x} = 137.03 )</td>
</tr>
<tr>
<td>Census Tract #</td>
<td>Observed Deaths</td>
<td>Expected Deaths</td>
<td>Standardized Mortality Ratio</td>
<td>Standard Deviation = 63.46</td>
<td>Scatter Class Limits</td>
</tr>
<tr>
<td>--------------</td>
<td>----------------</td>
<td>----------------</td>
<td>-----------------------------</td>
<td>-----------------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>26</td>
<td>9</td>
<td>0.77</td>
<td>129.45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>8</td>
<td>0.71</td>
<td>123.91</td>
<td></td>
<td></td>
</tr>
<tr>
<td>101</td>
<td>3</td>
<td>0.27</td>
<td>120.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>7</td>
<td>0.65</td>
<td>118.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>7</td>
<td>0.67</td>
<td>115.08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>9</td>
<td>0.87</td>
<td>114.30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>5</td>
<td>0.54</td>
<td>102.30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>8</td>
<td>0.90</td>
<td>93.18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>4</td>
<td>0.45</td>
<td>97.36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>3</td>
<td>0.34</td>
<td>95.74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>8</td>
<td>0.94</td>
<td>94.24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>4</td>
<td>0.49</td>
<td>89.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>9</td>
<td>1.16</td>
<td>86.08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>3</td>
<td>0.40</td>
<td>82.38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>4</td>
<td>0.44</td>
<td>77.26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>4</td>
<td>0.58</td>
<td>76.26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>1</td>
<td>0.14</td>
<td>75.53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>0.30</td>
<td>73.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>0.47</td>
<td>70.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>0.83</td>
<td>66.91</td>
<td></td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>2</td>
<td>0.42</td>
<td>47.35</td>
<td></td>
<td>50</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>0.50</td>
<td>41.02</td>
<td></td>
<td>25</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>0.48</td>
<td>22.98</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Provincial Norm 73.57 (-1)
and are in the middle to upper income brackets. Again, these patterns are similar to their male counterparts and indicate a clustered map pattern. No census tracts had statistically significant lower than expected S.M.R. values.

Female Lung Cancer Mortalities: Findings

According to studies recently completed by the Canadian Cancer Society in Sarnia and London, female lung cancer is much higher overall, than male lung cancer (Stavraky, 1978). The reasons for this as yet have not been determined.

Female lung cancer mortalities are summarized in Table 15. These data represent census tracts in Windsor which have high or low S.M.R. values for lung cancer. These mortalities are also illustrated in Figure 15. Standardized mortality rates, based on less than 20 observed deaths, may prove to be unreliable as a measure of the actual mortality (Howe, 1972).

Figure 15 depicts very high S.M.R. values (exceeding +2 standard deviations) within one census division, census tract 12, located in the central portion of Windsor, which has an S.M.R. value of 912.40. Other areas illustrating high S.M.R. values (exceeding +1 standard deviation) include tracts 19, with an S.M.R. value of 409.84, tract 101 with an S.M.R. value of 353.70, and tract 8 which has an S.M.R. value of 347.20. These areas are very similar to the male pattern, but the rates are approximately 34% higher than the male rates. This may be due to the fact that the actual number of observed female deaths were lower than the male observed deaths, thus the S.M.R. values may be elevated because of the sample size. The higher rates, however, corroborate studies completed by the Canadian Cancer
<table>
<thead>
<tr>
<th>Census Tracts</th>
<th>Observed Mortalities</th>
<th>Expected Mortalities</th>
<th>Standardized Mortality Ratio</th>
<th>Windsor Standard Deviations = 151.33</th>
<th>Scatter Class Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>2</td>
<td>0.037</td>
<td>912.41</td>
<td>422.15 (+2)</td>
<td>500</td>
</tr>
<tr>
<td>19</td>
<td>2</td>
<td>0.061</td>
<td>409.84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>101</td>
<td>3</td>
<td>0.106</td>
<td>353.77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>0.036</td>
<td>347.22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>8</td>
<td>0.528</td>
<td>189.39</td>
<td></td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>8</td>
<td>0.529</td>
<td>189.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>0.209</td>
<td>179.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>4</td>
<td>0.294</td>
<td>170.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>6</td>
<td>0.464</td>
<td>161.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>2</td>
<td>0.156</td>
<td>160.26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>3</td>
<td>0.247</td>
<td>151.82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>43</td>
<td>4</td>
<td>0.355</td>
<td>140.85</td>
<td></td>
<td>150</td>
</tr>
<tr>
<td>32</td>
<td>4</td>
<td>0.370</td>
<td>135.14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>4</td>
<td>0.408</td>
<td>122.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>2</td>
<td>0.217</td>
<td>115.21</td>
<td></td>
<td>119.49</td>
</tr>
<tr>
<td>18</td>
<td>2</td>
<td>0.230</td>
<td>108.69</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>0.359</td>
<td>104.45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>3</td>
<td>0.379</td>
<td>98.94</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>3</td>
<td>0.384</td>
<td>97.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>3</td>
<td>0.398</td>
<td>94.22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>3</td>
<td>0.402</td>
<td>93.28</td>
<td></td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>2</td>
<td>0.268</td>
<td>93.28</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Provincial Norm = 100
<table>
<thead>
<tr>
<th>Census Tracts</th>
<th>Observed Mortalities</th>
<th>Expected Mortalities</th>
<th>Standardized Mortality Ratio</th>
<th>Windsor Standard Deviations = 151.33</th>
<th>Scatter Class Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>36</td>
<td>2</td>
<td>0.290</td>
<td>86.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>2</td>
<td>0.296</td>
<td>84.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>2</td>
<td>0.301</td>
<td>83.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>1</td>
<td>0.154</td>
<td>81.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>2</td>
<td>0.317</td>
<td>78.86</td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>2</td>
<td>0.335</td>
<td>74.63</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>0.185</td>
<td>67.57</td>
<td></td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>2</td>
<td>0.391</td>
<td>63.94</td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>1</td>
<td>0.217</td>
<td>57.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>2</td>
<td>0.451</td>
<td>55.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>1</td>
<td>0.296</td>
<td>42.22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>1</td>
<td>0.313</td>
<td>39.93</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>0.435</td>
<td>28.74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>1</td>
<td>0.454</td>
<td>27.53</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>1</td>
<td>0.462</td>
<td>27.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>1</td>
<td>0.479</td>
<td>26.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>0.522</td>
<td>23.94</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0.168</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0.133</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0.194</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>0.040</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>0.172</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>0</td>
<td>0.112</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Census of Canada, 1976, Series A and B
Windsor and Essex County Health Unit
Author
MORTALITIES OF ISCHAEMIC HEART DISEASE—FEMALE (1970-1977)

FIGURE 14
MORTALITIES OF LUNG CANCER, FEMALE (1970-1977)

FIGURE 15
Society in that female lung cancer is much higher overall than the male rates. Tracts reporting statistically significant higher than expected rates (Figure 16) corroborate the map patterns depicting tracts at or above +1 standard deviation with the exception of the exclusion of tracts 101 and 8.

There are no areas reporting statistically significant lower than expected standardized mortality rates at or below -1 standard deviation. The fact that significant S.M.R. values exist only at or above +1 standard deviation would then suggest that the values are randomly skewed towards the positive side. The values are random because of the distribution positively and negatively from the provincial norm, but only the positive values are significantly high.

In an attempt to determine possible causal factors related to female lung cancer, new lifestyle, ethnic and socio-economic variables were collected and mapped. The resultant patterns were then statistically tested to determine if the variables were significant as possible causal factors. The results of this analysis are found in the following section.

Lifestyle Characteristics of the Respondents in Aggregate: Female

This section will examine the lifestyle characteristics of the female respondents in aggregate, in order to identify patterns which may indicate a correlation of coincidence that could then be compared to the mortality information. As previously stated, while these maps may have similar patterns to that of the mortality data, they do not imply a relationship between the patterns of lifestyle and death. Summary tables of the data used to produce these maps may be found in the Appendix.
Figure 17, percent of female respondents with high blood pressure, coincides with tracts which have mortalities above the provincial norm for both disease types. The tracts with a high incidence of females with high blood pressure are 27, 37, 13, 17, 33, and 35. This pattern is very similar to that of the male rates of high blood pressure. Also, tracts reporting high ischaemic heart mortalities above the provincial norm are 32, 8, 28, 7, 33, and 35. The tracts coinciding with high blood pressure incidence are 33 and 35.

Tracts reporting lung cancer mortalities above the provincial norm are 8, 19, 12, 33, 34, and 25. The only tract coinciding with high blood pressure incidence is tract 33. Unlike the male case, tract 33 is the only one which coincides with both lung cancer and ischaemic heart mortalities.

The next map, percentage of heavy smoking females (Figure 18), depicts four tracts with a fairly high incidence of heavy smoking females; those being tracts 35, 41, 3, and 8.

This pattern by itself is somewhat surprising because recent studies, done by the Canadian Cancer Society and Health and Welfare Canada, have indicated a steady increase in the percentage of females who smoke. This pattern does, however, coincide with the pattern of high blood pressure incidence as well as the incidence of ischaemic heart disease and lung cancer mortalities. The tract coinciding with high blood pressure incidence is 35, and the tracts coinciding with heart mortalities are 3, 35, and 8. When the pattern of heavy smoking females is compared to that of the male rates, there appears to be no coincidence in the patterns. This combined with other factors may prove to be very important as causal fac-
tors for the incidence of disease mortality.

Figure 19, percentage of overweight females, depicts tracts with the high rates. These tracts are 45, 4, 5, 1, 101, 35, 38, 39, and 9. The tracts coinciding with high blood pressure incidence are 35 and 39. The only tract coinciding with the high rates of heavy smoking females is 35. The areas coinciding with ischaemic heart mortalities are 101, 11, 35, 38, and 39. Also, tract 38 coincides with the mortality rates of lung cancer above the provincial norm.

The next map, percentage of underweight females (Figure 20), reports high incidence in tracts 9, 30, 26, 10, 100, and 2. Areas coinciding with the rates of high blood pressure incidence are 100 and 2. There are no tracts coinciding with the high rates of heavy smoking females. Also, tract 100 is the only one coinciding with high rates of overweight females. The tracts coinciding with the high rates of ischaemic heart mortality are 9, 30, and 26. None of the areas reporting high incidence of cancer coincide with that of underweight females.

At this point a definite pattern becomes evident in that census areas located in the Southern and Western portions of the city consistently report high incidence rates of lifestyle patterns which may affect the mortality rates found in these areas. Also, tracts 33, 35, 8, and 3 specifically report higher than average rates of mortality and high rates of lifestyle characteristics which may be associated with the mortality patterns.

Figure 21, depicting the percentage of females consuming the birth control pill for more than four years, and who also smoke, coincides with tracts depicting high rates of disease mortality. This variable was
FIGURE 19

PERCENTAGE OF RESPONDENTS OVERWEIGHT-FEMALE.
PERCENTAGE OF RESPONDENTS UNDERWEIGHT-FEMALE

FIGURE 20
chosen because previous studies done by the American Cancer Society and reports from the Surgeon General of the United States stated that disease incidence, especially cancer, increased with the usage of the birth control pill. Also, when tobacco combined with the consumption of the birth control pill, rates of heart disease and cancer increased threefold, compared to women who do not smoke and do not consume the birth control pill. With this in mind, let us look at the patterns of coincidence.

Census tracts depicting a high percentage of women consuming the birth control pill are 38, 19, 27, 16, 23, 18, 26, 36, 31, 24, 21, 28, 100, 20, 41, 22, 9, 3, 17, 29, and 8. The tracts coinciding with the patterns of high blood pressure are 17, 27, 28, and 100. Tracts coinciding with the patterns of heavy smoking females are 8, 42, and 3. The areas depicting high rates of overweight females are 43, 4, 5, 1, 101, 11, 35, 38, 39, and 9. Census areas reporting high rates of consumption of the birth control pill, which coincide with overweight females, are 38, and 9. Areas which coincide with underweight females are 9, 26, and 100.

When comparing the ischaemic heart mortalities to the high rates of consumption of the birth control pill, tracts 38, 21, 23, 24, 28, 9, and 8 are similarly high. Also, tracts compared to the high rates of cancer mortality are 8, 3, 19, 42, 38, and 22.

The final map, Figure 22, depicts the prevalence of diagnosed heart disease as reported from the questionnaire data. Tracts which have high rates are 40, 39, 37, 33, 12, and 11. Tracts reporting high rates of high blood pressure which coincide with the prevalence of heart disease are 37, 33, and 12. Areas depicting high rates of heavy smokers do not coincide with the prevalence of diagnosed heart disease. The patterns of
overweight females associated with high patterns of diagnosed heart disease include 4, 11, and 39. Tracts reporting a high percentage of females who are underweight do not coincide with the prevalence of diagnosed heart disease. The only tract which reports both a high rate of consumption of the birth control pill, and prevalence of diagnosed heart disease is tract 23.

Census tracts 40, 39, 37, 33, and 11 correspond with the high rates of ischaemic heart disease mortalities. When the high rates of lung cancer mortality are compared to the pattern of diagnosed heart disease, only tracts 33 and 12 correspond with each other.

In summary, it should be stressed that although these variables may suggest some distinct patterns, they do not suggest that these variables will be statistically significant causal factors for the occurrence of ischaemic heart disease and lung cancer mortalities for females in Windsor. In order to determine statistically if these variables may be causal in nature, they were tested using factor and regression analyses. The results of these analyses are found in the following section.

Ischaemic Heart Factor Variables Findings: Female

Factor Analysis. This analysis, as in the male case, was used to locate a smaller number of underlying factors contained in the original set of variables collected. The factor analysis tested three assumptions: 1) That the 16 lifestyle variables affect the mortality variable and significant factors may be identified as such; 2) That the four ethnic variables affect the mortality variable, and 3) That the socio-economic variables may be identified as significant factors.

To test the first assumption, only the lifestyle variables were con-
sidered. The eigenvalues of the resultant factors express their relative importance to the mortality variable. Seven factors were established in the initial run, of which five were found to have significant eigenvalues (Table 16). The first factor identified was young smokers, where variable 51 (% of female respondents who smoke are less than 35 years of age), and variable 62 (% of females with bronchitis) were loaded significantly into this first factor. The second factor identified was female consuming the birth control pill, who smoke. Variable 3 (per capita alcohol consumption), variable 65 (% of women who have smoked for more than six, but less than 15 years) were loaded significantly into this second factor. The third factor identified heavy smokers where V53 (% of middle aged women who smoke) and V58 (% of females who are heavy smokers) were both loaded significantly into this third factor. The fourth factor identified was smokers with respiratory disease where V64 (% of females reporting various respiratory diseases) and V71 (% of females who have smoked for more than 25 years) were both loaded significantly into this fourth factor. The fifth factor identified diabetics where V51 (% of females on a diet) and V61 (% of females with diabetes) were both loaded significantly into this fifth factor. The final two factors, as previously mentioned, did not have significant eigenvalues. Therefore, the cutoff point was after factor five. Also, since the first five factors were identified as being significant, the first assumption was accepted.

The second assumption tested used only the ethnic variables. The eigenvalues provided four factors, two of which were significant (Table 17). The first factor identified was percent North European, where V48 (% North European), and V49 (% German) were loaded significantly into
TABLE 16

Eigenvalues for the Lifestyle Factors

<table>
<thead>
<tr>
<th>Factor</th>
<th>Eigenvalue</th>
<th>Pct of Var</th>
<th>Cum Pct</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.62423</td>
<td>25.3</td>
<td>25.3</td>
</tr>
<tr>
<td>2</td>
<td>1.93957</td>
<td>18.7</td>
<td>44.0</td>
</tr>
<tr>
<td>3</td>
<td>1.70252</td>
<td>16.4</td>
<td>60.5</td>
</tr>
<tr>
<td>4</td>
<td>1.30482</td>
<td>12.6</td>
<td>73.0</td>
</tr>
<tr>
<td>5</td>
<td>1.26027</td>
<td>12.2</td>
<td>85.2</td>
</tr>
<tr>
<td>6</td>
<td>0.88391</td>
<td>8.5</td>
<td>93.7</td>
</tr>
<tr>
<td>7</td>
<td>0.65013</td>
<td>6.3</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Source: SPSS Factor Analysis, 1978 Questionnaire Data

Quartimax Rotated Factor Matrix

<table>
<thead>
<tr>
<th>Variable</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
<th>Factor 4</th>
<th>Factor 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>On a diet</td>
<td>0.05377</td>
<td>-0.10219</td>
<td>0.02638</td>
<td>-0.06809</td>
<td>0.53379</td>
</tr>
<tr>
<td>Alcohol consumption</td>
<td>-0.24015</td>
<td>0.60249</td>
<td>0.06841</td>
<td>0.20665</td>
<td>-0.32938</td>
</tr>
<tr>
<td>Young smokers</td>
<td>0.93012</td>
<td>-0.11109</td>
<td>0.00029</td>
<td>-0.29147</td>
<td>0.04413</td>
</tr>
<tr>
<td>Middle age smokers</td>
<td>-0.35318</td>
<td>-0.04845</td>
<td>0.62434</td>
<td>-0.01855</td>
<td>-0.1258</td>
</tr>
<tr>
<td>Old smokers</td>
<td>-0.01017</td>
<td>-0.09670</td>
<td>0.27054</td>
<td>0.23405</td>
<td>-0.03675</td>
</tr>
<tr>
<td>Heavy smokers</td>
<td>0.05502</td>
<td>0.18426</td>
<td>0.83705</td>
<td>0.11509</td>
<td>-0.01310</td>
</tr>
<tr>
<td>Diagnosed heart disease</td>
<td>-0.23220</td>
<td>-0.12234</td>
<td>0.07186</td>
<td>-0.06345</td>
<td>0.16502</td>
</tr>
<tr>
<td>Diabetics</td>
<td>-0.04039</td>
<td>0.09280</td>
<td>-0.03380</td>
<td>0.09610</td>
<td>0.92137</td>
</tr>
<tr>
<td>Bronchitis</td>
<td>0.74339</td>
<td>0.02460</td>
<td>-0.21721</td>
<td>0.18361</td>
<td>0.01795</td>
</tr>
<tr>
<td>Pneumonia</td>
<td>0.19594</td>
<td>0.10625</td>
<td>0.22137</td>
<td>0.16173</td>
<td>-0.19259</td>
</tr>
<tr>
<td>Respiratory diseases</td>
<td>0.06243</td>
<td>0.25643</td>
<td>-0.10050</td>
<td>0.82146</td>
<td>-0.07725</td>
</tr>
<tr>
<td>Consumption of birth control pill</td>
<td>0.09259</td>
<td>0.77942</td>
<td>0.01540</td>
<td>0.03838</td>
<td>0.04441</td>
</tr>
<tr>
<td>6-15 years of smoking</td>
<td>-0.01611</td>
<td>0.69252</td>
<td>0.56409</td>
<td>-0.16537</td>
<td>0.09978</td>
</tr>
<tr>
<td>&gt; 25 years of smoking</td>
<td>-0.15622</td>
<td>-0.22524</td>
<td>0.18979</td>
<td>0.60503</td>
<td>0.08114</td>
</tr>
<tr>
<td>Overweight</td>
<td>0.26894</td>
<td>-0.35756</td>
<td>0.03875</td>
<td>0.10748</td>
<td>-0.03904</td>
</tr>
<tr>
<td>Underweight</td>
<td>-0.15103</td>
<td>0.04734</td>
<td>-0.27618</td>
<td>0.15179</td>
<td>0.01946</td>
</tr>
</tbody>
</table>

Source: SPSS Factor Analysis, 1978 Questionnaire Data
### TABLE 17

Eigenvalues for the Ethnic Variables

**Female Heart Disease**

<table>
<thead>
<tr>
<th>Factor</th>
<th>Eigenvalue</th>
<th>Pct of Var</th>
<th>Cum Pct</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.69197</td>
<td>33.8</td>
<td>33.8</td>
</tr>
<tr>
<td>2</td>
<td>1.08472</td>
<td>21.7</td>
<td>55.5</td>
</tr>
<tr>
<td>3</td>
<td>0.81445</td>
<td>26.3</td>
<td>81.8</td>
</tr>
<tr>
<td>4</td>
<td>0.77539</td>
<td>18.2</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Source: SPSS Factor Analysis, 1978 Questionnaire Data

---

**Quartimax Rotated Factor Matrix**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Factor 1</th>
<th>Factor 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>British and French</td>
<td>0.29439</td>
<td>0.36553</td>
</tr>
<tr>
<td>North European</td>
<td>0.45820</td>
<td>0.07354</td>
</tr>
<tr>
<td>German</td>
<td>0.33013</td>
<td>-0.26601</td>
</tr>
<tr>
<td>East European</td>
<td>0.37001</td>
<td>-0.27396</td>
</tr>
</tbody>
</table>

Source: SPSS Factor Analysis, 1978 Questionnaire Data
this factor. The second factor identified was percent British and French, where this variable was loaded significantly into this second factor. Since three of the four ethnic variables were loaded significantly, the second assumption was accepted.

The third assumption tested used only the socio-economic variables. The eigenvalues provided three factors, two of which were identified as being significant (Table 18). The first factor identified was blue collar workers where V30 (% managerial) loaded significantly, but was a negative loading. V34 (% factory labour) and V35 (% income less than $12,000) also loaded significantly as positive values. V37 (% income less than $30,000) also loaded significantly but the loading was negative. The second factor identified was lower to middle income as these two variables were loaded significantly into this second factor. The third assumption was also accepted because six of eight socio-economic variables loaded significantly into these two factors.

With these new factors in mind, regression analysis was used to test the null hypothesis which states that there is no significant correlation between lifestyle, ethnic and socio-economic variables and incidence of female ischaemic heart mortalities in Windsor.

Regression Analysis Female Heart Disease Findings

The results of the regression analysis are summarized in Table 19. The overall regression equation accounted for 44.28% of the variation explained for the incidence of heart disease mortalities in females. The first five variables entered into the equation accounted for 37.19% of the variation for the disease occurrence. These independent variables, in order of entry, are: X1 (% managerial), X8 (% North European), X6 (%
TABLE 18

Eigenvalues for the Socio-Economic Factors
Female Heart Disease

<table>
<thead>
<tr>
<th>Factor</th>
<th>Eigenvalue</th>
<th>Pct of Var</th>
<th>Cum Pct</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.62633</td>
<td>50.6</td>
<td>50.6</td>
</tr>
<tr>
<td>2</td>
<td>1.65882</td>
<td>32.0</td>
<td>82.6</td>
</tr>
<tr>
<td>3</td>
<td>0.90094</td>
<td>17.4</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Source: SPSS Factor Analysis, 1978 Questionnaire Data

Quartimax Rotated Factor Matrix

<table>
<thead>
<tr>
<th>Variable</th>
<th>Factor 1</th>
<th>Factor 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Managerial</td>
<td>-0.89050</td>
<td>-0.00175</td>
</tr>
<tr>
<td>Clerical</td>
<td>0.24095</td>
<td>0.15126</td>
</tr>
<tr>
<td>Farming</td>
<td>0.02132</td>
<td>0.28858</td>
</tr>
<tr>
<td>Mining</td>
<td>0.05127</td>
<td>-0.22470</td>
</tr>
<tr>
<td>Factory Labour</td>
<td>0.82029</td>
<td>-0.06418</td>
</tr>
<tr>
<td>Income &lt; $12,000</td>
<td>0.47682</td>
<td>-0.83168</td>
</tr>
<tr>
<td>Income $12,000-$30,000</td>
<td>0.00763</td>
<td>0.94389</td>
</tr>
<tr>
<td>Income &gt; $30,000</td>
<td>-0.62093</td>
<td>-0.04110</td>
</tr>
</tbody>
</table>

Source: SPSS Factor Analysis, 1978 Questionnaire Data
TABLE 19

Relationship between Female Ischaemic Heart Disease Mortalities: 1970-1977 and Selected Lifestyle, Ethnic and Socio-Economic Variables by Census Tract

<table>
<thead>
<tr>
<th>Variable</th>
<th>'r' (Simple Correlation Coefficient)</th>
<th>B (Partial Regression Coefficient)</th>
<th>F-Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1 Managerial</td>
<td>-0.4588</td>
<td>-1.1660</td>
<td>3.481*</td>
</tr>
<tr>
<td>X8 North European</td>
<td>0.3235</td>
<td>1.0785</td>
<td>1.094</td>
</tr>
<tr>
<td>X6 Income &gt; $30,000</td>
<td>-0.4240</td>
<td>-2.1751</td>
<td>4.759*</td>
</tr>
<tr>
<td>X7 British/French</td>
<td>-0.1670</td>
<td>-1.4090</td>
<td>2.817*</td>
</tr>
<tr>
<td>X11 Young smokers</td>
<td>-0.0960</td>
<td>-0.8444</td>
<td>2.473*</td>
</tr>
</tbody>
</table>

r = Simple correlation coefficient
B = Partial regression coefficient
F = Partial F-statistic for significance
R² = Coefficient of determination = 0.3719
F-critical value = 2.41
* = Significant at the 0.05 significance level

Regression equation: \( (Y_c) = 220.86 - 1.1660 \times (X_1) + 1.0785 \times (X_8) - 2.1751 \times (X_6) - 1.4090 \times (X_7) - 0.8444 \times (X_{11}) \)
of respondents earning more than $30,000 per annum), X7 (% British and French), and X11 (% of young smokers). As a total these variables were found to be significant at the 0.05 level of significance. However, only X1, X6, X7, and X11 had significant partial F-values. Although these variables had significant partial F-values, their associated B values were all negative, indicating an inverse relationship with the occurrence of ischaemic heart disease mortality in females for the city of Windsor. The low R-square value and the vagueness of the significant negative B values indicates that the independent variables used in this analysis negates inferences concerning causal relationships from being made. These factors, combined with new variables must be considered before any firm causal relationship can be made concerning the female ischaemic heart disease mortalities.

Lung Cancer Factor Variables Findings: Female

Factor Analysis. This analysis, as with the ischaemic heart, was used to locate a smaller number of underlying factors contained in the initial set of variables. This analysis tested three assumptions: 1) That the 15 lifestyle variables affect the mortality variable and significant factors may be identified as such; 2) That the four ethnic variables affect the mortality variables, and 3) That the socio-economic variables affect the mortality variable and may be identified as significant factors.

To test the first assumption, the lifestyle variables were considered. The eigenvalues of the resultant factors established seven new factors, of which five were significant (Table 20). The first factor identified was young smokers where variable 51 (% of young smokers) and variable 62 (% of respondents with bronchitis) were loaded significantly into this
TABLE 20

Eigenvalues for the Lifestyle Factors
Female Lung Cancer

<table>
<thead>
<tr>
<th>Factor</th>
<th>Eigenvalue</th>
<th>Pct of Var</th>
<th>Cum Pct</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.71223</td>
<td>27.4</td>
<td>27.4</td>
</tr>
<tr>
<td>2</td>
<td>1.96421</td>
<td>19.8</td>
<td>47.2</td>
</tr>
<tr>
<td>3</td>
<td>1.67847</td>
<td>16.9</td>
<td>64.1</td>
</tr>
<tr>
<td>4</td>
<td>1.27870</td>
<td>12.9</td>
<td>77.0</td>
</tr>
<tr>
<td>5</td>
<td>1.01692</td>
<td>10.3</td>
<td>87.3</td>
</tr>
<tr>
<td>6</td>
<td>0.63382</td>
<td>6.4</td>
<td>93.7</td>
</tr>
<tr>
<td>7</td>
<td>0.62421</td>
<td>6.3</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Source: SPSS Factor Analysis, 1978 Questionnaire Data

Quartimax Rotated Factor Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
<th>Factor 4</th>
<th>Factor 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diet</td>
<td>0.05777</td>
<td>0.05582</td>
<td>-0.09822</td>
<td>-0.10890</td>
<td>-0.02373</td>
</tr>
<tr>
<td>Alcohol Consumption</td>
<td>-0.29182</td>
<td>0.04453</td>
<td>0.57097</td>
<td>0.22401</td>
<td>-0.28544</td>
</tr>
<tr>
<td>Young smokers</td>
<td>0.94286</td>
<td>-0.01586</td>
<td>-0.07359</td>
<td>-0.26528</td>
<td>0.07704</td>
</tr>
<tr>
<td>Middle age smokers</td>
<td>-0.29669</td>
<td>0.68827</td>
<td>-0.04467</td>
<td>-0.04615</td>
<td>0.10493</td>
</tr>
<tr>
<td>Old age smokers</td>
<td>-0.00216</td>
<td>0.26847</td>
<td>-0.06585</td>
<td>0.25284</td>
<td>0.54330</td>
</tr>
<tr>
<td>Heavy smokers</td>
<td>0.03572</td>
<td>0.84629</td>
<td>0.20462</td>
<td>0.13925</td>
<td>-0.17364</td>
</tr>
<tr>
<td>Diagnosed cancer</td>
<td>-0.03614</td>
<td>-0.22705</td>
<td>-0.07864</td>
<td>0.16624</td>
<td>0.26301</td>
</tr>
<tr>
<td>Bronchitis</td>
<td>0.74667</td>
<td>-0.24359</td>
<td>0.02647</td>
<td>0.19827</td>
<td>-0.25907</td>
</tr>
<tr>
<td>Pneumonia</td>
<td>0.21011</td>
<td>0.19895</td>
<td>0.11352</td>
<td>0.19589</td>
<td>-0.12140</td>
</tr>
<tr>
<td>Respiratory disease</td>
<td>0.04245</td>
<td>-0.10819</td>
<td>0.25145</td>
<td>0.86982</td>
<td>0.14033</td>
</tr>
<tr>
<td>Consumption of birth</td>
<td>0.09748</td>
<td>0.00076</td>
<td>0.77006</td>
<td>0.05298</td>
<td>0.15126</td>
</tr>
<tr>
<td>control pill</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6-15 years of smoking</td>
<td>-0.00160</td>
<td>0.53663</td>
<td>0.71096</td>
<td>-0.16508</td>
<td>-0.04810</td>
</tr>
<tr>
<td>&gt; 25 years of smoking</td>
<td>-0.14549</td>
<td>0.20452</td>
<td>-0.22992</td>
<td>0.54903</td>
<td>0.20783</td>
</tr>
<tr>
<td>Overweight</td>
<td>0.25489</td>
<td>0.04321</td>
<td>-0.36582</td>
<td>0.11917</td>
<td>-0.07369</td>
</tr>
<tr>
<td>Underweight</td>
<td>-0.12330</td>
<td>-0.25927</td>
<td>0.06319</td>
<td>0.16115</td>
<td>0.66823</td>
</tr>
</tbody>
</table>

Source: SPSS Factor Analysis, 1978 Questionnaire Data
factor. The second factor identified heavy smokers. Variable 55 (% of middle aged smokers), V58 (% of heavy smokers), and V68 (% of respondents who have smoked for more than six but less than 15 years) were loaded significantly into this second factor. The third factor identified women on the pill who smoke and drink as these variables were all loaded significantly into this factor. The fourth factor identified women with respiratory problems as variables 64 (% of women with respiratory disease), and 71 (% of women who have smoked for more than 25 years) both loaded significantly into the fourth factor. The fifth factor identified overweight middle aged smoking females as these variables were loaded significantly into this fifth factor. The final two factors, as previously mentioned, did not have significant eigenvalues. Therefore, the cutoff point was after the fifth factor.

The second assumption used only the ethnic variables. Two factors were established of which only one had significant eigenvalues (Table 21). The factor identified was North European, in that this variable and percent German were loaded significantly into the factor. Since two of the four factor variables were loaded significantly, the second assumption was accepted.

The third assumption used only the socio-economic variables expected to be associated with female cancer mortalities. Four factors were established for this run, of which three had significant eigenvalues (Table 22). The first factor identified was clerical workers where V31 (% clerical) and V36 (income between $12,000 and $30,000) were identified as significant variables with high factor loadings. The second factor identified was females in the middle income bracket as this was the only significant factor loading for this second factor. The third factor
TABLE 21

Eigenvalues for the Ethnic Factors

Female Lung Cancer

<table>
<thead>
<tr>
<th>Factor</th>
<th>Eigenvalue</th>
<th>Pct of Var</th>
<th>Cum Pct</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.28776</td>
<td>69.3</td>
<td>69.3</td>
</tr>
<tr>
<td>2</td>
<td>0.57086</td>
<td>30.7</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Source: SPSS Factor Analysis, 1978 Questionnaire Data

Quartimax Rotated Factor Matrix

<table>
<thead>
<tr>
<th>Variable</th>
<th>Factor 1</th>
<th>Factor 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>British/French</td>
<td>0.00583</td>
<td>-0.34633</td>
</tr>
<tr>
<td>North European</td>
<td>-0.40934</td>
<td>0.72381</td>
</tr>
<tr>
<td>German</td>
<td>0.44330</td>
<td>-0.01827</td>
</tr>
<tr>
<td>East European</td>
<td>-0.05457</td>
<td>0.60762</td>
</tr>
</tbody>
</table>

Source: SPSS Factor Analysis, 1978 Questionnaire Data
TABLE 22

Eigenvalues for the Socio-Economic Variables
Female Lung Cancer

<table>
<thead>
<tr>
<th>Factor</th>
<th>Eigenvalue</th>
<th>Pct of Var</th>
<th>Cum Pct</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.64725</td>
<td>43.3</td>
<td>43.3</td>
</tr>
<tr>
<td>2</td>
<td>1.72547</td>
<td>28.2</td>
<td>71.5</td>
</tr>
<tr>
<td>3</td>
<td>1.04864</td>
<td>17.2</td>
<td>88.7</td>
</tr>
<tr>
<td>4</td>
<td>0.69055</td>
<td>11.3</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Source: SPSS Factor Analysis, 1978 Questionnaire Data

Quartimax Rotated Factor Matrix

<table>
<thead>
<tr>
<th>Variable</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Managerial</td>
<td>0.93884</td>
<td>0.02486</td>
<td>0.18028</td>
</tr>
<tr>
<td>Clerical</td>
<td>-0.00583</td>
<td>0.14654</td>
<td>-0.04654</td>
</tr>
<tr>
<td>Farming</td>
<td>-0.12754</td>
<td>0.54569</td>
<td>0.13181</td>
</tr>
<tr>
<td>Mining</td>
<td>-0.01196</td>
<td>-0.22972</td>
<td>0.03974</td>
</tr>
<tr>
<td>Factory Labour</td>
<td>-0.90790</td>
<td>-0.11487</td>
<td>-0.19601</td>
</tr>
<tr>
<td>Income &lt; $12,000</td>
<td>-0.29428</td>
<td>-0.84486</td>
<td>-0.34084</td>
</tr>
<tr>
<td>Income $12,000-$30,000</td>
<td>0.11507</td>
<td>0.95322</td>
<td>0.26353</td>
</tr>
<tr>
<td>Income &gt; $30,000</td>
<td>0.28991</td>
<td>-0.01754</td>
<td>0.89459</td>
</tr>
</tbody>
</table>

Source: SPSS Factor Analysis, 1978 Questionnaire Data
identified women in managerial positions as this variable and the upper income variable were loaded significantly into the third factor.

With these new factors in mind, and their associated significant factor variables, regression analysis was used to determine the best fitting model associated with the null hypothesis. The null hypothesis states that there is no significant correlation between the lifestyle, ethnic, and socio-economic variables and the dependent mortality variable in Windsor.

Since a significant number of factors were identified to fit the model, the three assumptions were accepted.

The results of the regression analysis are found in Table 23. The overall regression equation accounted for 57.63% of the variation explained for the incidence of female lung cancer mortalities in Windsor. Of the 17 variables entered, the first six accounted for 53.88% of the variation. These independent variables, in order of entry, are: X10 (% German), X17 (% females who are underweight), X15 (% of females on the pill for more than four years who smoke), X4 (% clerical), X13 (% of heavy smoking females), and X9 (% of North Europeans). As a total, these variables were found to be significant at the 0.05 level of significance and their associated partial F-values were also significant at the 0.05 level of significance. The first and third variables entered has strong positive B values, indicating a positive relationship between these independent variables and the dependent variable. Although the other variables were significant their associated B-values were negative, indicating an inverse relationship between them and the dependent variable. As with the heart results, these factors combined with new independent variables, must be considered before any firm causal relationship can be made con-
TABLE 23

Relationship between Female Lung Cancer Disease Mortalities: 1970-1977 and Selected Lifestyle, Ethnic and Socio-Economic Variables by Census Tract

<table>
<thead>
<tr>
<th>Variable</th>
<th>'r' (Simple Correlation Coefficient)</th>
<th>B (Partial Regression Coefficient)</th>
<th>F-Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>X10 German</td>
<td>0.4304</td>
<td>30.6286</td>
<td>16.329*</td>
</tr>
<tr>
<td>X17 Underweight</td>
<td>-0.3751</td>
<td>-13.2067</td>
<td>13.642*</td>
</tr>
<tr>
<td>X15 Birth control pill</td>
<td>-0.1270</td>
<td>4.1122</td>
<td>7.337*</td>
</tr>
<tr>
<td>X4 Clerical</td>
<td>-0.2035</td>
<td>-19.1064</td>
<td>4.062*</td>
</tr>
<tr>
<td>X13 Heavy smoker</td>
<td>0.0192</td>
<td>-3.7969</td>
<td>3.645*</td>
</tr>
<tr>
<td>X9 North European</td>
<td>-0.3650</td>
<td>-4.3875</td>
<td>3.570*</td>
</tr>
</tbody>
</table>

r = Simple correlation coefficient
B = Partial regression coefficient
F = Partial F-statistic for significance
R² = Coefficient of Determination = 0.5380
F-critical value = 2.42
* = Significant at the 0.05 significance level

Regression equation: \( Yc = 38.278 + 30.6286 \times X10 - 13.2067 \times X17 + 4.1122 \times X15 - 19.1064 \times X4 - 3.7969 \times X13 - 4.3875 \times X9 \)
cerning the female lung cancer mortalities.
CHAPTER V

CONCLUSIONS AND IMPLICATIONS FOR FUTURE RESEARCH

Medical geography is a field of study that encompasses two extensive yet related spheres of geographic inquiry. The first pertains to the analyzing and delineating of all types of medical facilities. The second subfield is concerned with the collection of health related data in order to accurately depict disease incidences and to determine the relevant spatial associations and environmentally linked health-risk factors.

This study aimed at depicting the second of these two investigative types in that it examined the spatial variations of heart disease and lung cancer mortality incidences in the city of Windsor. Data were obtained from the death certificate information made available by the Essex County Metropolitan Windsor Health Unit. These data, relevant for all individuals aged 15 through 64, were separated by sex, transformed to standardized mortality rates, mapped, and their patterns statistically tested for each of Windsor's 45 census tracts.

Also, additional lifestyle data were obtained by means of a questionnaire survey administered randomly throughout the city of Windsor, during the summer of 1978. The questionnaire accounted for a total sample of 1,845 households, or approximately 4% of the total households, for each
census tract in Windsor. These data were also dichotomized by sex, transformed into frequency tables, mapped, and their patterns statistically tested.

For each sex in both a city-wide and census tract context, higher than expected premature deaths as represented by S.M.R.'s were obtained. In addition, significant areal variations and clustering of map patterns for higher and lower than expected mortality incidence rates were observed. In both the male and the female example, elevated ischaemic heart mortalities were found to occur in census tracts near and/or along the Detroit river; tracts essentially located in the older down town core area and the down river westside portion of the city. Lower than expected ischaemic heart mortalities, again for both sexes, were noted for those tracts situated in the newer, suburban, southern and eastern sections of the city. Also, the areas depicting the high rates of lifestyle incidence, coincide with the areas depicting the high rates of mortality incidence for ischaemic heart disease. When the lifestyle variables were subjected to factor and multiple regression analyses, the variables found to be most significant as possible causal factors were, for the male case, young smokers, high stress occupational variables (managerial, clerical, and farming), and also the variable depicting smokers who have smoked for more than 25 years. The coefficient of determination, $R^2$, for male ischaemic heart disease was 0.2871. This would indicate that, although the above mentioned variables were found to be significant, the degree of association between them and heart mortality was weak. This may be due to the limited number of variables which were available for this study. The variables found to be significant for the female case were managerial,
percent North European, percent British/French, income greater than $30,000, and percentage of young smokers. The coefficient of determination for the female case was 0.3719, again indicating a stronger, but still weak relationship between the variables and the standardized mortality rates.

The relationship between male lung cancer mortalities and the selected lifestyle variables was found to be stronger than the heart relationship. The variables accounted for an overall coefficient of determination of 0.4010. The variables found to be significant were, again, young smokers, North Europeans, managerial, farming, and the ethnic variable, percent German. Of these the only variable which had a strong positive relationship was young smokers; all the rest indicated an inverse relationship. The relationship between the female lung cancer mortalities and the selected variables produced the highest degree of association between the significant variables and the standardized mortality rates of lung cancer in Windsor. The coefficient of determination accounted for 0.5386 of the incidence of lung cancer mortality in females in Windsor. The six variables which were found to be significant at the 0.05 level were: percent German, percent underweight, consumption of the birth control pill, percent clerical, heavy smokers, and North European. The two variables which had a strong positive relationship were percent German and consumption of the birth control pill. The underweight variable, although strong in confidence, may be an effect rather than a cause of lung cancer mortality in females in Windsor.

All of the above mentioned variables for both the male and female case for both disease types, were found to have high rates of incidence in the
same census tracts. Although these variables correlated significantly with the standardized mortality rates for both disease types, they may not themselves be directly involved as causal factors. New and more specific data must be generated in order to define direct causal relationships. The most advantageous type of data would be complete medical histories of individuals, occupational history, dietary patterns, and mobility data. There must be a greater availability of this type of data pertaining to small areal units. This would then allow for more conclusive cause-related inferences to be drawn. Similar temporal and methodologically oriented studies must be made for cities and regions in Canada and the United States to determine whether or not Windsor is a unique case in the context of lifestyle and disease patterns. Finally, the construction of a continuously evolving computer data bank containing mortality and morbidity data and a wide range of socio-economic and environmental variables be made freely accessible to all concerned individuals.
Dear Sir/Madam

The following questionnaire is part of a Canada Works Project sponsored by the Department of Employment and Immigration. It is supported by the Metropolitan Windsor Health Unit, and the University of Windsor Departments of Home Economics and Geography. This survey aims to obtain specific health, recreational, nutritional and shopping aspects of Windsor's population in order to aid in further planning activities. The survey has been prepared under the supervision of Dr. P. C. Innes, Chairman of the Geography Department, University of Windsor.

During the next four months, over 1,000 Windsor households will be interviewed. No names are required for the questionnaire. All information will be transformed into statistics and individual privacy, will be complete. Street names are necessary only for mapping purposes.

A final report will be submitted to the government in September, 1978 and copies of the project will be made available to the public at various municipal libraries and at the University of Windsor.

Your cooperation in answering the questions as fully and accurately as possible will be greatly appreciated. If there are any problems or questions pertaining to this questionnaire, please contact the Geography Department, University of Windsor, 253-4222, ext. 446.

Yours sincerely,

F. C. Innes, Ph.D.
Departmental Chairman

Eric Christou
Project Manager

If we fail to pick up your questionnaire within 1 week, please mail it to: Eric Christou
Geography Department
University of Windsor
1. Age ______ (years)
2. Sex: Male Female
3. Weight ______ (pounds)
4. Height ______ (feet)
5. What is your marital status?
   1. Married
   2. Widowed
   3. Divorced
   4. Separated
   5. Single
   6. Common-law
6. What is the present size of your family living at home?
   1. One
   2. Two
   3. Three
   4. Four
   5. Five
   6. Six
   7. Other (specify)
7. What ethnic group do you feel your family belongs to?
   1. Canadian or American
   2. British (British Isles)
   3. Franch
   4. Italian
   5. German
   6. Polish, Ukrainian & other Eastern European
   7. Southern European (Greek, Spanish, Yugoslav, etc.)
   8. Northern and Western Eur.
   9. Other (specify)
8. What religion, if any, do you practice?
   1. Roman Catholic
   2. Protestant
   3. Orthodox
   4. Jewish
   5. Muslim
   6. Hindu
   7. Buddhist
   8. Other (specify)
   9. None
9. What language is most commonly spoken in your home?
10. How long have you resided at this address?
    1. less than 1 year
    2. 1 - 5 years
    3. 6 - 10 years
    4. 11-15 years
    5. 16-20 years
    6. over 20 years
11. What is your previous residential history? Please complete the following table indicating the most recent address first.
    | Address (Street) | Years Residing at That Address |
    |------------------|--------------------------------|
    |                  | 19 - 19                       |
    |                  | 18 - 19                       |
    |                  | 19 - 19                       |
    |                  | 19 - 19                       |
    |                  | 19 - 19                       |
12. Are you planning to relocate in the near future?
   1. Yes
   2. No

13. What is your approximate yearly family income?
   1. less than $6,000
   2. $6,001-$12,000
   3. $12,001-$18,000
   4. $18,001-$30,000
   5. $30,001-$42,000
   6. $42,001-$54,000
   7. $54,000 and over

14. The following table investigates the occupations, job duties, and possible workplace health hazards encountered by all employed members of your household. The topic of health/physical precautions refers to such aspects as the wearing of e.g. ear protectors, eye guards, filter masks, special protective clothing, etc. Please complete this table to the best of your ability.
13. Please complete the blank spaces in the following table.

<table>
<thead>
<tr>
<th>Town or City</th>
<th>Province</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>Place of birth of household head</td>
<td>/</td>
<td></td>
</tr>
<tr>
<td>Place of birth of spouse</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Place of birth of parents on male side: Father</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Place of birth of parents on female side: Father</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

14. Has anyone in the household ever used birth control pills?
   1. Yes
   2. No

17. If 'yes' (14) for how long?
   1. less than 1 year
   4. 1 - 11 years
   2. 1 - 3 years
   5. over 11 years
   3. 4 - 7 years

18. Do you, or anyone in the household, suffer from high blood pressure?
   1. Yes
   2. No

19. If 'yes' (18) please indicate family member
   Number of years with high blood pressure

20. Does anyone in the household smoke?
   1. Yes
   2. No

22. If yes (20) please fill in the following table.
   (Note: for amounts smoked, 1 = very infrequently, or light smoking e.g. less than 10 cigarettes/day, 2 = medium smoking, e.g. 4 to 1 pack/day, 3 = about 1 pack/day and over)

<table>
<thead>
<tr>
<th>Household member</th>
<th>Age</th>
<th>Years smoking</th>
<th>Cigarette usage</th>
<th>Cigar usage</th>
<th>Pipe usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
22. How would you rate the health of your household, generally?
1. Poor
2. Fair
3. Good
4. Excellent
5. Don't know

23. Are there any chronic illnesses, disabilities, affecting any household members? (Such illnesses include: kidney disease, heart disease, cancer, respiratory ailments, diabetes, arthritis, high cholesterol level)
1. Yes
2. No

24. If 'yes' (§ 23) please fill in the following table.

<table>
<thead>
<tr>
<th>Household member</th>
<th>Age</th>
<th>Illness</th>
<th>Duration with Illness</th>
<th>Occupation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

25. Are there any family members, or relatives (not residing in your household) suffering from, or who have died from, chronic illnesses?
1. Yes
2. No

If 'yes' (§ 25) please fill in the following table.

<table>
<thead>
<tr>
<th>Relation</th>
<th>Age</th>
<th>Illness</th>
<th>Duration with Illness</th>
<th>Occupation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

26. Has any member of your household ever had any of the following illnesses? (If 'yes', please indicate with an 'x' in appropriate box)

<table>
<thead>
<tr>
<th>Illness</th>
<th>Household Member (specify member and age)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Asthma</td>
<td></td>
</tr>
<tr>
<td>2. Pneumonia</td>
<td></td>
</tr>
<tr>
<td>3. Arthritis</td>
<td></td>
</tr>
<tr>
<td>4. Allergy</td>
<td></td>
</tr>
<tr>
<td>5. Arthritis</td>
<td></td>
</tr>
<tr>
<td>6. Any disease(s) of the heart as diagnosed by a physician</td>
<td></td>
</tr>
</tbody>
</table>
37. Does your household include any pets?
   1. Yes  2. No

38. If 'yes' (#37) please indicate type of pet.
   1. Dog  4. Turtle
   2. Cat  5. Hamster
   3. Bird (type)  6. Other (specify)

39. Has anyone in your household been bitten by an animal during the last 2 years?
   1. Yes  2. No

40. If 'yes' (#39) please indicate type of animal.
   1. Dog  4. Squirrel
   2. Cat  5. Other (specify)
   3. Rodent (e.g., rats, mice)  6. Don't Know

41. Has anyone in your household seen any rodents (e.g., rats, mice) in your neighbourhood during the past 2 years?
   1. Yes  2. No

42. Do any household members drink alcoholic beverages?
   1. Yes  2. No
   If 'no', skip to question #34.

43. If the heads of household drink alcoholic beverages, where is most of the drinking done?
   1. In the house  4. At friends' houses
   2. In Taverns  5. At parties
   3. In restaurants  6. Other (specify)

44. Are alcoholic beverages consumed during the meal?
   1. Commonly  3. Never
   2. Occasionally  4. Prohibited

45. How often, on the average, are the following beverages consumed?
   (Please fill in the table only where most appropriate)

<table>
<thead>
<tr>
<th>Number of bottles, glasses Per:</th>
<th>Day</th>
<th>Week</th>
<th>Month</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wine</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spirits (e.g., rye, whiskey, gin, etc.)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

46. Do any members of your household take prescription drugs?
   1. Yes  2. No
37. If 'yes' (¶36) by whom and for what reason?

________________________________________________________________________

38. Do any members of your household use marijuana?
1. __ Yes
2. __ No

39. If 'yes' (¶38) how frequently?

At least once a: 

(check appropriate box)

S
E
W
K
M
O
A
P
R

40. Nutritional Practices

The following questions pertain to your personal eating habits. Place a check beside the appropriate response(s) or fill in the answer when necessary.

41. Does anyone in your household use any vitamins or tonics?
1. __ No
2. __ Yes

42. If 'yes' (¶40) are they?
1. __ Single vitamin or mineral (e.g. vitamins C, E, iron, calcium, etc.)
2. __ Multivitamins
3. __ Multivitamins and minerals
4. __ Food supplements (e.g. wheat germ, rose hip syrup, kelp, etc.)

43. Do you consume soft drinks?
1. __ Yes
2. __ No

44. If 'yes' (¶42) are they?
1. __ Low calorie soft drinks
2. __ Sugared soft drinks

45. If 'yes' (¶44) how many bottles or cans (1 drink = 12 oz.) are consumed in an average day?
1. __ One
2. __ Two
3. __ Three
4. __ Four
5. __ Five
6. __ Six
7. __ More than six (specify)_______

46. Do you drink coffee or tea?
1. __ Yes
2. __ No

47. If 'yes' (¶46) how many cups per day?
1. __ One
2. __ Two
3. __ Three
4. __ Four
5. __ Five
6. __ Six
7. __ More than six (specify)_______
47. If 'yes' (44) is milk or sugar added?

<table>
<thead>
<tr>
<th>Milk</th>
<th>Sugar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

48. Do you normally have?

1. Breakfast
2. Midmorning snack
3. Lunch
4. Afternoon snack
5. Dinner
6. Late night snack
7. Other (specify) __________

49. What do your snacks normally consist of?

1. Not applicable
2. Fruit
3. Sandwich
4. Crackers/Cheese
5. Chips or others
6. Cookies
7. Donuts

50. Do you, or any household member, follow any type of special diet?

1. Yes
2. No

50b. Is this diet being followed on the advice of a doctor?

1. Yes
2. No

51. If 'no' (50b) does your doctor know you are following this diet?

1. Yes
2. No

52. Is the diet as:

1. Low calorie diet
2. Diabetic diet
3. High protein diet
4. Low fat diet
5. Low carbohydrate diet
6. Vegetarian diet
7. Low cholesterol diet
8. Low salt diet
9. Special diet (name) __________

53. If you are trying to reduce your weight, is your diet successful?

1. Yes
2. No

54. How long has this diet been followed?

1. One month
2. 1-3 months
3. 1-2 years
4. 6 mos. - 1 year
5. More than 3 years

55. Does your household regularly consume a quantity of traditional ethnic type foods?

1. Yes
2. No

56. If 'yes' (45) what type? (e.g. Italian, Chinese, Greek, German, etc.)
57. What type of bread is most regularly consumed by your household?
   1. ___ Commercial white bread  
   2. ___ Home made white bread  
   3. ___ Rye bread  
   4. ___ Whole wheat bread  
   5. ___ Soya bread  
   6. ___ Health food store bread  
   7. ___ Other (specify)  

58. What type of cereal is most usually consumed by your household?
   1. ___ Commercial granola  
   2. ___ Home made granola  
   3. ___ Commercial breakfast cereals  
   4. ___ Oatmeal  
   5. ___ Other (specify)  

59. Are rice and potatoes a regular part of your household’s diet?
   1. ___ Yes  
   2. ___ No  

60. Are pasta and macaroni products a regular part of your household’s diet?
   1. ___ Yes  
   2. ___ No  

61. How many meat servings (3-4 oz. = 1 serving) do you have in an average day?
   1. ___ One  
   2. ___ Two  
   3. ___ Three  
   4. ___ Other (specify)  

62. How many fish servings (3-4 oz. = 1 serving) do you have in an average day?
   1. ___ One  
   2. ___ Two  
   3. ___ Three  
   4. ___ Four  
   5. ___ Other (specify)  

63. How many fruit servings (1/2 cup = 1 serving) do you have daily?
   1. ___ One  
   2. ___ Two  
   3. ___ Three  
   4. ___ Other (specify)  

64. What type of fruit do you eat most often?
   1. ___ Citrus (e.g. oranges, lemons, grapefruit)  
   2. ___ Non citrus (e.g. pears, apples, peaches)  
   3. ___ Both  

65. How many vegetable servings (1/2 cup = 1 serving) do you have daily?
   1. ___ One  
   2. ___ Two  
   3. ___ Three  
   4. ___ Other (specify)  

66. What kind of vegetables do you eat most often?
   1. ___ Yellow & red  
   2. ___ Green leafy  
   3. ___ Green non-leafy  
   4. ___ None  
   5. ___ Other (specify)
67. How many glasses of milk (8 oz. = 1 glass) do you have daily?
1. One
2. Two
3. Three
4. Other (specify)

68. What type of milk do you drink most frequently?
1. Homo milk
2. 2% milk
3. Skim milk
4. Buttermilk
5. Powdered milk
6. Other (specify)

69. Do you consume the following dairy products?
1. Ice cream
2. Yogurt
3. Cheese
4. Sour cream
5. Cottage Cheese

---

70. What kinds of oil or shortening do you use in food preparation?
1. Olive Oil
2. Peanut oil
3. Corn oil, vegetable oil
4. Margarine
5. Rendered animal fat
6. Butter
7. Vegetable shortening
8. Rape seed oil
9. Soya bean oil
10. Other (specify)

71. What type of heating equipment is used in food preparation in your home?
1. Electric stove
2. Gas stove
3. Oven
4. Toaster
5. Microwave oven
6. Barbeque grill
7. Other (specify)

72. Who, for the most part, prepares the meals in your household?
1. Household head
2. Spouse
3. Other (specify)

73. How often (average per week) do you eat outside your home?
1. Once
2. Twice
3. Three times
4. Four times
5. Five times
6. More than 5 (specify)

74. Where do you usually eat (outside the home)?
1. Restaurant
2. Fast food store
3. Work cafeteria
4. Packed lunch
75. Do you shop at:
   1. __ Downtown
   2. __ Shopping malls
   3. __ Other (specify, e.g. Ottawa Street)

76. If you shop at malls, which malls do you most usually frequent?
   1. __ Devonshire Mall
   2. __ Tecumseh Mall
   3. __ University Mall
   4. __ Other (specify)

77. Do you often use the Metropolitan Detroit area for shopping?
   1. __ Yes
   2. __ No

78. If 'yes' (77) please specify why.

79. Do you think Windsor’s downtown parking facilities are adequate?
   1. __ Yes
   2. __ No

80. Do you use the downtown parking garage?
    1. __ Yes
    2. __ No

81. Do you use Ouellette Avenue for parking?
    1. __ Yes
    2. __ No

82. Would you like to see Ouellette Avenue converted into a downtown shopping mall (with auto traffic banned)?
    1. __ Yes
    2. __ No

83. Would you prefer to see continued expansion of parkland development along the riverfront?
    1. __ Yes
    2. __ No

84. Where do you usually obtain the majority of your food and drink for home preparation?
   1. __ Supermarkets
   2. __ Various small stores
   3. __ Friends and neighbours
   4. __ Other (specify)

85. In an average week, at how many different stores do you shop for food products?
   1. __ One
   2. __ Two
   3. __ Three
   4. __ Four
   5. __ Other (specify)

86. How often do you frequent the following types of stores (check where app.)
At least once: Daily; Weekly; Bi-Monthly; Monthly; Rarely; Never?

<table>
<thead>
<tr>
<th>Store Type</th>
<th>Daily</th>
<th>Weekly</th>
<th>Bi-Monthly</th>
<th>Monthly</th>
<th>Rarely</th>
<th>Never</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small Grocery Store</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delicatessen</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bakery</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fruit Market</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Meat Market</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>City Market</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vegetable Market</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supermarket</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health Food Store</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fish/Seafood Shop</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other (specify)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
87. How many automobiles does your household have?
   1.____ One
   2.____ Two
   3.____ Three
   4.____ Four
   ___ Other (specify)____

88. Do you own or rent a vacation home?
   1.____ Yes
   2.____ No

89. If 'yes' (# 88), what type?
   1.____ Cottage
   2.____ Summer House
   3.____ Motels
   4.____ Other (specify)____

90. Have you ever travelled abroad (outside North America)?
   1.____ Yes
   2.____ No

91. If 'yes' (# 90), where and when (most recent first)?
   Place
   Date
   1.____
   2.____
   3.____
   4.____

92. Have you travelled in Canada (outside Ontario)?
   1.____ Yes
   2.____ No

93. If 'yes' (# 92), where and when (most recent first)?
   Place
   Date
   1.____
   2.____
   3.____
   4.____

94. What is the mode of transportation you most frequently use to
   travel when on vacation?
   1.____ Car
   2.____ Airplane
   3.____ Train
   4.____ Bus
   5.____ Camper
   6.____ Boat
   7.____ Other (specify)____

95. Do you have school-age children?
   1.____ Yes
   2.____ No

96. If 'yes' (# 95), please fill in the table with checks where appropriate.

<table>
<thead>
<tr>
<th>Age Limit</th>
<th>Male</th>
<th>Female</th>
<th>Elementary School</th>
<th>Secondary School</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10-15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16-18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19-21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

97. Where, most frequently, do your children play?
   1.____ Backyard
   2.____ Neighbourhood
   3.____ Street
   4.____ Schoolyard
   5.____ Park(s)
   6.____ Other (specify)____
98. How close are you to the nearest park?

1. __ 1 mile  
2. __ 2 mile  
3. __ 2 miles 
4. __ 3 miles 
5. __ 4 miles 
6. __ more than 4 miles

99. Which park(s) do you and your children most frequently visit?

<table>
<thead>
<tr>
<th>Park Name</th>
<th>Average Time Spent for Visit</th>
<th>Weekly Frequency</th>
<th>Distance Traveled</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

100. Would you or your family use the neighbourhood park(s) more frequently if they were closer?

1. __ Yes  
2. __ No

101. Are the facilities at your neighbourhood park(s) adequate?

1. __ Yes  
2. __ No

102. If 'no' (§ 103 what improvement(s) would you like to have added to your neighbourhood park(s)?

_________________________________________________________________________

103. What facilities do you feel are most attractive at your neighbourhood park(s)? (Number your preferences in order of importance)

1 = very attractive
2 = fair
3 = poor

1. __ Greenspace
2. __ Play area
3. __ Swimming
4. __ Skating
5. __ View/Attractiveness
6. __ Tot-lot/Children's Playground
7. __ Sport Facilities (e.g. baseball, football,)
8. __ Other (specify) ______________________

104. In general, do you feel that there are enough parks to serve your needs?

1. __ Yes  
2. __ No

105. Would you be willing to pay higher taxes for the improvement of your neighbourhood park(s)?

1. __ Yes  
2. __ No

THANK YOU FOR YOUR COOPERATION AND TIME.
APPENDIX B

NUMBER OF HOUSEHOLDS SURVEYED BY CENSUS TRACT
<table>
<thead>
<tr>
<th>Census Tract #</th>
<th>Expected # of Questionnaire</th>
<th>Actual # Received</th>
<th>% of Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>56</td>
<td>45</td>
<td>80</td>
</tr>
<tr>
<td>2</td>
<td>42</td>
<td>32</td>
<td>76</td>
</tr>
<tr>
<td>3</td>
<td>60</td>
<td>50</td>
<td>83</td>
</tr>
<tr>
<td>4</td>
<td>88</td>
<td>57</td>
<td>65</td>
</tr>
<tr>
<td>5</td>
<td>67</td>
<td>38</td>
<td>57</td>
</tr>
<tr>
<td>6</td>
<td>59</td>
<td>40</td>
<td>68</td>
</tr>
<tr>
<td>7</td>
<td>12</td>
<td>11</td>
<td>92</td>
</tr>
<tr>
<td>8</td>
<td>5</td>
<td>5</td>
<td>100</td>
</tr>
<tr>
<td>9</td>
<td>107</td>
<td>53</td>
<td>50</td>
</tr>
<tr>
<td>10</td>
<td>117</td>
<td>20</td>
<td>17</td>
</tr>
<tr>
<td>11</td>
<td>50</td>
<td>34</td>
<td>68</td>
</tr>
<tr>
<td>12</td>
<td>9</td>
<td>9</td>
<td>100</td>
</tr>
<tr>
<td>13</td>
<td>70</td>
<td>61</td>
<td>87</td>
</tr>
<tr>
<td>14</td>
<td>63</td>
<td>52</td>
<td>83</td>
</tr>
<tr>
<td>15</td>
<td>53</td>
<td>45</td>
<td>85</td>
</tr>
<tr>
<td>16</td>
<td>74</td>
<td>58</td>
<td>78</td>
</tr>
<tr>
<td>17</td>
<td>72</td>
<td>60</td>
<td>83</td>
</tr>
<tr>
<td>18</td>
<td>100</td>
<td>83</td>
<td>83</td>
</tr>
<tr>
<td>19</td>
<td>23</td>
<td>20</td>
<td>87</td>
</tr>
<tr>
<td>20</td>
<td>72</td>
<td>56</td>
<td>78</td>
</tr>
<tr>
<td>21</td>
<td>97</td>
<td>83</td>
<td>91</td>
</tr>
<tr>
<td>22</td>
<td>61</td>
<td>57</td>
<td>93</td>
</tr>
<tr>
<td>23</td>
<td>53</td>
<td>45</td>
<td>85</td>
</tr>
<tr>
<td>24</td>
<td>85</td>
<td>78</td>
<td>92</td>
</tr>
<tr>
<td>25</td>
<td>53</td>
<td>44</td>
<td>91</td>
</tr>
<tr>
<td>26</td>
<td>91</td>
<td>78</td>
<td>87</td>
</tr>
<tr>
<td>27</td>
<td>27</td>
<td>24</td>
<td>89</td>
</tr>
<tr>
<td>28</td>
<td>83</td>
<td>74</td>
<td>89</td>
</tr>
<tr>
<td>29</td>
<td>71</td>
<td>55</td>
<td>77</td>
</tr>
<tr>
<td>30</td>
<td>78</td>
<td>53</td>
<td>68</td>
</tr>
<tr>
<td>31</td>
<td>40</td>
<td>35</td>
<td>88</td>
</tr>
<tr>
<td>32</td>
<td>62</td>
<td>35</td>
<td>55</td>
</tr>
<tr>
<td>33</td>
<td>50</td>
<td>16</td>
<td>32</td>
</tr>
<tr>
<td>34</td>
<td>47</td>
<td>27</td>
<td>57</td>
</tr>
<tr>
<td>35</td>
<td>88</td>
<td>85</td>
<td>97</td>
</tr>
<tr>
<td>36</td>
<td>62</td>
<td>54</td>
<td>87</td>
</tr>
<tr>
<td>37</td>
<td>29</td>
<td>16</td>
<td>55</td>
</tr>
<tr>
<td>38</td>
<td>39</td>
<td>30</td>
<td>77</td>
</tr>
<tr>
<td>39</td>
<td>35</td>
<td>16</td>
<td>46</td>
</tr>
<tr>
<td>40</td>
<td>45</td>
<td>16</td>
<td>36</td>
</tr>
<tr>
<td>41</td>
<td>42</td>
<td>14</td>
<td>33</td>
</tr>
<tr>
<td>42</td>
<td>37</td>
<td>27</td>
<td>73</td>
</tr>
<tr>
<td>43</td>
<td>60</td>
<td>55</td>
<td>92</td>
</tr>
<tr>
<td>100</td>
<td>61</td>
<td>36</td>
<td>59</td>
</tr>
<tr>
<td>101</td>
<td>42</td>
<td>33</td>
<td>79</td>
</tr>
</tbody>
</table>

Totals       | 2887                         | 1923            | 66.61%        |
Mean         | 64.15                        | 42.73           |               |
REFERENCES


VITA AUCTORIS

Name: Michael Arthur Desmarais

Birthdate: October 19, 1954

Birthplace: Windsor, Ontario, Canada

Education: B.A., Geography, University of Windsor, 1979

M.A., Geography, University of Windsor, 1981