STUDY OF PATTERNS OF MEDICAL CARE UTILIZATION USING COMPUTER ALGORITHMS

NAN LI

0-612-54931-3

Memorial
University of Newfoundland

This is to authorize the Dean of Graduate Studies to deposit two copies of my thesis/report entitled

Study of Patterns of Medical Care Utilization Using Computer Algorithms

in the University Library, on the following conditions. I understand that I may choose only ONE of the Options here listed, and may not afterwards apply for any additional restriction. I further understand that the University will not grant any restriction on the publication of thesis/report abstracts.

(After reading the explanatory notes at the foot of this form, delete TWO of (a), (b) and (c), whichever are inapplicable.)

The conditions of deposit are:

(a) that two copies are to be made available to users at the discretion of their custodians,

OR

(b) ~~that access to, and quotation from, this thesis/report is to be granted only with my written permission for a period of one year from the date on which the thesis/report, after the approval of the award of a degree, is entrusted to the care of the University, namely, ________ 19____, after which time the two copies are to be made available to users at the discretion of their custodians,~~

OR

(c) ~~that access to, and quotation from, this thesis/report is to be granted only with my written permission for a period of ________ years from the date on which the thesis/report, after approval for the award of a degree, is entrusted to the care of the University; namely, ________, 19____; after which time two copies are to be made available to users at the discretion of their custodians.~~

Date Jan. 12, 2000

Signed [signature]

[signature]
Dean of Graduate Studies

Witnessed by [signature]

NOTES

1. *Restriction (b) will be granted on application, without reason given.*

   However, applications for restriction (c) must be accompanied with a detailed explanation, indicating why the restriction is thought to be necessary, and justifying the length of time requested. Restrictions required on the grounds that the thesis is being prepared for publication, or that patents are awaited, will not be permitted to exceed three years.

   Restriction (c) can be permitted only by a Committee entrusted by the University with the task of examining such applications, and will be granted only in exceptional circumstances.

2. *Thesis writers are reminded that, if they have been engaged in contractual research, they may have already agreed to restrict access to their thesis until the terms of the contract have been fulfilled.*

STUDY OF PATTERNS OF MEDICAL CARE UTILIZATION USING COMPUTER ALGORITHMS

By

Nan Li

A thesis submitted to the school of Graduate Studies
in partial fulfillment of the requirements for the degree of
Master of Science (Medicine)

Division of Community Health
Faculty of Medicine
Memorial University of Newfoundland
January, 2000

St. John's Newfoundland

Abstract

The main purpose of this study is to investigate and characterize patterns of medical care utilization, specifically, to discover if there are associations between patterns of medical care utilization and demographic variables, socioeconomic status, and area of residence for three diagnostic groups (cardiovascular disease, mental disease and chronic respiratory conditions) across three geographical categories which divides the island into three areas by degree of urbanization. This study focuses on ambulatory physician visits which included outpatient, emergency room, and office visits, and excludes physician visits to hospitalized patients and visits to nursing home residents. It uses secondary data generated by the Newfoundland Panel on Health and Medical Care (NPHMC). The NPHMC began with a cross-sectional telephone survey in 1994-95 (random single-stage cluster sample of households selected by RDD); respondents who gave written consent to access medical care databases became the utilization panel. This panel was then linked to utilization databases – hospital separations and physicians' claims – using the provincial health insurance number, for a seven-year period (April 1, 1992 – March 31, 1999). The analysis in this thesis includes 678 subjects for cardiovascular disease, 402 subjects for mental disease, and 942 subjects for chronic respiratory conditions, 20 years or older, residing in the province of Newfoundland, Canada, and followed for four years (April 1, 1992 – March 31, 1996).

Patterns of utilization in the three diagnostic groups were investigated by dividing the four-year study period into 16 trimesters and using SAS programs to identify five main patterns of care: non-episodic, isolated episodes, and continuous episodes over 2-5, 6-11, or 12-16 consecutive trimesters. Respondents were assigned to one of three classifications by urbanization level: metropolitan St. John's, other urban, and remote, which acted as a proxy for level of medical care resources.

The hypothesis that demographic and socioeconomic variables, and geographical categories were associated with patterns of medical care utilization was tested separately for hospitalizations and physicians' visits. Residents in metropolitan St. John's are more likely to have continuous episodes than those in other urban or remote areas no matter what diagnosis they have. Residents in remote areas are more likely to be hospitalized than those in urban areas. The respondents who are older and have lower socioeconomic scores are more likely to have continuous episodes of GP visits for the three diagnoses. Generally, respondents with any of the three diagnoses are more likely to be older, have poorer health and lower socioeconomic scores, and live in urban areas as compared to those without the three diagnoses.

Descriptive and multivariate analyses clarify the complex association with demographic and socioeconomic variables, as well as with the urban variable. It appears from the analysis that even in a society with universal medical coverage there is an increase in continuous episodes for specialist visits in the urban areas of the three diagnostic groups, and an increase in the continuous episodes for GP visits in the urban areas for all subjects. The subjects in the remote areas have a higher probability of being hospitalized for the three diagnostic groups. As expected, access has significant influence on patterns of specialist visits. This study provides some evidence to increase knowledge

of the factors which determine patterns of medical care utilization for the three diagnostic groups.

Acknowledgments

I owe great thanks to my thesis supervisor, Dr. Jorge Segovia, who introduced me to the field of health services research, and was supportive and genuinely interested in my academic and intellectual development. I am indebted to my supervisor for his supervision, help and encouragement through the entire course of this study. A sincere word of appreciation is given to the members of my advisory committee, Dr. Doreen Neville and Dr. Veerabhadra Gadag, for their unfailing support and guidance during the entire thesis process.

I owe my gratitude to Ms. Alison C. Edwards. It is a pleasure to take this opportunity to express my sincere appreciation of all kinds of generous help from her. Without her assistance, the results presented here would never have been the same.

I would also like to extend my thanks to Dr. Roy F. Bartlett for his expertise, invaluable comments and enthusiastic help throughout this program.

I am grateful to all friends and members of Division of Community Health for their support along the way. Moral and financial support from both the Faculty of Medicine and the School of Graduate Studies throughout this program is greatly appreciated. I also wish to extend my thanks to Ms. Pauline Cole for her cordial administrative help and support throughout my graduate program.

Finally, my love and appreciation go to my husband, Zhihu Ding, and my parents. This thesis would not have been possible without their love, support and patience.

I warmly thank each of you.

Table of contents

Page

List of tables

List of figures

Chapter 1 Introduction

Are rural people disadvantaged with regard to medical care services? Are there any differences in the use of medical care services among various socio-economic groups? Despite universal and comprehensive public health insurance schemes (such as the hospital and medical insurance programs in Canada), and considerable improvements in health service delivery to rural communities, these questions still remain unanswered.

In an attempt to answer these questions, a study of the utilization of medical care services – hospitals and physicians' visits – by a panel of adults residing in the island of Newfoundland for a four-year period (1992-1996) is evaluated in this thesis. It will explore the influence of access on the amount and patterns of medical care utilization for three diagnostic groups, and the associations between patterns of care and demographic and socio-economic variables, as well as health status.

1.1 Background of the study

Hospitalizations and physicians' consultations are major contributors to health care costs. They represent a combination: physician - hospital, which is the core of the traditional pattern of acute medical care. This pattern is still extremely important, despite the current dominance of chronic diseases and changes such as reductions in lengths of stay, ambulatory surgery, and more emphasis being placed on community-based services.

Medical care utilization is affected by demographic and socio-economic variables, current health status, lifestyle factors, access to resources, and so on. The association among socio-economic status, access to care, health and disease has been the subject of numerous papers. The theory and methodology relating to these associations are complex. The issues such as urban/rural differences which correspond to structured access to care, or socio-economic differences in services utilization are multifaceted and must be addressed in the context of a dynamic environment. One problem of current health services research is that services are often studied in isolation, focusing on only one type of service use, like only studying hospitalizations or physician consultations. This approach fails to capture the importance of the constellation of services available to and used by individuals; consequently, potential substitution or complementary effects are ignored. In some cases, these patterns may be interpreted as access problems when they simply reflect alternative methods of providing desired services.

Despite the profuse literature on medical care utilization, there are not many studies that use a longitudinal design. There is a lack of longitudinal studies that address significant policy questions such as the followings: What is the degree of variation in individuals' utilization of medical care services over time? How is medical need related to patterns of utilization? And how are patterns of use for different diagnoses related to different geographical locations? Medical care utilization as a longitudinal phenomenon has not been extensively studied, in part because data collection is complex and expensive. There is very little in terms of specific indicators of utilization as a longitudinal event.

Utilization is a dynamic phenomenon. Study of patterns of medical care utilization (episodes of hospitalizations and physicians' consultations) requires developing and testing complex conceptual models in order to assess the contribution of demographic, socio-economic and health status variables and their relative influence. Empirically, this task requires the development of specific computer algorithms, as common statistical computer packages are not designed for this type of analysis; there is a need to develop methods to study patterns of care over longer periods of time.

The database generated by the Newfoundland Panel on Health and Medical Care (NPHMC) (Segovia *et al.*, 1997) using the entire island provides a greater variability for some variables, and also enables analysis by different geographical aggregations and various socio-economic groups.

1.2 Purpose of the study

The purpose of this study is to identify the possible differences in patterns of medical care utilization (hospitalizations and physicians' consultations) across three geographical categories. This geographical aggregation divides the island into three areas by degree of urbanization, which corresponds to the amount and complexity of medical care resources or structured access. This study also identifies differences of patterns across various socio-economic groups, health status, and demographic variables using three diagnostic groups: cardiovascular disease, mental disease, and chronic respiratory conditions.

Episode-based comparisons are methodologically more complex than comparisons based on number of events of care, since procedures must be developed for defining episodes and identifying the services contained in an episode. Perhaps because of this additional complexity, episode-based studies have not often been carried despite their conceptual advantages.

In this thesis, episodes of care for the selected diagnostic groups are studied using an empirical definition of episodes of care over a four-year period by different geographical categories of urbanization. The following goals are accomplished:

1. To identify patterns of medical care utilization (GPs, specialists, and hospitals), identifying temporal continuity of care for all subjects who had medical care utilization in four-year study period for any diagnosis and diagnostic continuity of care for cardiovascular disease, mental disease, and chronic respiratory conditions.
2. To identify differences in various patterns of medical care utilization with respect to health status, demographic, socio-economic status, and residence variables.
3. To compare the patterns of medical care utilization for cardiovascular disease, mental disease, and chronic respiratory conditions, and to determine typical patterns of care for these three diagnoses corresponding to socio-demographic characteristics and health status.
4. To build models to assess what factors determine whether individuals have certain patterns of medical care utilization (hospitalizations, GP and specialist consultations) for cardiovascular disease, mental disease, and chronic respiratory conditions separately.

1.3 Significance of the study

In this study, the concepts of continuity and episodes of care will be used to explore patterns of utilization for different diagnostic groups. Although the method requires complex computer programs, the identification of patterns of care provides a better characterization of the process of care than just the number of events.

This study explores the relationships between health status, demographic and socio-economic variables, and patterns of medical care utilization for cardiovascular disease, mental disease, and chronic respiratory conditions separately. This material may be useful for making policy decisions about the relative priority among programs and services in the treatment of certain diseases. It can be useful to discover models that would help in understanding patterns of medical care utilization in different categories of individuals, and to probe into the possible association between this and patterns of medical care utilization for the three diagnoses separately. Three diagnostic groups may have different "intensity" and "mix" (patterns) of utilization. Better understanding of the predictors of patterns of medical care utilization among different regions and users with different diagnoses is important for policy makers and practical uses.

The data for this study are generated by NPHMC, and are obtained from medical care utilization databases. Understanding the factors that lead to the utilization of medical care and patterns of service use may assist in improving access to health care for under-served populations. A longitudinal design permits the measurement of patterns of utilization over a relatively long period of time and results in a more realistic and dynamic representation of the patient population. Modelling with this longitudinal measure of utilization is likely to result in more precise, accurate, and innovative results.

Chapter 2 Review of literature

The determinants of medical care utilization are complex, multidimensional, and are not completely understood. The vast literature on the correlates of health care utilization attests to the impossibility of addressing all aspects in one study. As a result, the following review is focused only on the essential findings of selected empirical studies that have addressed issues related to those evaluated in this present study.

2.1 Medical care utilization

Utilization studies are part of the discipline of Health Service research. Significant reviews include Greenlick *et al.* (1968), McKinlay (1972), Anderson (1973), Hershey *et al.* (1975), and Muller (1986). In Canada there is a body of work related to the introduction of universal medical insurance (Enterline *et al.*, 1973; Beck, 1973; Beck, 1974; Siemiatycki *et al.*, 1980). Some studies include analysis of utilization from the Canada Health Survey for hospital use (Broyles *et al.*, 1983) and ambulatory care (Manga, 1987); both using a framework derived from an economic approach (Newhouse and Phelps, 1976). Recently, utilization studies in Canada have included designs using administrative medical care data banks, a feature of the insurance system that offers promising opportunities for research. Excellent examples include the paper by Beck and Horne (1980) regarding co-payment in Saskatchewan, and the many articles published by Roos and Roos in Manitoba. In the last 10 years, the number of studies has become too large and diversified to give a complete review in a few short paragraphs. A significant issue in health services research is the effective use of its results by policy makers and politicians; Mechanic (1979) and Bice (1980) have reviewed and interpreted the controversy surrounding this issue.

One of the central themes of utilization studies is how the class gradient in utilization is affected by reduced financial barriers to care. The relation of personal health history to economic disability has long been known. Poverty is a persistent condition of our society. So long as poverty remains, the issue of equity in health care is kept alive. In Canada, however, a universal health insurance system has been legislated. Utilization research may be used to analyze its effects. Broyles *et al.* (1983) conclude that Canada's Medicare has resulted in equitable distribution of care, since medical needs and socio-demographic characteristics – not economic factors – have determined both probability and amount of physician care. They note that in various countries, installation of more comprehensive insurance weakens the link between use of physician services and ability to pay.

Findings on the relationships between socio-economic status (SES) and medical care utilization are more scarce and less clear. In the 1970s, Andersen and Aday (1978) found that educational level has two relatively small effects on medical care utilization, which at first sight appear to be contradictory. Firstly, a direct effect is found, showing that higher educated individuals make more use of health services. Secondly, an indirect effect is found, with higher education leading to less illness and fewer physician visits. Other research reveals that lower SES appears to be related to more use of health services for diagnosis and treatment, while higher SES appears to be related to more use of preventive services (Wan and Odell, 1981; Adler *et al.*, 1993).

The study of Black *et al.* (1995) on the utilization of hospital resources in Winnipeg shows that with respect to use of all hospital care, persons residing outside of the major urban centre are 30% more likely to have a hospitalization on an inpatient or surgical outpatient basis than persons living within Winnipeg. Rural residents have much shorter hospital stays, however, with an average length of stay for inpatient care 30% lower than for Winnipeg residents. With respect to use of short- versus long-stay care, analysis of utilization rates by length of stay reveals that Winnipeg and non-Winnipeg regions have fundamentally different patterns of hospital care utilization. Non-Winnipeg residents receive considerably more short-stay separations and days of hospital care than Winnipeg residents. They are 47% more likely to be hospitalized, have rates of separation that are 66% higher, and have rates of use of hospital days that are 37% higher than that of Winnipeg residents. By contrast, for long-stay care, Winnipeg residents have rates of hospitalization 48% higher than non-Winnipeg residents. Average lengths of stay are 23% longer and total hospital days per 1 000 residents are 79% higher for Winnipeg residents than rural residents.

2.2 Access to medical care

There is an abundant literature with studies linking problems with health care access, differences across socio-economic groups, and health consequences. Access is a complex concept encompassing several dimensions. According to Aday *et al.* (1980) "access may be defined as those dimensions which describe the potential and actual entry of a given population group to the health care delivery system." Weissman and Epstein (1993) define access as "the attainment of timely, sufficient, and appropriate health care of adequate quality such that health outcomes are maximized." One dimension of access is related to amount, type, and organizational features of the health care system; this is called "structural access". Structural access is measurable as an ecological variable by obtaining indexes such as physicians (GPs, specialists) and hospital beds (by type of hospital) per 1 000 population; the smaller the area, the more likely the validity of this approach. Gwen (1998) mentions that financial barriers are not the only reasons why people might be deterred from using medical services. Services may be in short supply, and geographical location is also a barrier to use. Rural dwellers may have to travel long distances to obtain services, and even farther to find non-medical health services. None of the countries in the European Community have achieved a fully equitable geographical distribution of services (Maynard, 1981), and this is certainly the case in Australia, New Zealand, Canada, and the United States. In Canada, policy changes in the public sector have raised new concerns about access to care in rural locations. For example, the closure of rural hospitals and short hospital stays may indicate reduced access to and less use of hospital services by rural people. In addition, studies of outpatient physician utilization show rural people have fewer visits to specialists than that of urban people.

Dansky *et al.* (1998) identify differences in hospital days, home health visits, and physician office visits across five geographical categories: 1) large metropolitan – core counties, 2) large metropolitan – fringe counties, 3) medium/lesser metropolitan counties, 4) urban non-metropolitan counties, 5) completely rural counties, using data from the Medicare Current Beneficiary Survey in the United States. Observation of the data reveals few consistent patterns of service use differentials. Residents in the urban categories (large metropolitan core, large metropolitan fringe, and medium metropolitan) tend to have more inpatient covered days and office visits than residents in the rural categories (non-metropolitan urbanized and completely rural). Residents in the completely rural category tend to have more post-acute (home health and skilled nursing facility) use than residents in other categories.

Socio-economic gradients in access to services have been mitigated by the introduction of universal health insurance in Canada (Enterline *et al*., 1973), but persons with higher levels of education and household income continue to report relatively greater use of specialist services (Enterline *et al*., 1973; McIsaac *et al*., 1997).

According to the study of Tataryn *et al*. (1995) about the utilization of physician resources for ambulatory care in Winnipeg, Winnipeg residents are much more likely (37%) to receive a consultation during the study year than are rural residents. Winnipeg residents have greater access to specialists in general. The greatest differences are in Winnipeg residents' higher rates of access to specialists for non-consultative care; Winnipeg residents are three times as likely to be seen by a medical specialist as non-Winnipeg residents. Psychiatry is exceptional for producing the largest differences in access for physician use.

2.3 Longitudinal design

Despite the profuse literature on medical care utilization, there are not many studies that use a longitudinal design. Mossey *et al.* (1989) reports that only four longitudinal studies of physician or hospital use over a period of three or more years existed prior to 1988. Since then, one additional study has been published that meets their criteria. The four studies identified by Mossey *et al.* (1989) focus primarily on longitudinal consistency of medical care utilization. Two are restricted to older adults, and the other two involve adults of all ages. Using data from the Manitoba Longitudinal Study on Aging (MLSA), Mossey and Shapiro (1985) examine physician use among older adults over an eight-year period. They have found that 60% of those who survived and remained in the community had annual numbers of physician visits that varied by no more than four visits in at least six of the eight study years. Among those who died in the first year of the study, 56% met similar criteria for consistent physician utilization levels. In the Colorado Medicare Study (CMS), McCall and Wai (1983) examine the consistency of physician utilization levels for older adults who were continuously enrolled for four years. Stump *et al.* (1995) studies the consistency of or changes in physician utilization over time. Hierarchical multivariable regression analysis of data on the 2 430 older adults, who were enrolled in the Longitudinal Study of Aging (LSOA) in United States and successfully reinterviewed in 1986, 1988, and 1990 is used to model changes in the number of physician visits between 1984 and 1990 based on the predisposing, enabling, and need (including functional status) characteristics measured in 1984, and subsequent changes in functional status. Tomiak *et al.* (1998) profiled health care utilization by disabled and non-disabled individuals in the Canadian province of Manitoba to evaluate the association between health care utilization and disability, using longitudinal data on individual encounters with the Manitoba health care system from 1983 to 1990. Associations between severity of disability, number of prior chronic conditions, and prospective utilization were examined using multivariate regressions. The findings of this study emphasize the importance of incorporating measures of disability in health services research. Roos *et al.* (1989) tracks usage of hospitals and nursing homes by a representative sample of elderly persons over a 16-year period beginning in 1970 in Manitoba, Canada. This study addresses issues like: To what extent do individuals who are high users at one point in time remain high users over subsequent time periods? What is the risk of any elderly person becoming a high user? What are the usage patterns of individuals with different expenditure levels over a 16-year period?

2.4 Episodes

There is extensive literature on episodes of care. With respect to health care episodes, Solon *et al.* (1967) suggest care episodes as a more useful method of analyzing medical care utilization.

As mentioned in Hornbrook *et al.*'s study in 1985, "episodes, as measures of output, circumscribe every service, and only those services, required to diagnose and treat each well-defined health problem. That is, episodes define boundaries for summing the total ending points to the medical care process by specifying the beginning and ending points and the course of an illness or problem. One advantage of concentrating on an episode as an output is that potentially it can specify definitive and desirable outcomes, thus facilitating appropriate valuation of the final product of the health care system." Solon *et al.* (1967) define an episode of medical care as "a block of one or more medical services received by an individual during a period of relatively continuous contact with one or more providers of service, in relation to a particular medical problem or situation". Unlike other researchers, Solon *et al.* (1967) distinguish an episode of medical care from an episode of illness. They propose that organizing medical care utilization into "episodes" is a more useful way of analyzing medical care utilization than analyzing it by units of service for each component of care. Kessler *et al.* (1980) define an episode of care as "a sequence of services involved in the diagnosis, treatment and evaluation of a problem". Feldstein (1966) emphasizes the need to use the episode approach in research on the demand for medical care. Components used together in the treatment of an illness must be considered both complementary and interchangeable. A more restrictive concept of an episode of care is defined by Young and Fisher (1980), who define the start of the care episode as admission to a hospital. Hospital stays less than 15 days apart are considered to be part of the same disease/illness episode.

The episode approach to identifying medical care utilization and costs has been used for a variety of research purposes. Costs, outcomes, and the quality of care of episodes of illnesses or care have been compared among different providers and in different settings

(Moscovice 1977; Wright *et al.*, 1977; Kane *et al.*, 1978; Lohr and Brook 1980). German *et al.* (1976) and Salkever *et al.* (1976) use episodes of illness as the indicator by which to examine access to care.

Kessler *et al.* (1980) describes a methodology for generating episodes of psychiatric care given a data set with a small amount of routinely collected data present in many medical information systems. Both demographic and medical characteristics are significantly associated with health services resources use as defined by the number of visits in an episode. A model predicting recurrent episodes of care is also described. The general utility of this approach and the substantive implications of the specific results are discussed.

The study of Wolinsky *et al.* (1994) purposes to prospectively model hospital resource use for a representative sample of older adults over seven years. After linking their administrative records and interview data, the use of Medicare-reimbursed hospital resources during 1984 through 1990 by the 7 527 Longitudinal Study of Aging (LSOA) respondents are prospectively assessed using a two-part design. First, logistic regression is used to model whether a hospital episode occurred. Second, among those having had hospital episodes, ordinary least squares (OLS) regression is used to model the number of episodes. Part of their results show that greater numbers of episodes are mostly associated with prior hospital and physician utilization, and poorer perceived health.

As mentioned in the study of Hornbrook *et al.* (1985), major limitations to the further development and application of the episode approach in research are unavailability of data and the manner in which data are recorded. Questions have been raised about the validity and reliability of the data from all of the sources previously identified, particularly diagnostic information and services rendered. Diagnoses assigned to contacts for care reflect varying degrees of uncertainty because of the way medical record systems are maintained at present. Not all services ordered by providers are certain to be entered into the medical record. The use of multiple providers can also present a problem of missing data if a patient's medical record is not maintained centrally.

2.5 Health status

Health status is a major determinant of medical care utilization. It is a complex construct, including several dimensions. Review articles were written by Bergner (1985, 1987), Patrick and Bergner (1990), Young (1998) and Patrick and Erickson (1993). The main concern of these articles is the measurement of health status. Currently available instruments used to measure health status can be broadly classified into those focusing on individual health and those emphasizing on the health of populations or communities (Roos *et al.*, 1999). According to Young's book, various definitions of health exist. In operational terms, it is useful to divide health status into various components representing different concepts or domains: opportunity, perceptions, functional status, psychological, physical, impairment, and death/duration of life. The study of Segovia *et al.* (1989) has empirically defined three dimensions: physical, emotional, and subjective. Physical health status, including number of chronic conditions, disability, ADL's, and activity restriction, is a straightforward indicator of need within a medical dimension; as such it is predominantly biological. Subjective health status suggests a more diffuse, non-medical explanation of health and illness that corresponds to a personal interpretation of these complex concepts. The emotional dimension – including social networks – is also important to consider.

The previous studies demonstrate that there is a relationship among health status and socio-economic status and medical care utilization (Statistics Canada 1994a). Individuals with low socio-economic status are more likely to have poor health status, spend more days in the hospital, and, particularly in Canada, have more contact with physicians. In an early study (Black *et al.*, 1995) conducted in Manitoba, Canada, it is found that residents of regions with the highest use of hospitals (largely for medical admissions) also had the greatest socio-economic deprivation and the poorest health status. Another study (Roos and Mustard, 1997) has evaluated variations in health status across socio-economic groups and assessed the extent to which hospitals and physicians operating under a universally funded system deliver health care according to need. Across six of the 14 mortality measures they have found a more than twofold difference in rates, indicating that the health of Winnipeg residents decline with their socio-economic status. Individuals of poorer health status are hospitalized at a much higher rate than individuals with better health status, who in turn are hospitalized more frequently than residents of better health status. This study also shows that specialist physicians provide no more care to residents of higher health needs than to residents of lower health needs.

2.6 The three diagnostic groups

2.6.1 Cardiovascular disease

Few diseases have as great an impact on society as cardiovascular disease. It is the most disabling chronic condition, a common cause of mortality and morbidity and a major consumer of health care costs in Newfoundland and even in Canada. The number of people newly affected is unknown. Despite its impact on the individual, the family, and the society, several studies have focused only on hospitalizations. Little has been done in Newfoundland or Canada to anticipate health services needs and to study patterns of medical care utilization of cardiovascular disease. Not all patients with cardiovascular disease are admitted to hospital even though the universal health care system is available in Newfoundland and Canada. Some will be seen only in physicians' offices or emergency rooms, so the study of patterns of medical care utilization for cardiovascular disease just focusing on hospitalization is not enough. It should include other sources too. For those not admitted to hospitals, a search of physician claims could yield additional patterns, which may have more comprehensive information on health care use for cardiovascular disease. Actually, unlike other chronic diseases such as arthritis, multiple sclerosis or Parkinson's diseases, which do not have a predictable pattern of health care utilization, for cardiovascular disease the health care pattern is predictable. In addition, the changes in the occurrence of cardiovascular disease are likely to be reflected in the changes in the number of records in some databases like mortality, hospital discharges, and physician claims, which are the most valuable source of data for determining health care resource utilization.

According to the literature, the hospitalization rates for all cardiovascular disease have been monitored in two provinces (Saskatchewan and Quebec) in Canada. In Quebec, the study group looked at hospital discharges over 10 years (Walker *et al.*, 1981). They have concentrated on three of the ICD-9 codes for cerebrovascular diseases (431 for intracerebral hemorrhage, 434 and 436 for ischemic stroke). The main finding is that the hospitalization rate for intracerebral hemorrhage increases over time in all age groups for men, and in the two highest age groups for women. The rates of ischemic stroke declines in younger men but increases in older persons, although not dramatically. In Saskatchewan, the rates over a four-year period appear to have been relatively stable. The average annual age-standardized hospitalization rate for all cardiovascular disease for the nine-year period from 1981 to 1989 was 151 per 100 000 population over the age of 15 years. This rate is higher than the 127 per 100 000 reported from registry data in France, but lower than the rates reported from Scandinavian countries and from Japan (more than 200 per 100 000 population) (Mayo *et al.*, 1991).

2.6.2 Mental disease

Mental disease is usually cited as one of the common conditions of persons who use medical services repeatedly, but the precise estimation of the prevalence of mental disease in the general population is unknown in Newfoundland. Regarding this, the previous studies in Ontario for the population 15 to 64 years of age (Offord *et al.*, 1996) have shown that almost 1 in 5 (18.6%) Ontarians 15 to 64 years of age had at least one of the mental disorders; among the young (15 to 24 years of age), the prevalence rates were especially high, with approximately 1 in 4 having at least one mental disorder.

A general finding is that the total utilization for mental disease has increased and a high proportion of the population is entering the care system. Increasing awareness of mental health problems and increasing acceptance of mental health treatment may have led to substantial changes in mental health utilization. The apparently increased treatment prevalence or incidence raises the question of utilization of resources when the patient enters the care system. Patterns of care reflect the interaction of features of the psychiatric care system and the help-seeking behavior of the patient. The concept of pattern of care is central to the analysis of resource utilization and is composed of several dimensions, such as the type of care, setting and intensity, and duration of care (Sytema *et al.*, 1989). The study of patterns can explore the method for a more efficient utilization of resources. It is very useful for policy makers and they can rely on results of the study to allocate health resources and produce the most clinically desirable allocation of treatment resources.

Previous studies have also shown that a number of social and demographic characteristics of the population are associated with patterns of medical care utilization for mental disease (Wells *et al.*, 1986a; Wells *et al.*, 1986b; Sytema, 1991; Simon *et al.*, 1994; Temkin-Greener and Clark, 1988.). Authors of previous studies have consistently found age, sex, socio-economic status, and geographic location to be significantly associated with the use of mental health services. However, estimates from previous studies have been limited by one or more methodological concerns. First, some authors have limited their scope to certain provider groups, such as psychiatrists or mental health specialists. Second, few authors have examined the effects of socio-demographic factors while controlling comprehensively for differences in health status, assessed independently from use of services. Third, some authors have not used multivariate techniques to estimate effects. Such techniques can enhance the precision of estimates and allow statistical control for multiple factors related to use.

Shapiro *et al.* (1984) have found specialty mental health service use to be higher among females, never married respondents, and in some study sites among those aged 25-44 years. Leaf *et al.* (1985, 1988) have found that among those reporting needs, females, individuals aged 25-44 years, and the unmarried, are more likely to use outpatient mental health services.

2.6.3 Chronic respiratory conditions

Respiratory disease is a serious problem in Newfoundland. It is the third leading cause of death in both men and women in Newfoundland and in Canada in 1992, the latest year for which statistics are available (Statistics Canada, 1994b). The total category of chronic obstructive pulmonary disease (COPD), which includes chronic bronchitis, emphysema and asthma, has accounted for 152 deaths in Newfoundland and 2,490 hospitalizations in 1992. Asthma was responsible for eight deaths and nearly half (1019) of the COPD inpatient visits (Statistics Canada, 1995b). In Canada as a whole, the total cost of asthma in 1990 was established to be $504 million to $648 million (Krahn *et al.*, 1996). Most chronic lung diseases affect older people; asthma strikes all ages.

Respiratory conditions are a chronic condition of increasing prevalence and high cost to patients, their families, and society. They adversely affecte schooling, employment, physical activities, social interaction, personal relationship, and emotional well-being (Nocon and Booth, 1991). Research into the causes and consequences of this condition has increased during the last few years, yet the study on chronic respiratory conditions in Canada is not complete, and little is known in Newfoundland. Erzen *et al.* (1997) describe the relationship between income level and the prevalence and care patterns of physician-diagnosed asthma and total respiratory morbidity (TRM) in Winnipeg, Manitoba. This study shows that there is a tendency for a larger fraction of low-income than high-income quintile patients to be seen on multiple occasions. Patterns of care reveal excellent access for low-income quintiles. People with low income have more physician contacts for asthma and TRM than those with high income.

Chapter 3 Design and ethics

3.1 Design

This study is a secondary data analysis based on the Newfoundland Panel on Health and Medical Care (NPHMC), which is a seven-year project studying factors influencing medical care utilization before and after a substantial reorganization of the provincial health care system. Utilization is measured over a seven-year period (1992-99). The restructuring began in 1994, with the adoption of regionalization and the constitution of eight Institutional Boards and five Community Health Boards. The design of NPHMC includes an initial cross-sectional telephone health survey (1994-95), with the survey sample becoming a panel for a longitudinal study of utilization. The survey respondents consenting to the use of their health insurance number were linked with the medical care utilization database, for a period of seven years, three before and four after the survey. The midpoint of the survey (April 1, 1995) corresponds approximately with the initiation of the reorganization of provincial health care system. Therefore, the first three-year period of utilization represents the situation before the restructuring, while the following four-year period will measure possible change after the reorganization in the health care system. The NPHMC used a single-stage cluster (household) random sample for the island portion of the province of Newfoundland. For this analysis, four years of data will be used, 1992-96.

The null hypotheses are:

1. There is no relation between patterns of medical care utilization (for cardiovascular

disease, mental disease, and chronic respiratory conditions separately) and the level of urbanization.

2. There is no relation between patterns of medical care utilization (for cardiovascular disease, mental disease, and chronic respiratory conditions separately) and socio-economic factors.

3. There is no relation between patterns of medical care utilization (for cardiovascular disease, mental disease, and chronic respiratory conditions separately) and demographic variables: age and sex, as well as health status.

3.2 Ethics

This study was approved by Human Investigations Committee (HIC) of the Faculty of Medicine, Memorial University of Newfoundland. Participant consent has been obtained in writting for linkage to the medical care utilization data files.

Confidentiality is maintained by not using patient identifiers on any study documents or reports to be seen by anyone other than those who are directly involved in the study. This study uses the electronic databases with the arbitrary codes from the survey as identifier. Identification data (name, address, etc.) are not included in any analysis file. All participants have been informed of their rights to refuse and to withdraw from the study at any time during the interview or later. For further information on the survey design, see "Newfoundland Panel on Health and Medical Care: Adult Health Survey 1995 – Methodology and Descriptive Results" by Segovia *et al.* (1997).

Chapter 4 Methods

Access, defined as the amount and complexity of medical care resources, is measured by proxy, dividing the sample into three areas. Residence, when SES is controlled, acts as a proxy for structural access. These areas have different amounts and types of institutional care and they also differ in the amount of physicians. Of course, these areas also differ in demographic and socio-economic status. Access is studied by using three areas: 1) Metropolitan St. John's: tertiary care, GPs, and specialists. 2) Other urban (Corridor): regional hospitals, GPs, and some specialists. 3) Remote/rural (the rest): clinics, GPs (many under salary), no specialists.

Measures of central tendency and dispersion are obtained for consultations to general practitioners and specialists, number of hospitalizations and hospital days, for the whole sample and by gender for cardiovascular disease, mental disease, and chronic respiratory conditions. For the study of episodes and patterns, the four-year study period is divided into 16 trimesters. Specific programs are developed for using SAS to identify temporal continuity and diagnostic continuity.

4.1 Data

This NPHMC datafile includes 9,062 subjects with four years of medical care utilization, 20 years and older, residing in the province of Newfoundland, Canada. This population is served by tertiary or regional hospitals, and by general practitioners and specialists on fee-for-service salary. It should be noted that visits to salary physicians for the remote/rural area have been excluded, since this information is not collected in the physicians' claims database. As with all Canadian provinces, all permanent residents are covered by universal health insurance.

There are three data files generated by NPHMC: one contains the survey data, and the others contain data on medical care utilization: hospital separations and physicians' claims in separate data files, obtained by linkage with the medical care utilization databases through the MCP number (health insurance number).

Computer files for all data were maintained on the Digital UNIX V4.0D at Memorial University and linkage between the survey and utilization files was effected by matching on the identification number (unique for each subject in the survey).

This current study focuses on ambulatory physician visits which included outpatient, emergency room, and office visits, and excludes physician visits to hospitalized patients and visits to nursing home residents.

All records that are selected for this study in medical care utilization data files contain a principal diagnosis of cardiovascular disease, mental disease, or chronic respiratory conditions, coded according to the international Classification of Disease, 9th Revision, Clinical Modification (ICD-9-CM) code. All records also contain place of residence and some indicators of medical care utilization. These are matched with selected demographic, socio-economic, health status, and geographic variables from the survey data file. The study subjects for diagnostic groups are those who were resident in the province of Newfoundland for the four-year study period (1992-1996) with primary diagnosis as cardiovascular disease, mental disease, or chronic respiratory conditions. These three diagnostic groups have been selected because of their range in coverage of the population and their varied mix of use of GPs, specialists, and hospitalizations. In addition to these three subgroups of the NPHMC, the entire 9,062 respondents are analyzed to obtain an overall view of utilization pattern in the population.

4.2 Variables

This thesis is concerned with how the patterns of hospitalizations and physicians' consultations differ between subgroups, defined according to geographic categories, socio-demographic variables, or health status for different diagnostic groups.

The variables selected in this study are as following:

1. Demographic variables:

   1.1. SEX: nominal.

   1.2. AGE: three kinds of variables are used in modelling:

      (1). Continuous (AGE);

      (2). Integer: truncated the continuous age to the nearest integer (INTEAGE);

      (3). Grouped in 6 categories: 20-29, 30-39, 40-49, 50-59, 60-69, and 70+ (AGEGP6).

2. Socio-economic variables:

   2.1. Income adequacy: Five groupings of income adequacy are used in this study, with 1 to 5 representing very low, low, lower middle, upper middle, and high income.

Income adequacy (a household variable) is calculated using reported income, the number of individuals per household, and the provincial low income cutoff points (LICOs), following procedures used in the Canada Census (Segovia *et al.*, 1997). See Table A1 in Appendix for the income levels for different household sizes.

2.2. Educational level: Divided into five categories: less than high school, high school, trade school or diploma, university without degree, and university with degree.

2.3. Socio-economic score: Two variables are used in the analysis. One is SES9 with nine levels (1-9), which is a simple additive index, including education level and income adequacy. It is scored 1 to 9, representing different level of socio-economic status from low to high, and simply calculated by adding income and education levels together and then minus one. There are five levels for income and education separately, so there are nine levels for socio-economic score. The socio-economic score is also divided into three levels (SES3), from one to three indicating the socio-economic status from low to high based on SES9 through combining three levels together (SES3: 1, 1-3 of SES9; 2, 4-6 of SES9; 3, 7-9 of SES9).

3. Health status: Self-assessed health status (HSTAT) is used in this study. It is divided into four categories: 1. Excellent, 2. Good, 3. Fair, and 4. Poor. For this variable, the lower the value, the better the self-assessed health status.

4. Residence Area: URBAN: nominal variable in three categories: 1) Metropolitan St. John's, 2) Other Urban, and 3) Remote. When SES is controlled, residence is a proxy

for structural access.

5. Utilization:

5.1. Hospital separations: The information obtained from the hospitalization files includes the identification number, hospital location, admission date, separation type (alive or dead), length of stay, and up to three diagnoses coded by ICD-9-CM code.

5.2. Physicians' claims: The information obtained from physicians' claims for physician visits includes the identification number, diagnosis code, fee code, and a non-identifiable code for the physician. Only ambulatory visits are used in this study, which are defined as visits to doctor's office, patient's home, or hospital outpatient department. Emergency department visits are included in the outpatient category.

5.3. Diagnosis: The study subjects for cardiovascular disease analysis are selected from the databases of hospital separations, or GPs', or specialists' consultations for those subjects whose primary diagnosis coded as: ICD-9-CM 410, 411-414, 431, 434, and 436, including acute myocardial infarction, other acute and subacute forms of ischaemic heart disease, old myocardial infarction, angina pectoris, other forms of chronic ischaemic heart disease, intracerebral haemorrhage, occlusion of cerebral arteries, and acute but ill-defined cerebrovascular disease. The total number of subjects is 678.

The subjects for mental disease are selected by the primary diagnosis in the database of hospital separations and the database of physician consulations. The diagnostic codes are ICD-9-CM 295-298, 303, and 309, including schizophrenic psychoses, affective psychoses, paranoid states, other nonorganic psychoses, alcohol dependence syndrome, and adjustment reaction. The total number of subjects is 402.

The subjects for chronic respiratory conditions are selected from the databases of hospital separations, and GPs' or specialists' consultations for those subjects whose primary diagnosis is coded as: ICD-9-CM 490-494 and 496, including bronchitis, chronic bronchitis, emphysema, asthma, bronchiectasis, and chronic airways obstruction, not elsewhere classified. The total number of subjects is 942.

6. Development of measurements of continuity/episodes: SAS programs have been designed to identify temporal continuity and diagnostic continuity.

In this study, SPSS data files are transferred to SAS files, several variables are combined, and some abbreviated files containing only those variables satisfying selected criteria are created. The development of definitions for temporal and diagnostic continuity follows considerable investigation and development of new variables.

First of all, for GP visits, the four-year study period is divided into 16 trimesters, the corresponding 16 new variables (t1 - t16) are created separately representing the number of GP visits in each trimester from April 1, 1992 to March 31, 1996 for each diagnostic group. Then two new variables, totvisit (the total number of GP visits over

16 trimesters) and maxvisit (the maximum number of GP visits in any trimester), are created based on t1 - t16. According to these two variables, episodes are categorized into EPITYPE (11 categories from 0 to 10, shown in Appendix B). This variable is used to get a preliminary examination of the data.

In order to identify the possible differences in patterns of care for GP and specialist visits, other variables are created, such as bepichar and comb.

6.1 BEPICHAR: identifies trimesters with two or more visits. To construct this variable, make 16 variables as follows: If the first trimester has two or more visits, make the first variable=1, otherwise equal to 0; if the second trimester has two or more visits, make the second variable =10, otherwise equal to 0; if the third trimester has two or more visits, make the third variable=100, otherwise equal to 0, etc. This will give 16 variables with values from 1, 10, 100, 1000 up to 1 000 000 000 000 000 when all 16 trimesters have two or more visits. Summing these 16 variables results in a (maximum) 16 digit number with the right-most digit corresponding to the first trimester. Converting this variable to characters gives a 16 character variable with 0's and 1's. This gives a look of the trimesters with two or more visits over time.

Example 1: 0000000000100111 is a case where there are two or more visits in the first, second, third, and sixth trimesters. All other trimesters have one or no visit.

6.2. COMB: identify trimesters with two or more visits. Count the number of consecutive trimesters with two or more visits and create a 16 character variable where the first character implies "one" trimester with two or more visits, the second implies "two" consecutive trimesters with two or more visits, etc.

Example 1: 1011000000000000 is a case with "four" consecutive trimesters with two or more visits, "three" consecutive trimesters with two or more visits, and "one" single non-consecutive trimester with two or more visits. This corresponds to a case where BEPICHAR is 0011110001110010.

Example 2: 0000001100000000 is a case with "seven" and "eight" consecutive trimesters with two or more visits, i.e. there is only "one" trimester (16-(7+8)) that does not have two or more visits. This corresponds to a case where BEPICHAR is 1111111101111111.

The above was done separately for GP and specialist visits.

7. Definitions related to this process:

7.1. Episode:

An episode can be defined as a continuous series of contacts within a given observation period. In this study, an episode can be defined as two or more visits in a trimester.

7.2. Continuity:

3a: Temporal continuity: defined as two or more visits in a trimester or in two or more consecutive trimesters. Five main patterns have been identified: 1) Non-episodic: when no trimesters (T) have two or more visits; 2) Isolated episodes (within single trimester), when there are some non-consecutive T's with two or more visits (i.g. xx2xx5xxxxxxxx4x); 3) Continuous episodes: when there are some consecutive T's with 2 or more visits, which was further subdivided into: continuous short (two to five consecutive trimesters), medium (six to 11 consecutive trimesters), and long episodes (12 to 16 consecutive trimesters).

Example 1: Short (2-5 T's): xx3254xxxx72xxxx: This example shows that the first two trimesters had no visits, the third to the sixth had three, two, five and four visits respectively, hence there are four trimesters with two or more visits.

Medium (6-11 T's): xxx2368232xxxxxx: in this example, there are seven trimesters (from the fourth to the tenth) with two or more visits.

Long (12-16 T's): 274222633423xxxx: in this example, there are 12 trimesters with two or more visits (from the first to the twelfth).

3b: Diagnostic continuity: the same as temporal, but with the same diagnostic group (ICD-9-CM) for visits to GPs and specialists for cardiovascular disease, mental disease, and chronic respiratory conditions separately.

4.3 Analysis

The relationship between patterns of medical care utilization and the selected variables (age, sex, SES9, health status, and urban) mentioned above is examined individually. A multivariate analysis assesses their combined value as predicators of service demand. In the multivariate analysis, there are two dependent variables: 1) multi-level for GP and specialist visits, indicating different patterns of medical care utilization; 2) two levels for hospitalizations, indicating hospitalized or not.

Descriptive statistics are calculated, first for individual characteristics of all respondents in the NPHMC sample for each geographical category based on demographic, socio-economic variables, health status, and diagnosis. Then we examine the distribution of patterns of care (including the patterns of GP visits, patterns of specialist visits, and patterns of hospitalizations) by the selected variables including age, sex, health status, socio-economic score and urbanization for subjects with cardiovascular disease, mental disease, or chronic respiratory conditions separately.

All variables in the descriptive analysis are included in the multivariate analyses. Logistic regression models are used, as they are well suited to the analysis of categorical variables. Nominal logistic regression equations, which allow a multi-level dependent variable, and binary logistic regression equations are used to assess the effects of the independent variables and covariates as predictors of each pattern of service use. Table 4.1 shows the independent variables and their categories included in this analysis.

In the multivariate analysis, binary and nominal logistic regressions were performed to examine the association between demographic variables, socio-economic status, health status and urbanization, and patterns of GP or specialist visits, as well as the probability of being hospitalized for the selected three diagnostic groups. The dependent variables used in these logistic regressions are: 1) patterns of GP visits (0: non-episodic; 1: isolated episodic; and 2: continuous episodic), 2) pattern of specialist visits (0: non-episodic; 1:

isolated episodic; and 2: continuous episodic), and 3) Hospitalizations (0: not hospitalized; 1: hospitalized). Nominal logistic regression was separately performed on the patterns of GP visits and specialist visits for each study group. Binary logistic regression was performed on the hospitalization data. Table 4.2 shows the dependent variables and their categories included in this analysis.

Nominal logistic models are assessed by checking the log-likelihood and its probability (which in this case should be significant), as well as by inspecting the Pearson and Deviance values for the goodness-of-fit tests of model (which in this case should be insignificant). Binary logistic models are examined by checking the log-likelihood and its probability (which in this case should be significant), as well as primarily inspecting the Hosmer-Lemeshow value for the goodness-of-fit tests of model (which in this case should be insignificant).

Table 4.1 Logistic regression – independent variables and categories

Variables	Categories
Age	Agegp6:1: 20-29, 2: 30-39, 3: 40-49, 4: 50-59, 5: 60-69, 6: 70+ Age: continuous Inteage: integer type
Sex	1: male; 2: female
Health status	1: Excellent, 2: good, 3: fair, 4: poor
Socio-economic score	SES9:1 – 9: from the lowest to the highest
Urban	1: metropolitan St. John's, 2: Other urban, 3: Remote

Table 4.2 Logistic regression – dependent variables and categories

Service Use	Binary Logistic	Nominal Logistic
GP visits		Pattern: 0, 1, and 2
Specialist visits		Pattern: 0, 1, and 2
Hospitalizations	Hospitalized: 0: no; 1: yes	

Chapter 5 Results

5.1 General characteristics of study subjects

Tables 5.1-5.3 show demographic and socio-economic characteristics for cardiovascular disease, mental disease, and chronic respiratory conditions separately, as well as the mean of demographic and socio-economic characteristics by diagnostic groups. When comparisons of the six variables shown in the three tables for residence are examined separately for the three diagnostic groups, the level of income and education, as well as socio-economic score decrease from metropolitan St. John's to the remote area and there are significant differences among different areas for the three diagnostic groups respectively, meaning the respondents who live in St. John's have higher level of income and education, as well as socio-economic score than those who live in other urban and the remote for the three diagnostic groups respectively (For cardiovascular disease, the p-value for SES3 is 0.000, income is 0.000, education is 0.000. For mental disease, the p-value for SES3 is 0.000, income is 0.008, education is 0.017. For chronic respiratory conditions, the p-value for SES3 is 0.000, income is 0.000, education is 0.000). From the three tables, it can be clearly seen that the subjects in remote areas are significantly older than those in metropolitan St. John's and other urban for cardiovascular disease and mental disease (For cardiovascular disease, the p-value is 0.003; for mental disease, the p-value is 0.019). Health status has significant differences among three areas for cardiovascular disease and chronic respiratory conditions. Subjects in remote areas have poorer health status than those in metropolitan St. John's and other urban for the two diagnostic groups (the p-value of health status for cardiovascular disease is 0.000; for chronic respiratory condition, the p-value is 0.000). There is no significant influence of sex by urbanization for any of the three diagnostic groups. Upon comparison of the corresponding mean of each variable for the three diagnostic groups and all subjects who had medical care utilization for any diagnosis in the four-year study period separately (Table 5.1-5.4). It is found that the subjects with cardiovascular disease are much older and less healthy, as well as have lower income, education levels and socio-economic score than those subjects with mental disease and chronic respiratory conditions, regardless of where they live. The differences of age and health status among the three diagnostic groups compared to the whole sample are significant, in that the whole sample is much younger and has better health status than the three diagnostic groups (Table 5.4).

Table 5.1 Characteristics for cardiovascular disease

Characteristics		Sample						
	St John's		Other Urban		Remote		Total	
	No	%	No	%	No	%	No	%
Total	219	32.3	218	32.2	241	35.5	678	100
Sex[1] mean	1.45		1.48		1.44		1.45	
Male	121	32.7	113	30.5	136	36.8	370	100
Female	98	31.8	105	34.1	105	34.1	308	100
Age[2] mean	69.1		70.3		72.4		70.7	
20-39	9	31.0	14	48.3	6	20.7	29	100
40-59	84	41.4	60	29.6	59	29.1	203	100
60+	126	28.3	144	32.3	176	39.5	446	100
SES3[3] mean	1.91		1.49		1.19		1.51	
Low	69	17.3	130	32.7	199	50.0	398	100
Middle	92	47.7	67	34.7	34	17.6	193	100
High	49	66.2	19	25.7	6	8.1	74	100
Health Status[4] mean	1.33		1.37		1.53		1.41	
Exc-good	147	36.9	137	34.4	114	28.6	398	100
Fair-poor	72	25.7	81	28.9	127	45.4	280	100
Income[5] mean	3.41		2.79		2.45		2.86	
Very low	9	15.8	19	33.3	29	50.9	57	100
Low	40	17.6	75	33.0	112	49.3	227	100
Lower middle	64	30.6	74	35.4	71	34.0	209	100
Upper middle	49	52.1	29	30.9	16	17.0	94	100
High	48	61.5	19	24.4	11	14.1	78	100
Education[6] mean	2.34		1.81		1.3		1.8	
< High school	89	21.4	126	30.4	200	48.2	415	100
High school	37	37.8	39	39.8	22	22.4	98	100
Trade sch/diploma	48	53.3	30	33.3	12	13.3	90	100
Univ-no degree	19	54.3	15	42.9	1	2.9	35	100
Univ-with degree	26	65.0	8	20.0	6	15.0	40	100

[1]Chi-Sq: 1.036; DF: 2; P-Value: 0.596; [2]Chi-Sq: 16.355; DF: 4; P-Value: 0.003; [3]Chi-Sq: 126.727; DF: 4; P-Value: 0.000; [4]Chi-Sq: 20.867; DF: 2; P-Value: 0.000; [5]Chi-Sq: 90.365; DF: 8; P-Value: 0.000; [6]Chi-Sq: 103.331; DF: 8; P-Value: 0.000.

Table 5.2 Characteristics for mental disease

Characteristics		Sample						
	St John's		Corridor		Remote		Total	
	No	%	No	%	No	%	No	%
Total	182	45.3	98	24.4	122	30.3	402	100
Sex[1] mean	1.61		1.60		1.63		1.61	
Male	71	39.0	39	39.8	45	36.9	155	100
Female	111	61.0	59	60.2	77	63.1	247	100
Age[2] mean	53.8		51.6		54.6		53.5	
20-39	64	35.2	50	51.0	45	36.9	159	100
40-59	94	51.6	31	31.6	54	44.3	179	100
60+	24	13.2	17	17.3	23	18.9	64	100
SES3[3] mean	1.91		1.60		1.48		1.71	
Low	56	31.3	51	52.0	67	55.8	174	100
Middle	83	46.4	35	35.7	48	40.0	166	100
High	40	22.3	12	12.2	5	4.2	57	100
Health Status[4] mean	1.31		1.41		1.42		1.37	
Exc-good	125	68.7	58	59.2	71	58.2	254	100
Fair-poor	57	31.3	40	40.8	51	41.8	148	100
Income[5] mean	3.1		2.7		2.7		2.9	
Very low	19	10.6	14	14.3	11	9.2	44	100
Low	38	21.2	32	32.7	38	31.7	108	100
Lower middle	62	34.6	36	36.7	51	42.5	149	100
Upper middle	35	19.6	6	6.1	13	10.8	54	100
High	25	14.0	10	10.2	7	5.8	42	100
Education[6] mean	2.7		2.2		1.8		2.3	
< High school	47	25.8	41	41.8	66	54.1	154	100
High school	37	20.3	19	19.4	21	17.2	77	100
Trades sch/diploma	47	25.8	26	26.5	27	22.1	100	100

[1]Chi-Sq: 0.223; DF: 2; P-Value: 0.894; [2]Chi-Sq: 11.736; DF: 4; P-Value: 0.019; [3]Chi-Sq: 30.676; DF: 4; P-Value: 0.000; [4]Chi-Sq: 4.343; DF: 2; P-Value: 0.114; [5]Chi-Sq: 20.719; DF: 8; P-Value is 0.008; [6]Chi-Sq: 11.98; DF: 4; P-Value: 0.017.

Table 5.3 Characteristics for chronic respiratory conditions

Characteristics	Sample						Total	
	St John's		Corridor		Remote			
	No	%	No	%	No	%	No	%
Total	367	39.0	312	33.1	263	27.9	942	100
Sex[1] mean	1.61		1.58		1.54		1.58	
Male	145	36.6	130	32.8	121	30.6	396	100
Female	222	40.7	182	33.3	142	26.0	546	100
Age[2] mean	54.8		55.9		59.3		56.1	
20-39	149	42.3	117	33.2	86	24.4	352	100
40-59	130	39.4	113	34.2	87	26.4	330	100
60+	88	33.8	82	31.5	90	34.6	260	100
SES3[3] mean	2.01		1.61		1.44		1.72	
Low	88	21.9	154	38.4	159	39.7	401	100
Middle	177	46.1	124	32.3	83	21.6	384	100
High	93	66.0	33	23.4	15	10.6	141	100
Health Status[4] mean	1.23		1.34		1.38		1.31	
Exc-good	282	43.5	205	31.6	162	25.0	649	100
Fair-poor	85	29.0	107	36.5	101	34.5	293	100
Income[5] mean	3.4		2.8		2.5		3.0	
Very low	28	28.3	36	36.4	35	35.4	99	100
Low	53	23.1	84	36.7	92	40.2	229	100
Lower middle	113	33.9	121	36.3	99	29.7	333	100
Upper middle	93	61.2	40	26.3	19	12.5	152	100
High	71	62.8	30	26.5	12	10.6	113	100
Education[6] mean	2.7		2.0		1.8		2.3	
< High school	87	22.6	142	36.9	156	40.5	385	100
High school	83	43.0	70	36.3	40	20.7	193	100
Trades sch/diploma	92	47.7	65	33.7	36	18.7	193	100
Univ-no degree	60	61.9	17	17.5	20	20.6	97	100
Univ-with degree	45	60.8	18	24.3	11	14.9	74	100

[1]Chi-Sq: 2.681; DF: 2; P-Value: 0.262; [2]Chi-Sq: 9.105; DF: 4; P-Value: 0.059; [3]Chi-Sq: 109.507; DF: 4; P-Value: 0.000; [4]Chi-Sq: 18.827; DF: 2; P-Value: 0.000; [5]Chi-Sq: 101.61; DF: 8; P-Value: 0.000; [6]Chi-Sq: 102.013; DF: 8; P-Value: 0.000.

Table 5.4 Characteristics for the three diagnostic groups and all subjects

Characteristics	All subjects N=9062		Cardiovascular N=678		Mental N=402		Respiratory N=942	
	N	%	N	%	N	%	N	%
Sex[1] mean	1.53		1.45		1.61		1.58	
Male	4249	46.9	370	54.6	155	38.6	396	42.0
Female	4813	53.1	308	45.4	247	61.4	546	58.0
Age[2] mean	53.1		70.7		53.5		56.1	
20-39	3970	43.8	29	4.3	159	39.6	352	37.4
40-59	3440	38.0	203	29.9	179	44.5	330	35.0
60+	1652	18.2	446	65.8	64	15.9	260	27.6
SES3[3] mean	1.82		1.51		1.71		1.72	
Low	3398	37.5	398	59.8	174	43.8	401	43.3
Middle	3806	42.0	193	29.0	166	41.8	384	41.5
High	1746	19.3	74	11.1	57	14.4	141	15.2
Health Status[4] mean	1.20		1.41		1.37		1.31	
Exc-good	7221	79.7	398	58.7	254	63.2	649	68.9
Fair-poor	1841	20.3	280	41.3	148	36.8	293	31.1

[1]Chi-Sq: 5.913; DF: 3; P-Value: 0.116; [2]Chi-Sq: 86.499; DF: 6; P-Value: 0.000;
[3]Chi-Sq: 11.382; DF: 6; P-Value: 0.077; [4]Chi-Sq: 11.476; DF: 3; P-Value: 0.009

5.2 Mean utilization by residence for three diagnostic groups

When examining the means of GP visits, specialist visits, number of hospitalizations, and mean of length of stay for three diagnostic groups, there are some significant differences between areas for the three different diagnostic groups. Table 5.5 shows the mean of GP visits, specialist visits, and number of hospitalizations by urbanization for cardiovascular, mental, and respiratory diagnoses. From this table, for cardiovascular disease we can see there is very little difference between areas for GP visits (4.6 in St. John's, 4.9 in other urban, and 4.5 in remote), but there are significant differences in specialist visits and number of hospitalizations, as well as mean of length of stay. The subjects in St. John's have many more specialist visits for cardiovascular disease than those in other urban or remote areas, and they have fewer hospitalizations than those in other urban or remote.

For mental disease in different areas, there are significant differences between number of specialist visits and number of hospitalizations. Table 5.5 shows very clearly that the respondents in St. John's have more specialist visits than those in other areas, but respondents in remote areas have more hospitalizations than those in St. John's and many more LOS days. There are no major differences between areas for GP visits.

For chronic respiratory conditions in Table 5.5, there is no significant trend observed for GP visits and specialist visits. Hospitalizations are more common for other urban and remote.

Table 5.5 Mean of GP visits, specialist visits, number of hospitalizations and mean of LOS by urbanization for the three diagnostic groups

	St. John's		Other urban		Remote	
	Mean	N	Mean	N	Mean	N
Cardiovascular						
GP visits	4.6	183	4.9	192	4.5	183
Specialist visits*	3.0	99	0.9	74	1.1	95
Hospitalizations*	0.4	47	0.3	53	0.7	89
LOS*	18.3	47	11.5	53	19.8	89
Mental						
GP visits	2.3	90	1.3	65	1.9	82
Specialist visits*	7.8	108	2.6	31	2.0	38
Hospitalizations*	0.2	14	0.3	18	0.4	25
LOS	40.5	14	28.6	18	52.7	25
Respiratory						
GP visits	3.6	355	3.4	292	3.5	241
Specialist visits	0.4	39	0.3	33	0.3	31
Hospitalizations*	0.04	11	0.16	19	0.24	27
LOS	8.4	11	21.8	19	14.3	27

*indicating significant difference among areas by Bonferroni method in Post Hoc tests

5.3 Descriptive analysis for the three diagnostic groups

This section presents the descriptive analyses of patterns of GP or specialist visits for those subjects with cardiovascular disease, mental disease, and chronic respiratory conditions, as well as percentages of subjects with utilization for the three diagnostic groups and all subjects. In terms of the analysis of patterns of care, five kinds of episodes are identified as discussed in the methodology section (p34). From tables 5.6-5.8, it can be seen that most subjects, whatever the diagnosis group being studied, have *no episodes* or *isolated episodic care*, less have *short continuous*, and few have *medium and long continuous episodic care* for GP or specialist visits. Since the number of subjects with *short, medium* and *long continuous episodes* of care are few for the three diagnostic groups, these three episode categories were combined together for the following descriptive analysis. The results will show three kinds of patterns of care for GP and specialist visits: *no episodes*, *isolated episodes*, and *continuous episodes*. The next section focuses on the association between patterns of medical care utilization and demographic, socio-economic variables including sex, age, socio-economic score, the degree of urbanization, as well as health status.

Table 5.9 shows the differences of percentage of subjects with utilization for the three diagnostic groups and all subjects with utilization in the four-year study period. The respondents with chronic respiratory conditions have more GP visits and less specialist visits and hospitalizations compared to the other two diagnostic groups and all subjects. Respondents with mental disease have fewer GP visits than other groups. Respondents with cardiovascular disease have more hospitalizations compared to other groups.

Table 5.6 Distribution of patterns of care for subjects with cardiovascular disease in four-year study period by gender (%)

	GP Visits			Specialist Visits		
	All	Male	Female	All	Male	Female
N of All	678	370	308	678	370	308
N of Subj. with visit	558	299	259	268	165	103
Patterns						
No Episodes	324 (58.1)	172 (57.5)	152 (58.7)	176 (65.7)	103 (62.4)	73 (70.9)
Isolated Episodes	143 (25.6)	76 (25.4)	67 (25.9)	71 (26.5)	49 (29.7)	22 (21.4)
Cont. Episodes						
Short	75(13.4)	44(14.7)	31(12.0)	19(7.1)	11(6.7)	8(7.8)
Medium	14(2.5)	6(2.0)	8(3.1)	1(0.4)	1(0.6)	-
Long	2(0.4)	1(0.3)	1(0.4)	1(0.4)	1(0.6)	-

Table 5.7 Distribution of patterns of care for subjects with mental disease in four-year study period by gender (%)

	GP Visits			Specialist Visits		
	All	Male	Female	All	Male	Female
N of All	402	155	247	402	155	247
N of Subj. with visit	237	82	155	177	77	100
Patterns						
No Episodes	173 (73.0)	64 (78.1)	109 (70.3)	64 (36.2)	23 (29.9)	41 (41)
Isolated Episodes	50(21.1)	12(14.6)	38(24.5)	56(31.6)	26(33.8)	30(30)
Cont. Episodes						
Short	13(5.5)	5(6.1)	8(5.2)	42(23.7)	19(24.7)	23(23)
Medium	-	-	-	6(3.4)	4(5.2)	2(2)
Long	1(0.4)	1(1.2)	-	9(5.1)	5(6.5)	4(4)

Table 5.8 Distribution of patterns of care for subjects with chronic respiratory conditions in four-year study period by gender (%)

	GP Visits			Specialist Visits		
	All	Male	Female	All	Male	Female
N of All	942	396	546	942	396	546
N of Subj. with visit	888	377	511	103	36	67
Patterns						
No Episodes	598 (67.3)	260 (69.0)	338 (66.1)	80 (77.7)	30 (83.3)	50 (74.6)
Isolated Episodes	229(25.8)	87(23.1)	142(27.8)	15(14.6)	3(8.3)	12(17.9)
Cont. Episodes						
Short	51(5.7)	22(5.8)	29(5.7)	7(6.8)	3(8.3)	4(6.0)
Medium	9(1.0)	7(1.9)	2(0.4)	1(0.97)	-	1(1.5)
Long	1(0.1)	1(0.3)	-	-	-	-

Table 5.9 Percentage of subjects with utilization (GP or specialist visits or hospitlizations) for the three diagnostic groups and all subjects

Utilization	Cardiovascular	Mental	Respiratory	All
GP visits	82.3	58.9	94.3	90.1
Specialist visits	48.0	44.0	13.4	62.9
Hospitalizations	27.9	14.2	6.0	25.7

5.3.1 Cardiovascular disease

With respect to the association between patterns of GP visits and the selected variables for cardiovascular disease, some trends are identified. The association between patterns of GP visits and the variables, age, socio-economic score, and health status, are significant (the p-value for age is 0.005; for socio-economic score is 0.024; for health status is 0.018). There is a clear trend in the patterns of care for different age groups, in that the older subjects are more likely to have *isolated* or *continuous episodic care* and less likely to have *no episodic care* than the younger. When the association between patterns of GP visits by socio-economic status is examined, *continuous* or *isolated episodes* are more likely to occur in those in the lowest socio-economic group (Table 5.10). The association between patterns of GP visits for cardiovascular disease and health status is also strong. Table 5.10 shows a trend that subject with fair or poor health status are more likely to have *isolated* or *continuous episodic care* and less likely to have *non-episodic care* than those with excellent or good health status. There is no significant difference between gender for the pattern of GP visits.

For specialist visits, urbanization is the only variable showing a significant difference. Table 5.11 shows a very clear trend that respondents in St. John's are more likely to have *continuous episodic care* and less likely to have *non-episodic care* than those in other urban and remote.

Examination of the association between distribution of number of hospitalizations and the selected variables for cardiovascular disease (Table 5.12) revealed that the differences are significant for AGE, URBAN, SES3, and HSTAT respectively. Respondents who are older, live in remote areas, have lower socio-economic scores, or have poorer health status normally have more hospitalizations than those respondents who are younger, live in St. John's or other urban, have higher socio-economic scores, or have good health status. There is no significant difference between males and females.

Table 5.10 Patterns of GP visits for cardiovascular disease by selected variables in four years (%)

	N	N of subj. with visits	No Episodes	Isolated Episodes	Continuous Episodes
GP Visits	678	558(82.3)	324(47.8)	143(21.1)	91(13.4)
Sex[1]					
Male	370	299(80.8)	172(57.5)	76(25.4)	51(17.1)
Female	308	259(84.1)	152(58.7)	67(25.9)	40(15.5)
Age[2]					
20-39	29	22(75.9)	16(72.7)	4(18.2)	2(9.1)
40-59	203	156(76.8)	108(69.2)	31(19.9)	17(10.9)
60+	446	380(85.2)	200(52.6)	108(28.4)	72(19.0)
Urban[3]					
St. John's	219	183(83.6)	112(61.2)	44(24.0)	27(14.7)
Other urban	218	192(88.1)	109(56.8)	52(27.1)	31(16.2)
Remote	241	183(75.9)	103(56.3)	47(25.7)	33(18.1)
SES3[4]					
Low	398	327(82.2)	173(52.9)	88(26.9)	66(20.2)
Medium	193	162(83.9)	103(63.6)	39(24.1)	20(12.3)
High	74	57(77.0)	39(68.4)	14(24.6)	4(7.0)
Health Status[5]					
Exc-Good	398	332(83.4)	206(62.0)	83(25)	43(12.9)
Fair-Poor	280	226(80.7)	118(52.2)	60(26.5)	48(21.2)

GP visits: [1]Chi-Sq: 0.265; DF: 2; P-Value: 0.876; [2]Chi-Sq: 14.872; DF: 4; P-Value: 0.005; [3]Chi-Sq: 1.412; DF: 4; P-Value: 0.842; [4]Chi-Sq: 11.284; DF: 4; P-Value: 0.024; [5]Chi-Sq: 8.029; DF: 2; P-Value: 0.018.

Table 5.11 Patterns of specialist visits for cardiovascular disease by selected variables in four years (%)

	N	N of subjects with visits	No Episodes	Isolated Episodes	Continuous Episodes
Specialist Visits	678	268(39.5)	176(26.0)	71(10.5)	21(3.1)
Sex[1]					
Male	370	165(44.6)	103(62.4)	49(29.7)	13(7.9)
Female	308	103(33.4)	73(70.9)	22(21.4)	8(7.8)
Age[2]					
20-39	29	6(20.7)	5(83.3)	1(16.7)	-
40-59	203	87(42.9)	57(65.5)	23(26.4)	7(8.1)
60+	446	175(39.2)	114(65.1)	47(26.9)	14(8)
Urban[3]					
St. John's	219	99(45.2)	60(60.6)	25(25.3)	14(14.1)
Other urban	218	74(33.9)	49(66.2)	24(32.4)	1(1.4)
Remote	241	95(39.4)	67(70.5)	22(23.2)	6(6.3)
SES3[4]					
Low	398	154(38.7)	109(70.8)	37(24.0)	8(5.2)
Middle	193	73(37.8)	43(58.9)	22(30.1)	8(11.0)
High	74	35(47.3)	20(57.1)	10(28.6)	5(14.3)
Health Status[5]					
Exc-Good	398	144(36.2)	91(63.2)	41(28.5)	12(8.3)
Fair-Poor	280	124(44.3)	85(68.5)	30(24.2)	9(7.3)

SP visits: [1]Chi-Sq: 2.354; DF: 2; P-Value: 0.308; [2]Chi-Sq: 1.001; DF: 4; Chi-Sq cannot be computed; [3]Chi-Sq: 11.447; DF: 4; P-Value: 0.022; [4]Chi-Sq: 6.322; DF: 4; P-Value: 0.176; [5]Chi-Sq: 0.850; DF: 2; P-Value: 0.654.

Table 5.12 Distribution (and percentage) of hospitalizations for cardiovascular disease by the selected variables in four years

	N	N of Hosp.	No Hosp.	One visit	Two or more
Total	678	189(27.9)	489(72.1)	117(17.3)	72(10.6)
Sex[1]					
Male	370	115(31.1)	255(68.9)	68(18.4)	47(12.7)
Female	308	74(24.0)	234(76.0)	49(15.9)	25(8.1)
Age[2]					
20-39	29	4(13.8)	25(86.2)	-	4(13.8)
40-59	203	47(23.2)	156(76.8)	29(14.3)	18(8.9)
60+	446	138(30.9)	308(69.1)	88(19.7)	50(11.2)
Urban[3]					
St. John's	219	47(21.5)	172(78.5)	28(12.8)	19(8.7)
Other urban	218	53(24.3)	165(75.7)	38(17.4)	15(6.9)
Remote	241	89(36.9)	152(63.1)	51(21.2)	38(15.8)
SES3[4]					
Low	398	130(32.7)	268(67.3)	75(18.8)	55(13.8)
Middle	193	40(20.7)	153(79.3)	26(13.5)	14(7.3)
High	74	15(20.3)	59(79.7)	13(17.6)	2(2.7)
Health Status[5]					
Exc-good	398	101(25.4)	297(74.6)	69(17.3)	32(8.0)
Fair-poor	280	88(31.4)	192(68.6)	48(17.1)	40(14.3)

[1]Chi-Sq: 5.082; DF: 2; P-Value: 0.079; [2]Chi-Sq: 10.641; DF: 4; P-Value: 0.031; [3]Chi-Sq: 18.706; DF: 4; P-Value: 0.001; [4]Chi-Sq: 15.627; DF: 4; P-Value: 0.004; [5]Chi-Sq: 6.876; DF: 2; P-Value: 0.032.

5.3.2 Mental Disease

For mental disease, age, and socio-economic score have significant influence on patterns of GP visits (the p-value for age is 0.016; for socio-economic score the p-value is 0.040). Examination of patterns of GP visits for mental disease by age indicates a trend that the older are more likely to have *continuous episodic care* than the younger. Concerning the patterns of GP visits for mental disease by socio-economic score, there are also significant differences in the composition among each type of episodic care for GP visits in each socio-economic status category. Table 5.13 shows that the subjects with lower socio-economic status are more likely to have *continuous episodic care* than those with medium and higher socio-economic status. Most people are more likely to have *no episodes* for GP visits no matter what variable is investigated. The percentage of *no episodes* is from 63.5 to 81.7.

With regard to the patterns of specialist visits for mental disease, the five selected variables have no significant influence in this sample (Table 5.14).

Investigation of the association between distribution of number of hospitalizations and selected variables for mental disease, reveals that only URBAN, SES9, and HSTAT have significant influence (Table 5.15). The respondents who live in remote areas, have lower socio-economic scores, or have fair or poor health status have more hospitalizations than those respondents who live in St. John's, have higher socio-economic status, or have excellent or good health status. Sex and age have no significant influences on the number of hospitalizations.

Table 5.13 Patterns of GP visits for mental disease

by selected variables in four years (%)

	N	N of subj. with visits	No Episodes	Isolated Episodes	Continuous Episodes
GP Visits	402	237(59.0)	173(43.0)	50(12.4)	14(3.5)
Sex[1]					
Male	155	82(52.9)	64(78.1)	12(14.6)	6(7.3)
Female	247	155(62.8)	109(70.3)	38(24.5)	8(5.2)
Age[2]					
20-39	159	93(58.5)	76(81.7)	13(14.0)	4(4.3)
40-59	179	104(58.1)	66(63.5)	32(30.8)	6(5.8)
60+	64	40(62.5)	31(77.5)	5(12.5)	4(10.0)
Urban[3]					
St. John's	182	90(49.5)	66(73.3)	20(22.2)	4(4.4)
Other urban	98	65(66.3)	49(75.4)	13(20.0)	3(4.6)
Remote	122	82(67.2)	58(70.7)	17(20.7)	7(8.5)
SES3[4]					
Low	174	104(59.8)	83(79.8)	14(13.5)	7(6.7)
Middle	166	102(61.4)	65(63.7)	31(30.4)	6(5.9)
High	57	27(47.4)	22(81.5)	4(14.8)	1(3.7)
Health Status[5]					
Exc-Good	254	153(60.2)	117(76.5)	28(18.3)	8(5.3)
Fair-Poor	148	84(56.8)	56(66.7)	22(26.2)	6(7.1)

GP visits: [1]Chi-Sq: 3.343; DF: 2; P-Value: 0.188; [2]Chi-Sq: 12.167; DF: 4; P-Value: 0.016; [3]Chi-Sq: 1.675; DF: 4; P-Value: 0.795; [4]Chi-Sq: 10.041; DF: 4; P-Value: 0.040; [5]Chi-Sq: 2.650; DF: 2; P-Value: 0.266.

Table 5.14 Patterns of specialist visits for mental disease by selected variables in four years (%)

	N	N of subj. with visits	No Episodes	Isolated Episodes	Continuous Episodes
Specialist Visits	402	177(44.0)	64(15.9)	56(13.9)	57(14.2)
Sex[1]					
Male	155	77(49.7)	23(29.9)	26(33.8)	28(36.4)
Female	247	100(40.5)	41(41)	30(30)	29(29)
Age[2]					
20-39	159	67(42.1)	23(34.3)	21(31.3)	23(34.3)
40-59	179	88(49.2)	34(38.6)	24(27.3)	30(34.1)
60+	64	22(34.3)	7(31.8)	11(50)	4(18.2)
Urban[3]					
St. John's	182	108(59.3)	34(31.5)	34(31.5)	40(37.0)
Other urban	98	31(31.6)	11(35.5)	11(35.5)	9(29.0)
Remote	122	38(31.1)	19(50.0)	11(29.0)	8(21.1)
SES3[4]					
Low	174	72(41.4)	26(36.1)	24(33.3)	22(30.6)
Middle	166	68(41.0)	24(35.3)	25(36.8)	19(27.9)
High	57	36(63.2)	14(38.9)	6(16.7)	16(44.5)
Health Status[5]					
Exc-Good	254	107(42.1)	38(35.5)	33(30.8)	36(33.6)
Fair-Poor	148	70(47.3)	26(37.1)	23(32.9)	21(29.9)

SP visits: [1]Chi-Sq: 2.418; DF: 2; P-Value: 0.299; [2]Chi-Sq: 4.737; DF: 4; P-Value: 0.315; [3]Chi-Sq: 5.251; DF: 4; P-Value: 0.262; [4]Chi-Sq: 5.403; DF: 4; P-Value: 0.248; [5]Chi-Sq: 0.260; DF: 2; P-Value: 0.878.

Table 5.15 Distribution (and percentage) of hospitalizations for mental disease by selected variables in four years

	N	N of Hosp.	No Hosp.	One visit	Two or more
Total	402	57(14.2)	345(85.8)	37(9.2)	20(5.0)
Sex[1]					
Male	155	29(18.7)	126(81.3)	17(11.0)	12(7.7)
Female	247	28(11.3)	219(88.7)	20(8.1)	8(3.2)
Age[2]					
20-39	159	22(13.8)	137(86.2)	14(8.8)	8(5.0)
40-59	179	21(11.7)	158(88.3)	13(7.3)	8(4.5)
60+	64	14(21.9)	50(78.1)	10(15.6)	4(6.3)
Urban[3]					
St. John's	182	14(7.7)	168(92.3)	7(3.8)	7(3.8)
Other urban	98	18(18.4)	80(81.6)	13(13.3)	5(5.1)
Remote	122	25(20.5)	97(79.5)	17(13.9)	8(6.6)
SES3[4]					
Low	174	37(21.3)	137(78.7)	25(14.4)	12(6.9)
Middle	166	17(10.2)	149(89.8)	11(6.6)	6(3.6)
High	57	3(5.3)	54(94.7)	1(1.8)	2(3.5)
Health Status[5]					
Exc-good	254	31(12.2)	223(87.8)	25(9.8)	6(2.4)
Fair-poor	148	26(17.6)	122(82.4)	12(8.1)	14(9.5)

[1]Chi-Sq: 5.338; DF: 2; P-Value: 0.069; [2]Chi-Sq: 4.498; DF: 4; P-Value: 0.343; [3]Chi-Sq: 13.141; DF: 4; P-Value: 0.011; [4]Chi-Sq: 13.524; DF: 4; P-Value: 0.009; [5]Chi-Sq: 10.087; DF: 2; P-Value: 0.006.

5.3.3 Chronic respiratory conditions

The association between patterns of GP visits for chronic respiratory conditions and the five selected variables is also investigated. Age, socio-economic score, and health status have significant influence on patterns of GP visits (The p-value for age is 0.000; for socio-economic score is 0.015; for health status is 0.000). The examination of the patterns of GP visits by sex and urban does not show any significance. The influence of age on the patterns of GP visits shows a significant trend that the older are more likely to have *isolated* and *continuous episodic care* than the younger. The investigation of patterns of GP visits by socio-economic status also shows a significant trend that subjects with lower socio-economic status are more likely to have *continuous episodic care* and less likely to have *non-episodic care* than those with medium or higher socio-economic status (Table 5.16). Health status had significant influence on patterns of GP visits, in that subjects with fair or poor health status are much more likely to *have continuous episodic care* and less likely to have *non-episodic care* than those with excellent or good health status.

When patterns of specialist visits for chronic respiratory conditions is investigated by sex, age, socio-economic score, and self-assessed health status, no significant differences are found (Table 5.17). Urbanization is the only significant variable. With respect to patterns of specialist visits by urbanization, the respondents in St. John's are more likely to have *isolated* and *continuous episodic care* than those in other urban or remote areas.

Examination of the association between distribution of number of hospitalizations and the five selected variables for chronic respiratory conditions indicates that the five variables all have significant influence (Table 5.18). Respondents who are males, older, live in remote areas, have lower socio-economic scores, or have fair or poor health status have more numbers of hospitalizations than those respondents who are females, younger, live in St. John's or other urban, have higher socio-economic scores, or have excellent or good health status.

Table 5.16 Patterns of GP visits for chronic respiratory conditions by selected variables in four years (%)

	N	N of subj. with visits	No Episodes	Isolated Episodes	Continuous Episodes
GP Visits	942	888(94.3)	598(63.5)	229(24.3)	61(6.5)
Sex[1]					
Male	396	377(95.2)	260(69.0)	87(23.1)	30(8.0)
Female	546	511(93.6)	338(66.1)	142(27.8)	31(6.1)
Age[2]					
20-39	352	336(95.5)	244(72.6)	82(24.4)	10(3.0)
40-59	330	315(95.5)	215(68.3)	81(25.7)	19(6.0)
60+	260	237(91.2)	139(58.7)	66(27.9)	32(13.5)
Urban[3]					
St. John's	367	355(96.7)	239(67.3)	95(26.8)	21(5.9)
Other urban	312	292(93.6)	205(70.2)	67(22.9)	20(6.8)
Remote	263	241(91.6)	154(63.9)	67(27.8)	20(8.3)
SES3[4]					
Low	401	372(92.8)	242(65.1)	94(25.3)	36(9.7)
Middle	384	368(95.8)	251(68.2)	97(26.4)	20(5.4)
High	141	133(94.3)	97(72.9)	34(25.6)	2(1.5)
Health Status[5]					
Exc-Good	649	617(95.1)	436(70.7)	160(25.9)	21(3.4)
Fair-Poor	293	271(92.5)	162(59.8)	69(25.5)	40(14.8)

GP visits: [1]Chi-Sq: 3.253; DF: 2; P-Value: 0.197; [2]Chi-Sq: 27.641; DF: 4; P-Value: 0.000; [3]Chi-Sq: 3.391; DF: 4; P-Value: 0.495; [4]Chi-Sq: 12.276; DF: 4; P-Value: 0.015; [5]Chi-Sq: 38.682; DF: 2; P-Value: 0.000.

Table 5.17 Patterns of specialist visits for chronic respiratory conditions by selected variables in four years (%)

	No. of All	No. of subjects with visits	No Episodes	Isolated Episodes	Continuous Episodes
Specialist Visits	942	103(10.9)	80(8.5)	15(1.6)	8(0.85)
Sex[1]					
Male	396	36(9.1)	30(83.3)	3(8.3)	3(8.3)
Female	546	67(12.3)	50(74.6)	12(17.9)	5(7.5)
Age[2]					
20-39	352	32(9.1)	25(78.1)	7(21.9)	-
40-59	330	36(10.9)	28(77.8)	3(8.3)	5(13.9)
60+	260	35(13.5)	27(77.1)	5(14.3)	3(8.6)
Urban[3]					
St. John's	367	39(10.6)	24(61.5)	10(25.6)	5(12.8)
Other urban	312	33(10.6)	29(87.9)	2(6.1)	2(6.1)
Remote	263	31(11.8)	27(87.1)	3(9.7)	1(3.2)
SES[4]					
Low	401	39(9.7)	33(84.6)	4(10.3)	2(5.1)
Middle	384	40(10.4)	30(75)	8(20)	2(5)
High	141	20(14.2)	15(75)	3(15)	2(10)
Health Status[5]					
Exc-Good	649	57(8.8)	48(84.2)	6(10.5)	3(5.3)
Fair-Poor	293	46(15.7)	32(69.6)	9(19.6)	5(10.9)

SP visits: [1]Chi-Sq: 1.726; DF: 2; P-Value: 0.422; [2]Chi-Sq: 6.390; DF: 4; P-Value: 0.172; [3]Chi-Sq: 9.766; DF: 4; P-Value: 0.045; [4]Chi-Sq: 2.157; DF: 4; P-Value: 0.707; [5]Chi-Sq: 3.161; DF: 2; P-Value: 0.206.

Table 5.18 Distribution (and percentage) of hospitalizations for chronic respiratory conditions by selected variables in 4 years

	N	N of Hosp.	No Hosp.	One visit	Two or more
Total	942	57(6.1)	885(93.9)	39(4.1)	18(1.9)
Sex[1]					
Male	396	21(5.3)	375(94.7)	10(2.5)	11(2.8)
Female	546	36(6.6)	510(93.4)	29(5.3)	7(1.3)
Age[2]					
20-39	352	9(2.6)	343(97.4)	8(2.3)	1(0.3)
40-59	330	13(3.9)	317(96.1)	7(2.1)	6(1.8)
60+	260	35(13.5)	225(86.5)	24(9.2)	11(4.2)
Urban[3]					
St. John's	367	11(3.0)	356(97.0)	8(2.2)	3(0.8)
Other urban	312	19(6.1)	293(93.9)	12(3.8)	7(2.2)
Remote	263	27(10.3)	236(89.7)	19(7.2)	8(3.0)
SES3[4]					
Low	401	38(9.5)	363(90.5)	24(6.0)	14(3.5)
Middle	384	15(3.9)	369(96.1)	13(3.4)	2(0.5)
High	141	4(2.8)	137(97.2)	2(1.4)	2(1.4)
Health Status[5]					
Exc-good	649	20(3.1)	629(96.9)	13(2.0)	7(1.1)
Fair-poor	293	37(12.6)	256(87.4)	26(8.9)	11(3.8)

[1]Chi-Sq: 7.031; DF: 2; P-Value: 0.030; [2]Chi-Sq: 36.835; DF: 4; P-Value: 0.000; [3]Chi-Sq: 14.613; DF: 4; P-Value: 0.006; [4]Chi-Sq: 16.213; DF: 4;P-Value: 0.003; [5]Chi-Sq: 32.538; DF: 2; P-Value: 0.000.

5.4 Results of multiple analysis

The following figures show a summary of the results of logistic regression models, including binary and nominal models for GP visits, specialist visits, and hospitalizations in the three diagnostic groups. Socio-economic score is divided into nine levels in the following analysis.

5.4.1 Logistic models for the subjects with cardiovascular disease

The following nominal and binary logistic models investigate whether GP visits, specialist visits, and hospitalizations for subjects with cardiovascular disease is associated with any of SEX, AGE, URBAN, SES9, or HSTAT, using all subjects with medical care utilization for only cardiovascular disease in the four-year study period. These models show that AGE and SES9 is significant for patterns of GP visits. URBAN is the only significant variable for patterns of specialist visits. HSTAT, SEX, AGE, and URBAN are significant for hospitalizations. AGE is continuous in the following models of this section. Non-significant variables were removed from the model before producing the following figures.

In the following models, patterns 0, 1, and 2 represent respectively ***non-episodes***, *isolated episodes* and *continuous episodes*.

Figure 5.1 presents the estimated differences of pattern 2 (*continuous episodes*) and pattern 1 (*isolated episodes*) for GP visits, respectively, compared to pattern 0 (*non-episodes*) as the reference event. In this model, only the changes of socio-economic score and AGE have significant influence on patterns of GP visits for cardiovascular disease.

The positive coefficients for AGE indicate that subjects are more likely to have *continuous episodes* (in Logit (1) part) or *isolated episodes* (in Logit (2) part) as GP visits as they are older. The negative coefficients of socio-economic score indicate that subjects with lower socio-economic score are more likely to have *continuous* or *isolated episodes* as GP visits over *non-episodic care* compared to subjects with higher socio-economic scores. Although AGE is significant, the odds ratio is very close to one, indicating that a one year increase in age minimally effects patterns of GP visits for cardiovascular disease. A more meaningful difference is found if subjects with a larger age difference are compared. For socio-economic scores, the negative coefficient and the odds ratio indicate that subjects with lower socio-economic scores are more likely to have *isolated* or *continuous episodic care* than subjects with higher socio-economic scores.

The results of Figure 5.1 are also presented in Table 5.19. From this table, a clear trend is indicated that respondents who are older, or have lower socio-economic scores are more likely to have *continuous episodic care* than to have *non-episodic care*. For example, respondents who are over 70 years old are over five times more likely to have *continuous episodic care* than those respondents who are under 30 years old (0.22 versus 0.04). Respondents who have lower socio-economic scores have about three times more the probability of having *continuous episodes* than those respondents who have higher socio-economic scores (0.21 versus 0.07).

Response Information

Variable	Value	Count
pattern	0	315 (Reference Event): No episodes
	2	90: Continuous episodes
	1	141: Isolated episodes
	Total	546

546 cases were used

12 cases contained missing values

Logistic Regression Table

					Odds	95% CI	
Predicator	Coef	StDev	Z	P	Ratio	Lower	Upper
Logit 1:	(2/0)						
Constant	-2.5979	0.8408	-3.09	0.002			
Sescore	-0.19232	0.07733	-2.49	0.013	0.83	0.71	0.96
Age	0.02999	0.01061	2.83	0.005	1.03	1.01	1.05
Logit 2:	(1/0)						
Constant	-1.9986	0.6664	-3.00	0.003			
Sescore	-0.02806	0.05607	-0.50	0.617	0.97	0.87	1.09
Age	0.019971	0.008464	2.36	0.018	1.02	1.00	1.04

Log-likelihood = -512.951

Test that all slopes are zero: G = 26.921, DF = 4, P-Value = 0.000

Goodness-of-Fit Tests

	Chi-Square	DF	P
Method			
Pearson	1064.188	1070	0.544
Deviance	1013.766	1070	0.889

Figure 5.1 Nominal logistic model of patterns of GP visits for subjects with cardiovascular disease

Table 5.19 Mean event probability of GP visits for subjects with cardiovascular disease by the significant variables

Variables	Non-episodic	Isolated episodes	Continuous episodes
Age			
20-29	0.80	0.16	0.04
30-39	0.75	0.18	0.07
40-49	0.71	0.21	0.08
50-59	0.65	0.23	0.12
60-69	0.58	0.26	0.16
70+	0.49	0.29	0.22
Socio-economic Score			
Low	0.53	0.27	0.21
Middle	0.63	0.25	0.12
High	0.70	0.24	0.07

Figure 5.2 shows that only the change in area of residence has significant influence on the pattern of specialist visits for cardiovascular disease. For urban, the negative coefficient and the odds ratio indicate that subjects living in St. John's are more likely to have *continuous episodic care* than subjects living in other urban and remote areas.

The results of Figure 5.2 are also presented in Table 5.20. From this table, a clear trend is indicated that respondents who live in St. John's are two and half times more likely to have *continuous episodic care* for specialist visits than those respondents who live in remote areas (0.15 versus 0.06).

Response Information

Variable	Value	Count	
pattern	0	172	(Reference Event): No episodes
	2	21	: Continuous episodes
	1	69	: Isolated episodes
	Total	262	

262 cases were used

6 cases contained missing values

Logistic Regression Table

					Odds	95% CI	
Predicator	Coef	StDev	Z	P	Ratio	Lower	Upper
Logit 1: (2/0)							
Constant	-1.4214	0.2978	-4.77	0.000			
Urban							
Other urban	-2.450	1.053	-2.33	0.020	0.09	0.01	0.68
Remote	-0.9765	0.5201	-1.88	0.060	0.38	0.14	1.04
Logit 2: (1/0)							
Constant	-0.9249	0.2464	-3.75	0.000			
Urban							
Other urban	0.2318	0.3510	0.66	0.509	1.26	0.63	2.51
Remote	-0.1737	0.3483	-0.50	0.618	0.84	0.42	1.66

Log-likelihood = -210.961

Test that all slopes are zero: G = 12.976, DF = 4, P-Value = 0.011

* NOTE * No goodness-of-fit tests performed.

* The model uses all degrees of freedom.

Figure 5.2. Nominal logistic model of patterns of specialist visits for subjects with cardiovascular disease

Table 5.20 Mean event probability of specialist visits for subjects with cardiovascular disease by the significant variables

Variables	Non-episodic	Isolated episodes	Continuous episodes
Urban			
St. John's	0.61	0.24	0.15
Other urban	0.66	0.33	0.01
Remote	0.70	0.23	0.06

Examination of coefficients in Figure 5.3 reveals that HSTAT, SEX, AGE, and URBAN are significant. A more meaningful difference is found upon examination of odds ratios. The odds ratio for health status (1.26) indicates that when the subject's health status decreases one category from good to poor, the odds of a subject being hospitalized for cardiovascular disease increase by 1.26 times. The negative coefficient and odds ratio for SEX indicates that the odds of being hospitalized is 1.47 (1/0.68) times greater for males compared to females. The odds ratio for URBAN shows that the odds of a subject who lives in other urban and being hospitalized is 1.20 times (1.93 times for those living in remote areas) greater compared to those who live in St. John's. AGE has minimal influence on hospitalizations for cardiovascular disease.

The results of Figure 5.3 are also presented in Table 21. From this table, it can be seen that there is a trend that respondents who are older, males, have poorer health status, or live in remote areas are more likely to be hospitalized than those who are younger, females, have better health status, or live in St. John's and other urban. For example, respondents who have poor health status are two times more likely to be hospitalized than those respondents who have excellent health status (0.38 versus 0.19).

Link Function: Logit

Response Information

Variable	Value	Count
hosp	1	185 (Event): Hospitalized
	0	480: Not hospitalized
	Total	665

665 cases were used

13 cases contained missing values

Logistic Regression Table

Predicator	Coef	StDev	Z	P	Odds Ratio	95% CI Lower	95% CI Upper
Constant	-2.3166	0.5625	-4.12	0.000			
Hstat	0.2301	0.1135	2.03	0.043	1.26	1.01	1.57
Sex	-0.3892	0.1818	-2.14	0.032	0.68	0.47	0.97
Age	0.016019	0.007021	2.28	0.023	1.02	1.00	1.03
Urban							
Other urban	0.1840	0.2342	0.79	0.432	1.20	0.76	1.90
Remote	0.6576	0.2219	2.96	0.003	1.93	1.25	2.98

Log-Likelihood = -378.665

Test that all slopes are zero: G = 29.021, DF = 5, P-Value = 0.000

Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	651.559	656	0.542
Deviance	751.785	656	0.005
Hosmer-Lemeshow	8.710	8	0.367

Figure 5.3. Binary logistic model of hospitalizations for subjects with cardiovascular disease

Table 5.21 Mean event probability of hospitalizations for subjects with cardiovascular disease by the significant variables

Variables	Not Hospitalized	Hospitalized
Sex		
Male	0.69	0.31
Female	0.76	0.24
Age		
20-29	0.87	0.13
30-39	0.83	0.17
40-49	0.79	0.21
50-59	0.75	0.25
60-69	0.71	0.29
70+	0.68	0.32
Urban		
St. John's	0.79	0.21
Other urban	0.75	0.25
Remote	0.63	0.37
Health status		
Excellent	0.81	0.19
Good	0.75	0.25
Fair	0.68	0.32
Poor	0.62	0.38

5.4.2 Logistic model for subjects with mental disease

The following part investigates whether GP visits, specialist visits, and hospitalizations for subjects with mental disease are associated with the variables: SEX, AGE, URBAN, SES, or HSTAT, using subjects with medical care utilization only for mental disease in the four-year study period. Concerning the patterns of GP or specialist visits for mental disease, there are no suitable models to fit them using the selected five independent variables. The model of hospitalizations showed that SEX, URBAN, and SES9 were significant.

Examination of the estimated coefficients of Figure 5.4 reveals that socio-economic score, sex, and urban have p-value less than 0.05, indicating that there is sufficient evidence to conclude that the parameters are not zero using a significance level of $\alpha=0.05$. The negative coefficient and odds ratio of socio-economic score indicate that subjects with lower socio-economic scores are more likely to be hospitalized for mental disease than those with higher socio-economic scores. The negative coefficient for SEX indicates that males are more likely to be hospitalized than females. URBAN also has significant influence on hospitalizations. Subjects living in remote areas are much more likely to be hospitalized for mental disease than those living in St. John's.

The results of Figure 5.4 can be clearly shown in Table 22. From Table 22, It can be seen that the probabilities of being hospitalized are higher among those respondents who are males, live in remote areas, or have lower socio-economic scores.

Link Function: Logit

Response Information

Variable	Value	Count
hosp	1	57 (Event): Hospitalized
	0	340: Not hospitalized
	Total	397

397 cases were used

5 cases contained missing values

Logistic Regression Table

Predicator	Coef	StDev	Z	P	Odds Ratio	95% CI Lower	95% CI Upper
Constant	-0.4287	0.6293	-0.68	0.496			
Sescore	-0.25526	0.08791	-2.90	0.004	0.77	0.65	0.92
Sex	-0.6011	0.2969	-2.02	0.043	0.55	0.31	0.98
Urban							
Other urban	0.7545	0.3941	1.91	0.056	2.13	0.98	4.60
Remote	0.9097	0.3693	2.46	0.014	2.48	1.20	5.12

Log-Likelihood = -150.355

Test that all slopes are zero: G = 25.945, DF = 4, P-Value = 0.000

Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	38.575	45	0.739
Deviance	42.552	45	0.576
Hosmer-Lemeshow	10.940	8	0.205

Figure 5.4. Binary logistic model of hospitalizations for subjects with mental disease

Table 5.22 Mean event probability of hospitalizations for subjects with mental disease by the significant variables

Variables	Not Hospitalized	Hospitalized
Sex		
Male	0.83	0.17
Female	0.89	0.11
Urban		
St. John's	0.92	0.08
Other urban	0.84	0.16
Remote	0.80	0.20
Socio-economic score		
Low	0.77	0.23
Middle	0.88	0.12
High	0.94	0.06

5.4.3 Logistic models for subjects with chronic respiratory conditions

The following nominal and binary logistic models investigate whether GP visits, specialist visits, and hospitalizations for subjects with chronic respiratory conditions are associated with the variables: SEX, AGE, URBAN, SES, or HSTAT, using subjects with medical care utilization only for chronic respiratory conditions in the four-year study period. These models showed that AGE and health status are significant for patterns of GP visits. Urbanization and health status are significant for patterns of specialist visits. Health status, SEX, AGE, and URBAN are significant for hospitalizations. Age is continuous in the following models of this section.

Figure 5.5 shows the estimated differences of pattern 2 (*continuous episodes*) and pattern 1 (*isolated episodes*) for GP visits, respectively, compared to pattern 0 (*non-episodes*) as the reference event. In this model, only the change in variables health status and AGE have significant influence on patterns of GP visits for chronic respiratory conditions. The positive coefficients for AGE indicate that subjects are more likely to have *continuous episodes* (pattern 2) of GP visits over the *non-episodic care* (pattern 0) as age increases. The positive coefficients of health status indicate that subjects with poorer health status are more likely to have *continuous episodes* (pattern 2) of GP visits over the *non-episodic care* (pattern 0) compared to subjects with better health status. Although there is evidence that the parameter of AGE is not zero, the odds ratio is very close to one, indicating that a one year increase in age minimally affects pattern of GP visits for chronic respiratory conditions. A more meaningful difference is found if subjects with a larger age difference are compared.

The results of Figure 5.5 are also presented in Table 5.23. From this table, a trend can be found that respondents who have poor health status are 27 times more likely to have *continuous episodic care* of GP visits for chronic respiratory conditions than those respondents who have excellent health status (0.27 versus 0.01). Respondents who are older have a higher probability of *isolated* or *continuous episodes* than those respondents who are young (0.28 versus 0.23 for *isolated episodes*; 0.15 versus 0.02 for *continuous episodes*).

Response Information

Variable	Value	Count
pattern	0	590 (Reference Event): No episodes
	2	58: Continuous episodes
	1	225: Isolated episodes
	Total	873

873 cases were used
15 cases contained missing values

Logistic Regression Table

Predicator	Coef	StDev	Z	P	Odds Ratio	95% CI Lower	95% CI Upper
Logit 1:	(2/0)						
Constant	-6.5660	0.6675	-9.84	0.000			
Hstat	0.9605	0.1817	5.29	0.000	2.61	1.83	3.73
Age	0.035410	0.008778	4.03	0.000	1.04	1.02	1.05
Logit 2:	(1/0)						
Constant	-1.5432	0.3058	-5.05	0.000			
Hstat	0.0908	0.1070	0.85	0.396	1.10	0.89	1.35
Age	0.008042	0.004694	1.71	0.087	1.01	1.00	1.02

Log-likelihood = -664.018
Test that all slopes are zero: G = 58.962, DF = 4, P-Value = 0.000

Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	1642.489	1708	0.870
Deviance	1308.628	1708	1.000

Figure 5.5 Nominal logistic model of patterns of GP visits

for the subjects with chronic respiratory conditions

Table 5.23 Mean event probability of GP visits for subjects with chronic respiratory conditions by the significant variables

Variables	Non-episodic	Isolated episodes	Continuous episodes
Age			
20-29	0.75	0.23	0.02
30-39	0.72	0.25	0.03
40-49	0.69	0.26	0.05
50-59	0.65	0.27	0.08
60-69	0.62	0.27	0.11
70+	0.57	0.28	0.15
Health status			
Excellent	0.74	0.25	0.01
Good	0.70	0.26	0.04
Fair	0.63	0.26	0.11
Poor	0.49	0.24	0.27

Figure 5.6 shows the estimated differences of pattern 2 (*continuous episodes*) and pattern 1 (*isolated episodes*) for specialist visits, respectively, compared to pattern 0 (*non-episodes*) as the reference event. The first set: logit (1) is the parameter estimates of the change of *continuous episodes* (pattern 2) relative to the reference event, *non-episodes* (pattern 0). In this model, the p-values of HSTAT and URBAN are more than the acceptable α level, indicating that there is insufficient evidence to conclude that a change in HSTAT and URBAN affect *continuous episodes* for specialist visits over the *non-episodic care*. The second set: Logit(2) is the parameter estimates of the change of

isolated episodes (pattern 1) relative to the reference event, *non-episodes*. Both the p-values for HSTAT and URBAN indicate that there is sufficient evidence to conclude that a change in HSTAT or URBAN affected *isolated episodes* for specialist visits compared to the *non-episodic care*. The positive coefficients and odds ratio for HSTAT indicate that subjects with poorer health status are more likely to have *isolated episodes* for specialist visits. The negative coefficients and odds ratio for URBAN indicate that subjects living in St. John's are more likely to have *isolated episodes* for specialist visits over the *non-episodic care* compared to subjects living in other urban or remote areas.

The results of Figure 5.6 can be simply described by the mean of event probability in Table 24. When the health status of respondents decreased from excellent to poor, the probability of having *non-episodic care* decreased over one fold; the probability of having *isolated* or *continuous episodes* increased, over six-fold for *isolated episodes*, over one-fold for *continuous episodes*. Respondents who live in St. John's have a higher probability of *isolated* or *continuous episodes* than those respondents who live in remote areas.

Response Information

Variable	Value	Count
pattern	0	78 (Reference Event): No episodes
	2	6: Continuous episodes
	1	15: Isolated episodes
	Total	99

99 cases were used

4 cases contained missing values

Logistic Regression Table

Predicator	Coef	StDev	Z	P	Odds Ratio	95% CI Lower	Upper
Logit 1: (2/0)							
Constant	-2.691	1.355	-1.99	0.047			
Hstat	0.3023	0.5446	0.56	0.579	1.35	0.47	3.93
Urban							
Other urban	-0.7709	0.9919	-0.78	0.437	0.46	0.07	3.23
Remote	-1.398	1.241	-1.13	0.260	0.25	0.02	2.81
Logit 2: (1/0)							
Constant	-2.711	1.018	-2.66	0.008			
Hstat	0.8197	0.3995	2.05	0.040	2.27	1.04	4.97
Urban							
Other urban	-2.2526	0.8825	-2.55	0.011	0.11	0.02	0.59
Remote	-1.8666	0.8092	-2.31	0.021	0.15	0.03	0.76

Log-likelihood = -57.239

Test that all slopes are zero: G = 12.967, DF = 6, P-Value = 0.044

Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	9.216	16	0.904
Deviance	11.217	16	0.796

Figure 5.6 Nominal logistic model of patterns of specialist visits for the subjects with chronic respiratory conditions

Table 5.24 Mean event probability of specialist visits for subjects with chronic respiratory conditions by the significant variables

Variables	Non-episodic	Isolated episodes	Continuous episodes
Urban			
St. John's	0.58	0.34	0.08
Other urban	0.87	0.07	0.06
Remote	0.87	0.10	0.03
Health status			
Excellent	0.90	0.05	0.05
Good	0.84	0.11	0.06
Fair	0.75	0.19	0.06
Poor	0.62	0.31	0.07

When the estimated coefficients of Figure 5.7 are examined, it can be seen that HSTAT, SEX, AGE, and URBAN are significant predictors of hospitalizations. The positive coefficient and odds ratio of health status indicate that subjects with poorer health status are more likely to be hospitalized for chronic respiratory conditions than those with better health status. The positive coefficient for SEX indicates that females are more likely to be hospitalized than males. The positive coefficient for AGE indicates that the older are more likely to be hospitalized than the younger. URBAN also has significant influence on hospitalizations. Subjects living in the remote areas are much more likely to be hospitalized for chronic respiratory conditions than those living in St. John's.

The results of Figure 5.7 can also be described by the mean of event probability in

Table 5.25. When the health status of respondents decreases from excellent to poor, the probability of being hospitalized increases about 26 times (from 0.01 to 0.26). Respondents who live in remote areas have a higher probability for hospitalization than those respondents who live in St. John's (0.11 versus 0.03).

Link Function: Logit

Response Information

Variable	Value	Count
hosp	1	57 (Event): Hospitalized
	0	869: Not hospitalized
	Total	926

926 cases were used

16 cases contained missing values

Logistic Regression Table

Predicator	Coef	StDev	Z	P	Odds Ratio	95% CI Lower	Upper
Constant	-8.9489	0.9855	-9.08	0.000			
Hstat	0.9546	0.1869	5.11	0.000	2.60	1.80	3.75
Sex	0.6064	0.3069	1.98	0.048	1.83	1.00	3.35
Age	0.042611	0.009269	4.60	0.000	1.04	1.02	1.06
Urban							
Other urban	0.5010	0.4050	1.24	0.216	1.65	0.75	3.65
Remote	0.9098	0.3878	2.35	0.019	2.48	1.16	5.31

Log-Likelihood = -174.735

Test that all slopes are zero: G = 78.759, DF = 5, P-Value = 0.000

Goodness-of-Fit Tests

Method	Chi-Suare	DF	P
Pearson	1080.468	915	0.000
Deviance	349.470	915	1.000
Hosmer-Lemeshow	3.639	8	0.888

Figure 5.7 Binary logistic model of hospitalizations for the subjects with chronic respiratory conditions

Table 5.25 Mean event probability of hospitalizations for subjects with chronic respiratory conditions by the significant variables

Variables	Not Hospitalized	Hospitalized
Sex		
Male	0.95	0.05
Female	0.93	0.07
Age		
20-29	0.99	0.01
30-39	0.98	0.02
40-49	0.96	0.04
50-59	0.93	0.07
60-69	0.89	0.11
70+	0.84	0.16
Urban		
St. John's	0.97	0.03
Other urban	0.94	0.06
Remote	0.89	0.11
Health status		
Excellent	0.99	0.01
Good	0.97	0.03
Fair	0.90	0.10
Poor	0.74	0.26

Chapter 6. Conclusions

This study has been successful in achieving most of the objectives originally included in the proposal. The association between patterns of medical care utilization, including GP visits, specialist visits, and the probability of being hospitalized, and social and demographic variables, including sex, age, health status, socio-economic score, and urbanization, has been verified for three diagnostic groups, as well as for all subjects with utilization for any diagnosis in the four-year study period. Logistic regression is used to demonstrate the relationship between the five selected variables and the patterns of GP or specialist visits and the probability of being hospitalized.

Using temporal continuity and diagnostic continuity, five patterns of medical care utilization for GP and specialist visits have been identified. There are also some observations about patterns of care for the three diagnostic groups from the analysis. Chronic respiratory conditions are very common in all age groups; people with this disease are more likely to visit their GP for medical care rather than specialists or hospitals. By contrast, mental disease and cardiovascular disease may need more medical care; the subjects with mental disease or cardiovascular disease are more likely to visit specialists or hospitals. With respect to the mean number of GP and specialist visits, the number of hospitalizations, as well as the mean of length of stay (LOS) by urbanization for these three diagnoses, there is a clear difference in the mix of GPs, specialists, and hospitalizations by diagnostic groups. Generally subjects with cardiovascular disease have more visits to GPs and hospitalizations than those with mental disease and chronic respiratory conditions. Subjects with mental disease have more specialist visits and longer LOS than those subjects with either of the other two diagnoses. Subjects with chronic respiratory conditions have the least number of specialist visits and hospitalizations compared to the other two diagnostic groups. Comparisons of patterns of medical care utilization within the three major categories of services: GP visits, specialist visits, and hospitalizations by the five demographic socio-economic and health characteristics for three diagnostic groups show both differences and similarities.

Patterns of GP visits - The associations between patterns of GP visits and the five selected variables show both similarities and differences from the descriptive analysis for the three diagnostic groups. Across all three diagnostic groups considered, there are no obvious differences for the patterns of GP visits by urbanization. No matter where the study subjects live, they have the similar patterns of GP visits for the three kinds of diagnoses. A lower number of GP visits in the remote areas are expected due to the absence of data from visits to salary physicians. The association between patterns of GP visits and sex shows similar results for the three diagnostic groups, where there is no significant difference between males and females in terms of patterns of GP visits. The investigation of patterns of GP visits by socio-economic status shows that study subjects with lower socio-economic status are much more likely to have continuous episodic care and less likely to have non-episodic care than those with medium or higher socio-economic status for the three diagnostic groups. Health status of study subjects only has significant influence on the patterns of GP visits for cardiovascular disease and chronic respiratory conditions. Respondents with fair or poor health status are more likely to have continuous episodic care and less likely to have non-episodic care than those with excellent or good health status. In all diagnostic groups, there is very little difference in the patterns of utilization for GPs. Most subjects in the three diagnostic groups have non-episodic care, from 52.2% up to 79.8%, no matter which variable is considered. Access has no obvious influence on the patterns of GP visits for the three diagnostic groups. The multivariate analysis further examines the complex association between patterns of GP visits and demographic and socio-economic variables. With respect to cardiovascular disease, variable socio-economic score and age have significant influence on patterns of GP visits. For chronic respiratory conditions, health status and age are significant. Age is the common significant variable for patterns for these two diagnostic groups. Concerning mental disease, there were no models that proved satisfactory. GP visits for mental disease may be influenced by factors not considered in this study.

Patterns of specialist visits – The associations between patterns of specialist visits and the selected five variables are similar for cardiovascular disease and chronic respiratory conditions. No matter what analysis methods are used, descriptive or multivariate, no variables are significant for patterns of specialist visits for mental disease. Across the two diagnostic groups: cardiovascular disease and chronic respiratory conditions, urbanization is the only common and significant variable for patterns of specialist visits from the descriptive analysis. The subjects in St. John's are much more likely to have continuous episodic and less likely to have non-episodic care of specialist visits than those in other urban or the remote. The multivariate analysis further clarifies this relationship for the two diagnostic groups in that there is a clear gradient, with more continuity in St. John's, and less continuity in the remote areas. Despite this, health status has also significant influence on patterns of specialist visits for chronic respiratory conditions, particularly influence on the isolated episodes over the non-episodic care. As expected, access has some influence on continuity of care for specialists. For chronic respiratory conditions, most respondents have non-episodic care for patterns of specialist visits, and non-episodic care accounted for more than 60%, even up to 87.9%, among all types of episodic care for specialist visits.

Hospitalizations – In this study, hospitalizations are analyzed in two ways. Descriptive analysis is used on one hand to identify the distribution and percentage of hospitalizations. The multivariate analysis is used on the other hand to examine the association between the probability of being hospitalized and demographic and socio-economic variables. Across the three diagnostic groups, respondents who live in the remote areas have more single hospitalizations, and two or more hospitalizations than those who live in St. John's and other urban for the three diagnostic groups. The multivariate analysis shows some similarities across the three diagnostic groups for the probability of being hospitalized. The common finding is that cardiovascular disease and chronic respiratory conditions have the same significant variables: health status, age, sex, and urbanization. The odds of being hospitalized for the subjects who are older, females, have poorer health status, or live in the rest, is more than for those subjects who are younger, males, have good health status, or live in other areas, for these two diagnoses separately. Urbanization and sex are also the common significant variables for mental disease. Females and the subjects who live in remote areas are more likely to be hospitalized. Socio-economic score is a significant variable for mental disease where the subjects with lower socio-economic scores are much more likely to be hospitalized for mental disease.

Descriptive and multivariate analyses clarify the complex association among demographic geographic, and socio-economic variables, as well as residence. It appears from the analysis that even in a society with universal medical coverage there is an increase in the number of continuous episodes for specialist visits in St. John's for the three diagnostic groups and an increase in the continuous episodes for GP visits in the urban areas for all subjects. The subjects in the remote areas have a higher probability of being hospitalized for the three diagnostic groups. Even when economic barriers are removed by health insurance, differences related to SES and access persist. For all subjects, the continuous episodes increase in the lower socio-economic groups. As expected, access has significant influence on patterns of specialist visits.

In Newfoundland, the regions were remarkably divergent in physician supply. Newfoundland is a geographically large province with most of its population and its physicians living in St. John's. The rest of the population is thinly distributed in other urban and the remote areas. Inner city residents have, on average, higher physician contact rates than other Newfoundland residents. The access to ambulatory care and hospitalizations for the three diagnostic groups was not generally clear before this study. In this study, access was measured by structured access using the three access areas. It appears that there is a "substitution" factor, in which insufficient visits to specialists may be replaced by hospitalizations. For three areas used in this study, it is known that specialists are in short supply in the rural areas, and geographical location is also a barrier to use. Rural residents may have to travel long distances to obtain services and even farther to find non-medical health services. Thus, the residents in St. John's are more likely to have continuous episodes for specialist visits.

This study raises legitimate questions about the appropriate allocation of physician resources in a province with widely varied population density and different needs. What is an appropriate level and mix of GP and specialist supply? What are the differences of patterns of GP or specialist visits for the three diagnostic groups between urban and rural residents? This study improves our understanding of the influence of access, demographic and socio--economic variable, and health status on patterns of medical care utilization, measured as dynamic phenomena over time.

One of the most important outcomes of the study is methodological. It has been possible to design computer programs to study patterns of medical care utilization using the concept of episode, and to obtain more detailed information about GP visits, specialist visits, and hospitalizations for three diagnostic groups. A certain predictable pattern of medical care utilization has been verfied for the three diagnostic groups using the five access, health status, demographic, and socio-economic variables. The methods used in this study open the way to further and more refined studies for medical care utilization, and advance the understanding of differences in patterns of medical care utilization by providing more accurate profiles of the population in the context of multiple service use—hospitals, specialist visits, and GP visits based on the diagnoses: cardiovascular disease, mental disease, and chronic respiratory conditions.

The other important outcome is that a longitudinal study permits a better time sequence between the variables under study, and extends the medical care utilization experience of the subjects. The longitudinal study of medical care utilization provides an enriched data base with a larger absolute number of hospital users and greater variation in amount of utilization. The study of overall utilization and the study of special sub-groups is greatly enhanced by studying several years of data.

Limitations - A number of limitations must bring caution to the interpretation of the results. First, with respect to mental disease, there is no suitable model that fits patterns of GP or specialist visits. One reason for this situation is that there are few subjects in this study for mental disease so that there are more cells with less than five, even less than one case, when the subgroup of GP visits or specialist visits was further divided by the pattern variable; so there is a small sample size for mental disease. This introduces some difficulties in the analysis, and reduces the possibility of finding significant differences for patterns of medical care utilization among different variables. Secondly, missing values exist for some subjects concerning socio-economic score. Thirdly, the three diagnostic groups were selected by primary diagnoses given by GP, specialist, or hospital. Therefore, there is probably some bias in the diagnosis selection. Fourthly, health status and socio-economic score are two confounding variables for the analysis of patterns. The sample size in this study is not big enough for the three diagnostic groups to be used to study patterns in the same health status or socio-economic score level.

Future Research - The methodology in this study might be used in future research. The medical care utilization data of a longer time period or a larger population may be used in order to get a larger sample size for these three diagnostic groups, or other diagnostic groups in order to study patterns of medical care utilization in the same way. Another approach might be to examine patterns of medical care utilization across regions whose populations are of similar health status or in the same socio-economic groups. Further, the analysis of the number of physicians consulted and the number of locations visited within each type of pattern and within each utilization subgroup may be included, which will be very useful in the practical policy making.

This study is the first step in developing programs to be able to analyze patterns of care, going beyond the simple consideration of number of visits. The island portion of the province may be considered a natural community laboratory to study the effects of different levels of medical care resources on health status and medical care utilization. The results of this study provide useful information to understand patterns of care for the three diagnostic groups, as well as some information of all subjects with utilization in the four-year study period.

References

Aday LA, Andersen R and Fleming GV. (1980). *Health care in the US: Equitable for whom?* Beverly Hills: Sage publication.

Adler NE, Boyce WT, Chesney MA, Folkman S and Syme SL. (1993). Socio-economic inequalities in health: no easy solution. *Journal of the American Medical Association*, 269, 3140-3145.

Andersen R and Aday LA. (1978) Access to medical care in the U.S.: realized and potential. *Medical Care*, 16, 533-546.

Anderson JG. (1973). Health Services Utilization: Framework and Review. *Health Services Research*, 8(3), 184-199.

Beck RG and Horne JM. (1980). Utilization of Publicly Insured Health Services in Saskatchewan Before, During and After Co-payment. *Medical Care*, 18(8), 787-806.

Beck RG. (1973). Economic Class and Access to Physician Services Under Public Medical Care Insurance. *International Journal of Health Services*, 3(3), 341-355.

Beck RG. (1974). The Effects of Co-payment on the poor. *Journal of Human Resources*, 9(1), 129-143.

Bergner M. (1985). Measurement of health status. *Medical Care*, 23(5), 696-704.

Bergner M. (1987). Health status measures: An overview and guide for selection. *Annual Review of Public Health*, 8, 191-210.

Bice TW. (1980). Social Science and Health Services Research: Contributions to Public

Policy. *Milbank Memorial Fund Quarterly/ Health and Society*, 58(2), 73-200.

Black DC, Roos PN and Burchill AC. (1995). Utilization of hospital resources. *Medical Care*, 33(12), DS55-72.

Broyles RW, Manga P, Binder DA, Angus DE and Charette A. (1983). The Use of Physician Services Under a National Health insurance Scheme: An Examination of the Canada Health Survey. *Medical Care*, 21(11), 1037-1054.

Dansky KH, Brannon D, Shea DG, Vasey J and Dirani R. (1998). Profiles of hospital, physician, and home health service use by older persons in rural areas. *The Gerontologist*, 38(3), 320-330.

Enterline PE, Salter V, McDonald AD and McDonald JC. (1973). The Distribution of Medical Services Before and After "Free" Medical Care – The Quebec Experience. *The New England Journal of Medicine*, 289(22), 1174-1178.

Erzen D, Carriere KC, Dik N, Mustard C, Roos LL, Manfreda J and Anthonisen NR. (1997). Income level and asthma prevalence and care patterns. *Am J Respir Crit Care Med.*, 155, 1060-1065.

Feldstein PJ. (1966). Research on the demand for health services. *Milbank Memorial Fund Quarterly*, 44, 128-63.

German PS, Skinner EA, Shapiro S and Salkman DS. (1976). Preventive and episodic health care of inner-city children. *Journal of Community Health*, 2, 92-106.

Greenlick MR, Hurtado AV, Pope CR, Saward EW and Yoshioka SS. (1968). Determinants of Medical Care Utilization. *Health Services Research*, 3(4), 296-315.

Gwen G. (1998). Access to medical care under strain: new pressures in Canada and Australia. *Journal of Health Politics, Policy and Law*, 23(6), 905-947.

Hershey JC, Luft HS and Gianaris JM. (1975). Making Sense Out of Utilization Data. *Medical Care*, 13(10), 838-854.

Hornbrook MC, Hurtado AV and Johnson RE. (1985). Health care episodes: Definition, measurement and use. *Medical Care Review*, 42(2), 163-218.

Jackson GP. (1998). Asthma mortality by neighborhood of of domicile. *NZ Med J.*, 101, 593-5.

Kane RL, Gardner J, Wright DD, Woolley FR, Snell GF, Sundwall DN and Castle CH. (1978). Differences in the outcomes of acute episodes of care provided by various types of family practitioners. *Journal of Family Practice*, 6, 133-38.

Kessler LG, Steinwachs DM and Hankin JR. (1980). Episodes of psychiatric utilization. *Medical Care*, 18, 1219-27.

Krahn MD, Berka C, Langlois P and Detsky AS. (1996). Direct and indirect costs of asthma in Canada, 1990. *Can Med Assoc J.*, 154, 821-831.

Leaf PJ, Livingston MM, Tischler GL, Freeman DH, Weissman MM and Myers JK. (1985). Contact with health professionals for the treatment of psychiatric and emotional problems. *Medical Care*, 23, 1322-1327.

Leaf PJ, Tischler GL, Freeman DH, Weissman MM and Myers JK. (1988). Factors affecting the utilization of specialty and general medical mental health services. *Medical Care*, 26, 9-26

Lohr KN and Brook R.H. (1980). Quality of care in episodes of respiratory illness among medicaid patients in New Mexico. *Annuals of Internal Medicine*, 92, 99-106.

Manga P. (1987). *The Income Distribution Effects of Medical Insurance in Ontario.* Toronto: Ontario Economic Council.

Maynard A. (1981). The inefficiency and inequalities of health care system of Western Europe. *Social Policy and Administration*, 15(2), 145-163.

Mayo NE, Goldberg MS, Levy AR. (1991). Changing rates of stroke in the province of Quebec, Canada: 1981-1988. *Stroke*, 22, 590-595.

McCall N and Wai HS. (1983). An analysis of the use of medicine services by the continuously enrolled aged. *Medical Care*, 21(6), 567 – 585.

McKinlay JB. (1972). Some Approaches and Problems in the Study of the Use of Services – An Overview. *Journal of Health and Social Behavior*, 13, 115-152.

McIsaac W, Goel V and Naylor CD. (1997). Socio-economic status and visits to physicians by adults in Ontario, Canada. *Journal of Health Services Research and Policy*, 2, 4-102.

Mechanic D. (1979). *Future issues in Health Care: Social Policy and the Rationing of Medical Services*. New York: The Free press.

Moscovice I. (1977). A method for analyzing resource use in ambulatory care settings. *Medical Care*, 15, 1024-44.

Mossey JM and Shapiro E. (1985). Physician use by the elderly over an eight-year period. *American Journal of Public Health*, 75, 1333-1334.

Mossey JM, Havens B and Wolinsky FD. (1989). *The consistency of formal health care utilization: physician and hospital utilization. Aging and Health Care: Social Science and Policy Perspective.* New York: Routledge.

Muller C. (1986). Review of Twenty Years of Research on Medical Care Utilization. *Health Services Research*, 21(2), 129-144.

Newhouse JP and Phelps CE. (1976). New Estimates of Price and Income Elasticities of Medical Care Services. In: Rosett RN (Ed.). *The Role of Health insurance in the Health Services Sector*, page 1-261. New York: National Bureau of Economic Research.

Nocon A and Booth T. (1991). The social impact of asthma. *Fam Pract.*, 8, 37.

Offord DR, Boyle MH, Cambell D, Goering P, Lin E, Wong M and Racine YA. (1996). One-year prevalence of psychiatric disorder in Ontarians 15 to 64 years of age. Can *J Psychiatry*, 41, 559-563.

Patrick DL and Bergner M. (1990). Measurement of health status in the 1990s. *Annual Review of Public Health*, 11, 165-183.

Patrick DL and Erickson P. (1993). *Health status and health policy: Allocating resources to health care*. New York: Oxford University Press.

Roos NP and Mustard CA. (1997). Variation in health and health care use by socioeconomic status in Winnipeg, Canada: Does the system work well? Yes or no. *The Milbank Quarterly*, 75 (1), 89-111.

Roos NP, Black C, Roos LL, Frohlich N, DeCoster C, Mustard c, Brownell MD, Shanahan M, Fergusson P, Toll F, Carriere KC, Burchill C, Fransoo R, MacWilliam L, Bogdanovic B and Friesen D. (1999). Managing health services: how the population health information system (POPULIS) works for policymakers. *Medical Care*, 37(6), JS27-JS41.

Roos NP, Shapiro E and Tate R. (1989). Does a small minority of elderly account for a majority of health care expenditures?: A sixteen-year perspective. *The Milbank Quarterly*, 67(3-4), 347-369.

Salkever DS, German PS, Shapiro S, Horky R and Skinner EA. (1976). Episodes of illness and access to care in the inner city: a comparison of HMO and non-HMO populations. *Health Services Research*, 11, 252-270.

Segovia J, Bartlett RF and Edwards AC. (1989). An empirical analysis of the dimensions of health status indicators. *Social Science and Medicine*, 29(6), 761-768.

Segovia J, Edwards AC and Bartlett RF. (1997). *Newfoundland Panel on Health and Medical Care – Adult health survey 1995: Methodology and descriptive results*. St.John's, Newfoundland: Memorial University of Newfoundland.

Shapiro S, Skinner EA, Kessler LG, Korff MV, Pearl SG, Tischler GL, Leaf PJ, Benham L, Cottler L and Regier DA. (1984). Utilization of health and mental health services: three epidemiologic catchment area sites. *Arch Gen Psychiatry*, 47, 971-978.

Siemiatycki J, Richardson L and Pless IB. (1980). Equality in Medical Care Under National Health Insurance in Montreal. *The New England journal of Medicine*, 303(1), 10-15.

Simon GE, Korff MV and Durham ML. (1994). Predictors of outpatient mental health utilization by primary care patients in a health maintenance organization. *American Journal of Psychiatry*, 151(6), 908-913.

Solon JA, Feeney JJ, Jones SH, Rigg RD and Sheps CG. (1967). Delineating episodes of Medical Care, *American Journal of Public Health*, 57(3), 401-408.

Statistics Canada. (1994a). *Health status of Canadians*. Ottawa: General Social Survey Analysis Series.

Statistics Canada (1994b). *Causes of Death, 1992*. Ottawa, Canada: Ministry of Industry, Science and Technology.

Statistics Canada (1995a). *Hospital Morbidity. 1992-93*. Ottawa, Canada: Ministry of Industry, Science and Technology.

Statistics Canada. (1995b). *National Population Health Survey, 1994-95*. Ottawa, Canada: Health Statistics Division.

Stump TE, Johnson RJ and Wolinsky FD. (1995). Changes in physician utilization over time among older adults. *Journal of Gerontology: Social Sciences*, 50B(1), S45-S58.

Sytema S, Giel R and Ten Horn GHMM. (1989). Patterns of care in the field of mental health: Conceptual definitions and research methods. *Acta Psychiatr Scand*, 79, 1-10.

Sytema S. (1991). Social indicators and psychiatric admission rates: a case-register study in the Netherlands. *Psychological Medicine*, 21, 177-184.

Tataryn JD, Roos PN and Black DC. (1995). Utilization of physician resources for ambulatory care. *Medical Care*, 33(12), DS84-99.

Temkin-Greener H and Clark KT. (1988). Ethnicity, gender, and utilization of mental health services in a medicaid population. *Social Science and Medicine*, 26(10), 989-996.

Tomiak M, Berthelot JM and Mustard CA. (1998). A profile of health care utilization of the disabled population in Manitoba. *Medical Care*, 36(9), 1383-1397.

Walker AE, Robins M and Weinfeld FD. (1981). Clinical findings. In: Weinfeld, FD ed. The national survey of stroke. *Stroke*, 12(Suppl. 1), L13-L44.

Wan TTH and Odell BG. (1981) Factors affecting the use of social and health services among the elderly. *Aging and Sociarty*, 1, 95-113.

Weissman JS and Epstein AM. (1993). The insurance gap: Does it make a difference? *Annual Review of Public Health*, 14, 243 – 270.

Wells KB, Manning WG and Bernadette B. (1986a). A comparison of the effects of sociodemographic factors and health status on use of outpatient mental health services in HMO and Fee-for-service plans. *Medical Care*, 24(10), 949-960.

Wells KB, Manning WG, Duan NH, Newhouse JP and Ware JE. (1986b). Sociodemographic factors and the use of outpatient mental health services. *Medical Care*, 24(1), 75-85.

Wolinsky FD, Culler SD, Callahan CM and Johnson R. (1994). Hospital resource consumption among older adults: a prospective analysis of episodes, length of stay, and charges over a seven-year period. *Journal of Gerontology: Social Sciences*, 49(5), S240-S252.

Wright DD, Kane RL, Snell GF and Woolley FR. (1977). Costs and outcomes for different primary care providers. *Journal of the American Medical Association*, 238 (4), 46-50.

Young TK. (1998). *Population health: concepts and methods*. New York: Oxford University Press.

Young, KM and Fisher CR. (1980). Medicare episodes of illness: a study of hospital, skilled nursing facility, and home health agency care. *Health Care Financing Review*, 2, 1-23.

Appendix

Appendix A: Household income groups used to define income adequacy variable

	Household size				
	1	2	3	4	5 or more
Household income group					
Very low	Less than $10,000	Less than $10,000	Less than $10,000	Less than $15,000	Less than $15,000
Low	$10,000 to $14,999	$10,000 to $19,999	$10,000 to $19,999	$15,000 to $29,999	$15,000 to $29,999
Lower middle	$15,000 to $29,999	$20,000 to $39,999	$20,000 to $39,999	$30,000 to $49,999	$30,000 to $59,999
Upper middle	$30,000 to $39,999	$40,000 to $49,999	$40,000 to $59,999	$50,000 to $79,999	$60,000 to $79,999
High	$40,000 or more	$50,000 or more	$60,000 or more	$80,000 or more	$80,000 or more

Source: Russell Wilkins, Canadian Centre for Health Information, Statistics Canada

Appendix B: Types of episodes

11 groups based on following criteria:

0 = no visits

1 = total visits ≥ 1 and ≤ 5 and maximum number of visits in all trimesters <3

2 = total visits ≥ 1 and ≤ 5 and maximum number of visits in all trimesters ≥ 3

3 = total visits ≥ 6 and ≤ 12 and maximum number of visits in any trimesters <3

4 = total visits ≥ 6 and ≤ 12 and maximum number of visits in all trimesters ≥ 3

5 = total visits ≥ 13 and ≤ 23 and maximum number of visits in any trimesters <3

6 = total visits ≥ 13 and ≤ 23 and maximum number of visits in any trimesters ≥ 3 and number of trimesters with ≥ 3 is ≤ 2

7 = total visits ≥ 13 and ≤ 23 and maximum number of visits in any trimesters ≥ 3 and number of trimesters with ≥ 3 is ≥ 2

8 = total visits ≥ 24 and ≤ 52 and number of trimesters with ≥ 3 is ≤ 4

9 = total visits ≥ 24 and ≤ 52 and number of trimesters with ≥ 3 is > 4

10 = total visits > 52 or number of trimesters with ≥ 3 is > 8

Appendix C: Programs designed for this study

```
**************************************************************************
                              Program-1.sps
            Goal: To create gp.por,sp.por,hosp.por and total.por
**************************************************************************
***Explanation of variables used in this program:
***hh: household number; subj: subject number; servdate: date of service;
***spec: GP or specialist; diag: diagnosis group; diagcode: diagnosis code;
***yearmcp: fiscal year of service; util4yr: in all 4 years 1992-6

get file 'mcp.sav'
   /keep=hh subj servdate spec util4yr diag diagcode yearmcp
select if util4yr eq 1 and (yearmcp eq 9293 or yearmcp eq 9394 or yearmcp eq
                9495 or yearmcp eq 9596)  /select subjects who have utilization in 4 yrs

print format servdate(adate9)

temporary
select if spec=1   /only select GP visits
export outfile=gp.por   /save those subjects who have GP visits in file gp.por

temporary
select if spec=2   /only select specialist visits
export outfile=sp.por   /save those subjects who have specialist visits in file sp.por

***explanation of variables: disdate: date of discharge;  gpdx1: primary diagnosis code;
***los: length of stay

get file 'hosp.sav'
   /keep=hh subj disdate util4yr gpdx1 los
select if util4yr eq 1

print format disdate(adate9)
export outfile=hosp.por   /select those subjects who were hospitalized in 4 yrs (92-96) in
                          /file hosp.por


***Explanation of variables: agegp6: age in 6 groups;  SES9: education and income;
***urban: residence in three areas; hstat: health status;  hstatgp2: health status in 2 groups

get file 'panel.sav'
```

```
    /keep=hh subj age agegp6 sex SES9 urban hstat hstatgp2 util4yr

select if util4yr eq 1  /select those subjects who have medical care utilization in 4 years
compute inteage=trunc(age)  /to create a new variable inteage: integer type of age
export outfile=total.por  /

finish
```

****

Program-2.sas

Goal:To transfer SPSS portable files gp.por, sp.por, hosp.por, and total.por into SAS files: sasgp.ssd04,sassp.ssd04 and sashosp.ssd04 respectively

****

```
libname xyz 'file location';
filename abc 'gp.por';
proc convert spss=abc out=xyz.sasgp;
proc contents data=xyz.sasgp;
proc print data=xyz.sasgp;
run;

libname xyz 'file location';
filename abc 'sp.por';
proc convert spss=abc out=xyz.sassp;
proc contents data=xyz.sassp;
proc print data=xyz.sassp;
run;

libname xyz 'file location';
filename abc 'hosp.por';
proc convert spss=abc out=xyz.sashosp;
proc contents data=xyz.sashosp;
proc print data=xyz.sashosp;
run;

libname xyz 'file location';
filename abc 'total.por';
proc convert spss=abc out=xyz.sastotal;
proc contents data=xyz.sastotal;
proc print data=xyz.sastoatal;
run;
```

****

Program-3.sas

Goal: To create the data file which contains variables age inteage hstat hstatgp2 educgp5 incad5 sescore1 urban sex; to get files sasgp1.ssd04, sassp1.ssd04 and sashos1.ssd04 using the same program

****

```
libname xyz 'file location';

proc sort data=xyz.sastotal;
 by hh subj;
run;
proc sort data=xyz.sasgp;
 by hh subj;
run;

data xyz.merge1;
set xyz.sasgp;
retain flag 1;  /set a flag variable to indicate each subject in file sasgp.ssd04
output;
run;

data xyz.merge2;
merge xyz.sastotal xyz.merge1;
by hh subj;
run;

data xyz.sasgp1;
set xyz.merge2;
options missing=0;

if flag=1 then output;
run;

proc contents data=xyz.sasgp1;
proc print data=xyz.sasgp1;
run;
```

****

Program-4.sas

Goal: To get file logtotal.ssd04 which contains variable hosp, age, inteage, hstat, hstatgp2, educgp5, incad5, sescore1, sex, and urban for all subjects with medical care utilization, including GP visits, specialist visits, or hospitalizations in 4 years from 1992-96.

****

```
***Explanation of variables: indicate hospitalized or not;  inteage: integer type of age;
***hstat: health status;  hstatgp2: health status in 2 groups;
***educgp5: education in 5 groups;  incad5: income in 5 groups;
***sescore1: education and income;  urban: residence in three areas

libname xyz 'file location';
data xyz.sasgp2;
set xyz.sasgp1;
by hh subj;
retain _res1 _res2 _res3 _res4 _res5 _res6 _res7 _res8 _res9 _res10 _res11
     _res12 _res13 _res14 _res15  _res16 0;  /res1 to res16 are medium variables meaning
                                              /total GP visits in each trimester of total 16
                                              /trimesters

if first.subj then do;
  _res1=0;
  _res2=0;
  _res3=0;
  _res4=0;
  _res5=0;
  _res6=0;
  _res7=0;
  _res8=0;
  _res9=0;
  _res10=0;
  _res11=0;
  _res12=0;
  _res13=0;
  _res14=0;
  _res15=0;
  _res16=0;
  if first.subj & last.subj then do;
   if servdate >='01Apr92'd & servdate <='30Jun92'd then _res1=1;
   else if servdate >= '01Jul92'd & servdate <= '30Sep92'd then _res2=1;
   else if servdate >= '01Oct92'd & servdate <= '31Dec92'd then _res3=1;
```

```
    else if servdate >= '01Jan93'd & servdate <= '31Mar93'd then _res4=1;
    else if servdate >= '01Apr93'd & servdate <= '30Jun93'd then _res5=1;
    else if servdate >= '01Jul93'd & servdate <= '30Sep93'd then _res6=1;
    else if servdate >= '01Oct93'd & servdate <= '31Dec93'd then _res7=1;
    else if servdate >= '01Jan94'd & servdate <= '31Mar94'd then _res8=1;
    else if servdate >= '01Apr94'd & servdate <= '30Jun94'd then _res9=1;
    else if servdate >= '01Jul94'd & servdate <= '30Sep94'd then _res10=1;
    else if servdate >= '01Oct94'd & servdate <= '31Dec94'd then _res11=1;
    else if servdate >= '01Jan95'd & servdate <= '31Mar95'd then _res12=1;
    else if servdate >= '01Apr95'd & servdate <= '30Jun95'd then _res13=1;
    else if servdate >= '01Jul95'd & servdate <= '30Sep95'd then _res14=1;
    else if servdate >= '01Oct95'd & servdate <= '31Dec95'd then _res15=1;
    else if servdate >= '01Jan96'd & servdate <= '31Mar96'd then _res16=1;
    output;
  end;
end;
if servdate >='01Apr92'd & servdate <='30Jun92'd then _res1=_res1+1;
else if servdate >= '01Jul92'd & servdate <= '30Sep92'd then _res2=_res2+1;
else if servdate >= '01Oct92'd & servdate <= '31Dec92'd then _res3=_res3+1;
else if servdate >= '01Jan93'd & servdate <= '31Mar93'd then _res4=_res4+1;
else if servdate >= '01Apr93'd & servdate <= '30Jun93'd then _res5=_res5+1;
else if servdate >= '01Jul93'd & servdate <= '30Sep93'd then _res6=_res6+1;
else if servdate >= '01Oct93'd & servdate <= '31Dec93'd then _res7=_res7+1;
else if servdate >= '01Jan94'd & servdate <= '31Mar94'd then _res8=_res8+1;
else if servdate >= '01Apr94'd & servdate <= '30Jun94'd then _res9=_res9+1;
else if servdate >= '01Jul94'd & servdate <= '30Sep94'd then _res10=_res10+1;
else if servdate >= '01Oct94'd & servdate <= '31Dec94'd then _res11=_res11+1;
else if servdate >= '01Jan95'd & servdate <= '31Mar95'd then _res12=_res12+1;
else if servdate >= '01Apr95'd & servdate <= '30Jun95'd then _res13=_res13+1;
else if servdate >= '01Jul95'd & servdate <= '30Sep95'd then _res14=_res14+1;
else if servdate >= '01Oct95'd & servdate <= '31Dec95'd then _res15=_res15+1;
else if servdate >= '01Jan96'd & servdate <= '31Mar96'd then _res16=_res16+1;
if last.subj & ^first.subj then output;
run;

* The above is for GP visits
*****************************************************************************
data xyz.sassp2;
set xyz.sassp1;
by hh subj;
retain _res1 _res2 _res3 _res4 _res5 _res6 _res7 _res8 _res9 _res10 _res11
     _res12 _res13 _res14 _res15  _res16 0;  / res1 to res16 are medium variables meaning
```

```
                                        /total specialist visits in each trimester of
                                        /total 16 trimesters

if first.subj then do;
 _res1=0;
 _res2=0;
 _res3=0;
 _res4=0;
 _res5=0;
 _res6=0;
 _res7=0;
 _res8=0;
 _res9=0;
 _res10=0;
 _res11=0;
 _res12=0;
 _res13=0;
 _res14=0;
 _res15=0;
 _res16=0;
 if first.subj & last.subj then do;
  if servdate >='01Apr92'd & servdate <='30Jun92'd then _res1=1;
  else if servdate >= '01Jul92'd & servdate <= '30Sep92'd then _res2=1;
  else if servdate >= '01Oct92'd & servdate <= '31Dec92'd then _res3=1;
  else if servdate >= '01Jan93'd & servdate <= '31Mar93'd then _res4=1;
  else if servdate >= '01Apr93'd & servdate <= '30Jun93'd then _res5=1;
  else if servdate >= '01Jul93'd & servdate <= '30Sep93'd then _res6=1;
  else if servdate >= '01Oct93'd & servdate <= '31Dec93'd then _res7=1;
  else if servdate >= '01Jan94'd & servdate <= '31Mar94'd then _res8=1;
  else if servdate >= '01Apr94'd & servdate <= '30Jun94'd then _res9=1;
  else if servdate >= '01Jul94'd & servdate <= '30Sep94'd then _res10=1;
  else if servdate >= '01Oct94'd & servdate <= '31Dec94'd then _res11=1;
  else if servdate >= '01Jan95'd & servdate <= '31Mar95'd then _res12=1;
  else if servdate >= '01Apr95'd & servdate <= '30Jun95'd then _res13=1;
  else if servdate >= '01Jul95'd & servdate <= '30Sep95'd then _res14=1;
  else if servdate >= '01Oct95'd & servdate <= '31Dec95'd then _res15=1;
  else if servdate >= '01Jan96'd & servdate <= '31Mar96'd then _res16=1;
  output;
 end;
end;
if servdate >='01Apr92'd & servdate <='30Jun92'd then _res1=_res1+1;
else if servdate >= '01Jul92'd & servdate <= '30Sep92'd then _res2=_res2+1;
```

```
else if servdate >= '01Oct92'd & servdate <= '31Dec92'd then _res3=_res3+1;
else if servdate >= '01Jan93'd & servdate <= '31Mar93'd then _res4=_res4+1;
else if servdate >= '01Apr93'd & servdate <= '30Jun93'd then _res5=_res5+1;
else if servdate >= '01Jul93'd & servdate <= '30Sep93'd then _res6=_res6+1;
else if servdate >= '01Oct93'd & servdate <= '31Dec93'd then _res7=_res7+1;
else if servdate >= '01Jan94'd & servdate <= '31Mar94'd then _res8=_res8+1;
else if servdate >= '01Apr94'd & servdate <= '30Jun94'd then _res9=_res9+1;
else if servdate >= '01Jul94'd & servdate <= '30Sep94'd then _res10=_res10+1;
else if servdate >= '01Oct94'd & servdate <= '31Dec94'd then _res11=_res11+1;
else if servdate >= '01Jan95'd & servdate <= '31Mar95'd then _res12=_res12+1;
else if servdate >= '01Apr95'd & servdate <= '30Jun95'd then _res13=_res13+1;
else if servdate >= '01Jul95'd & servdate <= '30Sep95'd then _res14=_res14+1;
else if servdate >= '01Oct95'd & servdate <= '31Dec95'd then _res15=_res15+1;
else if servdate >= '01Jan96'd & servdate <= '31Mar96'd then _res16=_res16+1;
if last.subj & ^first.subj then output;
run;

* The above is for specialist visits
****************************************************************************

libname xyz 'file location';

data xyz.sashos2;
set xyz.sashos1;
by hh subj;
retain _res1 _res2 _res3 _res4 _res5 _res6 _res7 _res8 _res9 _res10 _res11
     _res12 _res13 _res14 _res15  _res16 0;  /res1 to res16 are medium variables meaning
                                              /total number of hospitalizations in each
                                              /trimester of total 16 trimesters

retain hosp 1;

if first.subj then do;
  _res1=0;
  _res2=0;
  _res3=0;
  _res4=0;
  _res5=0;
  _res6=0;
  _res7=0;
  _res8=0;
  _res9=0;
```

```
_res10=0;
_res11=0;
_res12=0;
_res13=0;
_res14=0;
_res15=0;
_res16=0;
if first.subj & last.subj then do;
  if disdate >='01Apr92'd & disdate <='30Jun92'd then _res1=1;
  else if disdate >= '01Jul92'd & disdate <= '30Sep92'd then _res2=1;
  else if disdate >= '01Oct92'd & disdate <= '31Dec92'd then _res3=1;
  else if disdate >= '01Jan93'd & disdate <= '31Mar93'd then _res4=1;
  else if disdate >= '01Apr93'd & disdate <= '30Jun93'd then _res5=1;
  else if disdate >= '01Jul93'd & disdate <= '30Sep93'd then _res6=1;
  else if disdate >= '01Oct93'd & disdate <= '31Dec93'd then _res7=1;
  else if disdate >= '01Jan94'd & disdate <= '31Mar94'd then _res8=1;
  else if disdate >= '01Apr94'd & disdate <= '30Jun94'd then _res9=1;
  else if disdate >= '01Jul94'd & disdate <= '30Sep94'd then _res10=1;
  else if disdate >= '01Oct94'd & disdate <= '31Dec94'd then _res11=1;
  else if disdate >= '01Jan95'd & disdate <= '31Mar95'd then _res12=1;
  else if disdate >= '01Apr95'd & disdate <= '30Jun95'd then _res13=1;
  else if disdate >= '01Jul95'd & disdate <= '30Sep95'd then _res14=1;
  else if disdate >= '01Oct95'd & disdate <= '31Dec95'd then _res15=1;
  else if disdate >= '01Jan96'd & disdate <= '31Mar96'd then _res16=1;
  output;
end;
end;
if disdate >='01Apr92'd & disdate <='30Jun92'd then _res1=_res1+1;
else if disdate >= '01Jul92'd & disdate <= '30Sep92'd then _res2=_res2+1;
else if disdate >= '01Oct92'd & disdate <= '31Dec92'd then _res3=_res3+1;
else if disdate >= '01Jan93'd & disdate <= '31Mar93'd then _res4=_res4+1;
else if disdate >= '01Apr93'd & disdate <= '30Jun93'd then _res5=_res5+1;
else if disdate >= '01Jul93'd & disdate <= '30Sep93'd then _res6=_res6+1;
else if disdate >= '01Oct93'd & disdate <= '31Dec93'd then _res7=_res7+1;
else if disdate >= '01Jan94'd & disdate <= '31Mar94'd then _res8=_res8+1;
else if disdate >= '01Apr94'd & disdate <= '30Jun94'd then _res9=_res9+1;
else if disdate >= '01Jul94'd & disdate <= '30Sep94'd then _res10=_res10+1;
else if disdate >= '01Oct94'd & disdate <= '31Dec94'd then _res11=_res11+1;
else if disdate >= '01Jan95'd & disdate <= '31Mar95'd then _res12=_res12+1;
else if disdate >= '01Apr95'd & disdate <= '30Jun95'd then _res13=_res13+1;
else if disdate >= '01Jul95'd & disdate <= '30Sep95'd then _res14=_res14+1;
else if disdate >= '01Oct95'd & disdate <= '31Dec95'd then _res15=_res15+1;
```

```
else if disdate >= '01Jan96'd & disdate <= '31Mar96'd then _res16=_res16+1;
if last.subj & ^first.subj then output;
run;
* The above is for hospitalizations
*************************************************************************
data xyz.sastot(keep=hh subj hosp agegp6 educgp5 incad5 hstat hstatgp2
    sescore1 sex urban);
merge xyz.sasgp xyz.sassp xyz.sashos;  /sasgp.ssd04 contains all GP visits; /sassp.ssd04
                                        contains all specialist visits; /sashos.ssd04
                                        contains all numbers of /hospitalizations.
                                        Merge the three files to /create a new file:
                                        sastot.ssd04 which /contains GP visits,
                                        specialist visits, or /hospitalizations
by hh subj;
options missing=0;
if hosp=missing then hosp=0;
output;
run;

libname xyz 'file location';

proc sort data=xyz.sastotal;
  by hh subj;
run;
proc sort data=xyz.sastot;
  by hh subj;
run;

data xyz.logtotal(keep=hh subj hosp agegp6 educgp5 incad5 hstat hstatgp2
    sescore1 sex urban age inteage);
merge xyz.sastotal xyz.sastot;
by hh subj;
options missing=0;
if hosp=missing then hosp=0;
output;
run;

proc freq data=xyz.logtotal;
  tables hosp;
run;
```

```
*****************************************************************************
```

Program-5.sas

Goal: Design this program to calculate frequency of three type of episodes: non-episodic, isolated episodes, or continuous episodes for GP and specialist visits separately for all subjects with medical care utilization in 4 years.

```
*****************************************************************************

libname xyz 'file location';

proc sort data=xyz.sasgp1;
  by hh subj;
run;

data xyz.sasresp;
set xyz.sasgp1;
by hh subj;
retain _res1 _res2 _res3 _res4 _res5 _res6 _res7 _res8 _res9 _res10 _res11
      _res12 _res13 _res14 _res15  _res16 0; /_res1 to _res16 are medium variables
                                             /meaning total GP visits in each trimester of
                                             /total 16 trimesters

if first.subj then do;
  _res1=0;
  _res2=0;
  _res3=0;
  _res4=0;
  _res5=0;
  _res6=0;
  _res7=0;
  _res8=0;
  _res9=0;
  _res10=0;
  _res11=0;
  _res12=0;
  _res13=0;
  _res14=0;
  _res15=0;
  _res16=0;
  if first.subj & last.subj then do;
    if servdate >='01Apr92'd & servdate <='30Jun92'd then _res1=1;
    else if servdate >= '01Jul92'd & servdate <= '30Sep92'd then _res2=1;
    else if servdate >= '01Oct92'd & servdate <= '31Dec92'd then _res3=1;
```

```
    else if servdate >= '01Jan93'd & servdate <= '31Mar93'd then _res4=1;
    else if servdate >= '01Apr93'd & servdate <= '30Jun93'd then _res5=1;
    else if servdate >= '01Jul93'd & servdate <= '30Sep93'd then _res6=1;
    else if servdate >= '01Oct93'd & servdate <= '31Dec93'd then _res7=1;
    else if servdate >= '01Jan94'd & servdate <= '31Mar94'd then _res8=1;
    else if servdate >= '01Apr94'd & servdate <= '30Jun94'd then _res9=1;
    else if servdate >= '01Jul94'd & servdate <= '30Sep94'd then _res10=1;
    else if servdate >= '01Oct94'd & servdate <= '31Dec94'd then _res11=1;
    else if servdate >= '01Jan95'd & servdate <= '31Mar95'd then _res12=1;
    else if servdate >= '01Apr95'd & servdate <= '30Jun95'd then _res13=1;
    else if servdate >= '01Jul95'd & servdate <= '30Sep95'd then _res14=1;
    else if servdate >= '01Oct95'd & servdate <= '31Dec95'd then _res15=1;
    else if servdate >= '01Jan96'd & servdate <= '31Mar96'd then _res16=1;
    output;
  end;
end;
if servdate >='01Apr92'd & servdate <='30Jun92'd then _res1=_res1+1;
else if servdate >= '01Jul92'd & servdate <= '30Sep92'd then _res2=_res2+1;
else if servdate >= '01Oct92'd & servdate <= '31Dec92'd then _res3=_res3+1;
else if servdate >= '01Jan93'd & servdate <= '31Mar93'd then _res4=_res4+1;
else if servdate >= '01Apr93'd & servdate <= '30Jun93'd then _res5=_res5+1;
else if servdate >= '01Jul93'd & servdate <= '30Sep93'd then _res6=_res6+1;
else if servdate >= '01Oct93'd & servdate <= '31Dec93'd then _res7=_res7+1;
else if servdate >= '01Jan94'd & servdate <= '31Mar94'd then _res8=_res8+1;
else if servdate >= '01Apr94'd & servdate <= '30Jun94'd then _res9=_res9+1;
else if servdate >= '01Jul94'd & servdate <= '30Sep94'd then _res10=_res10+1;
else if servdate >= '01Oct94'd & servdate <= '31Dec94'd then _res11=_res11+1;
else if servdate >= '01Jan95'd & servdate <= '31Mar95'd then _res12=_res12+1;
else if servdate >= '01Apr95'd & servdate <= '30Jun95'd then _res13=_res13+1;
else if servdate >= '01Jul95'd & servdate <= '30Sep95'd then _res14=_res14+1;
else if servdate >= '01Oct95'd & servdate <= '31Dec95'd then _res15=_res15+1;
else if servdate >= '01Jan96'd & servdate <= '31Mar96'd then _res16=_res16+1;
if last.subj & ^first.subj then output;
run;

data xyz.epires1;
set xyz.sasresp;
by hh subj;
options missing=0;
totres=sum(of _res1-_res16);  /The total number of GP visits in the 16 trimesters
maxres=max(of _res1-_res16);  /The maximum number of GP visits in the 16 trimesters
```

```
/* By using arrary, create a new character type variable 'bepichar' which
contains 16 characters. The 16 characters are the combination of '1' and
'0', which separately represent trimester>=2 visits and trimester<2 visits.*/

array res{16} _res1 _res2 _res3 _res4 _res5 _res6 _res7 _res8 _res9
        _res10 _res11 _res12 _res13 _res14 _res15 _res16; /set the 16 variables into a
                                                          /array
array bres{16} _bres1 _bres2 _bres3 _bres4 _bres5 _bres6 _bres7 _bres8 _bres9
        _bres10 _bres11 _bres12 _bres13 _bres14 _bres15 _bres16; /set another array
                                                                  /bres{16} using 16
                                                                  /new variables
```

***Variable BEPISOD: identify trimesters with 2 or more visits. To construct this variable, make 16 variables: bres{1} to bres{16} as follows: If the first trimester has 2 or more visits make the first bres{1}=1, otherwise equal to 0; if the second trimester has 2 or more visits make the second bres{2} =10, otherwise equal to 0; if the third trimester has 2 or more visits make the third bres{3}=100, otherwise equal to 0, etc. This will give 16 variables with values from 1, 10, 100, 1000 up to 1000000000000000 when every trimester has 2 or more visits. Summing these 16 variables to a new variable: bepisod results in a (maximum) 16 digit number. ***

```
number=1;  /just a temporary variable
if res{1} >=2 then bres{1}=1;  /get variable bres{i} from variable res{i}
if res{1} <2 then bres{1}=0;
do i=2 to 16;
  if res{i} >=2 then bres{i}=number*10;
  if res{i} <2 then bres{i}=0;
  number=number*10;
end;
bepisod=sum(of _bres1-_bres16);  /the total sum of variable _bres1 to _bres16
bepichal=put(bepisod,$16.);  /transfer variable bepisod into character type as a new
                              /variable bepichal
```

Variable BEPICHAR: Converting BEPISOD to characters gives a 16 characters variable with 0 and 1's. This gives a look of the trimesters with 2 or more visits over time. e.g. 0000000000100111 is a case where there are 2 or more visits in the first, second, third, and sixth trimesters. All other trimesters have 0 or 1 visit.

```
bepichar=left(bepichal);  /move each character of bepichal to the most left side
len=length(bepichar);  /Get the length of variable bepichar
do while(len<16);  /to put enough zero on the left side of variable bepichar to make the
```

```
                    /length of bepicha1 equal to 16
 bepichar='0'||bepichar;
 len=length(bepichar);
end;

/* To calculate how many '1' and how many '0' in variable 'bepichar' by using
two new variables 'num1' and 'num2' separately. */

num1=0;
num2=0;
do i=1 to 16;
 if res{i} >=2 then num1=num1+1;
 if res{i} <2 then num2=num2+1;
end;

/* Variable 'num' means the total number of trimester >=3. */

num=0; /This variable means how many number of trimester with 3 or more GP visits
do i=1 to 16;
 if res{i} >=3 then num=num+1;
end;
run;
```

***Variable COMB: identify trimesters with 2 or more visits. Count number of consecutive trimesters with 2 or more visits and turn into a 16 character variable where the first character implies "1" trimester with 2 or more visits , the second implies "2" trimesters with 2 or more visits, etc.

e.g. 1011000000000000 is a case with "4" consecutive trimesters with 2 or more visits, "3" consecutive trimesters with 2 or more visits, and "1" single non-consecutive trimester with 2 or more visits.

e.g. 0000001100000000 is a case with "7" and "8" consecutive trimesters with 2 or more visits, i.e. there is only "1" trimester (16-(7+8)) that does not have 2 or more visits.***

*The following part is to get variable COMB

```
libname xyz 'file location';
data xyz.numbof1;
set xyz.epires1;
array numb{16} numb1-numb16(0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0);

bepisod1=bepisod;
i=0;
```

```
flag=0;
do until(i>16);
  decimal=mod(bepisod1,10);
  i=i+1;
  if decimal=1 then do;
    flag=flag+1;
    bepisod1=int(bepisod1/10);
  end;
  else if decimal=0 then do;
    bepisod1=int(bepisod1/10);
    if flag>0 then do;
      numb{flag}=numb{flag}+1;
      flag=0;
    end;
  end;
end;
output;
do i=1 to 16;
  numb{i}=0;
end;
flag=0;
run;

data xyz.combres;
set xyz.numbofl;
array numb{16} numb1-numb16;
array cnumb{16} $1.;

do i=1 to 16;
  cnumb{i}=put(numb{i},$1.);
end;

comb=cnumb1||cnumb2||cnumb3||cnumb4||cnumb5||cnumb6||cnumb7||cnumb8||cnumb9||
   cnumb10||cnumb11||cnumb12||cnumb13||cnumb14||cnumb15||cnumb16;
output;
run;

proc freq data=xyz.combres;
  tables comb;
run;

/* the following part is in order to get variable "pattern" which contains three types: non-
```

```
episodic, isolated episodes, and continuous episodes */

libname xyz 'file location';
data xyz.nonepi xyz.isol xyz.cont;
set xyz.combres;

if cnumb1='0' and cnumb2='0' and cnumb3='0' and cnumb4='0' and cnumb5='0'
  and cnumb6='0' and cnumb7='0' and cnumb8='0' and cnumb9='0' and cnumb10='0'
  and cnumb11='0' and cnumb12='0' and cnumb13='0' and cnumb14='0' and
  cnumb15='0' and cnumb16='0' then output xyz.nonepi;

if cnumb1<>'0' and cnumb2='0' and cnumb3='0' and cnumb4='0' and cnumb5='0'
  and cnumb6='0' and cnumb7='0' and cnumb8='0' and cnumb9='0' and cnumb10='0'
  and cnumb11='0' and cnumb12='0' and cnumb13='0' and cnumb14='0' and
  cnumb15='0' and cnumb16='0' then output xyz.isol;

if cnumb6<>'0' or cnumb7<>'0' or cnumb8<>'0' or cnumb9<>'0' or cnumb10<>'0'
  or cnumb11<>'0' or cnumb12<>'0' or cnumb13<>'0' or cnumb14<>'0' or
  cnumb15<>'0' or cnumb16<>'0' or cnumb2<>'0' or cnumb3<>'0' or cnumb4<>'0'
  or cnumb5<>'0' then output xyz.cont;
run;

data xyz.nonepi1;
set xyz.nonepi(keep=hh subj agegp6 educgp5 hstat hstatgp2 incad5 sescore1 sex
          urban age inteage);
retain pattern 0;
output;
run;

data xyz.isol1;
set xyz.isol(keep=hh subj agegp6 educgp5 hstat hstatgp2 incad5 sescore1 sex
          urban age inteage);
retain pattern 1;
output;
run;

data xyz.cont1;
set xyz.cont(keep=hh subj agegp6 educgp5 hstat hstatgp2 incad5 sescore1 sex
          urban age inteage);
retain pattern 2;
output;
run;
```

```
data xyz.logsasgp;
set xyz.nonepi1 xyz.isol1 xyz.cont1;
output;
run;

proc freq data=xyz.logsasgp;
  tables pattern;
run;

proc contents data=xyz.logsasgp;
proc print data=xyz.logsasgp;
run;
```

****

Program-6.sps

Goal: To get the subjects who have cardiovascular disease, mental diseas and chronic respiratory condition separately, and then matched with file PANEL to get a new file which contains the total subjects who have these three kinds of diagnoses and we can investigate some characteristics of them further by demographic and socioeconomic variables such as sex, age, urban, health status, socioeconomic score and so on.

****

```
***Explanation of variables: dx1: ICD9_code;  los: length of stay. ***

get file 'hosp.sav'
sort cases by hh subj

temporary
select if (dx1 eq '410' or dx1 eq '411' or dx1 eq '412' or dx1 eq '413' or
  dx1 eq '414' or dx1 eq '4140' or dx1 eq '4141' or dx1 eq '4148' or dx1 eq
  '4149' or dx1 eq '431' or dx1 eq '434' or dx1 eq '4340' or dx1 eq '4341' or
  dx1 eq '4349' or dx1 eq '436')

aggregate outfile=strlos.sav
        /break=hh subj
        /los1 'Number of LOS for Stroke'=sum(los)

temporary
select if (dx1 eq '295' or dx1 eq '2950' or dx1 eq '2951' or dx1 eq '2952'
  or dx1 eq '2953' or dx1 eq '2954' or dx1 eq '2955' or dx1 eq '2956'
  or dx1 eq '2957' or dx1 eq '2958' or dx1 eq '2959' or dx1 eq '296'
  or dx1 eq '2960' or dx1 eq '2961' or dx1 eq '2962' or dx1 eq '2963'
  or dx1 eq '2964' or dx1 eq '2965' or dx1 eq '2966' or dx1 eq '2968'
  or dx1 eq '2969' or dx1 eq '297'  or dx1 eq '2970' or dx1 eq '2971'
  or dx1 eq '2972' or dx1 eq '2973' or dx1 eq '2978' or dx1 eq '2979'
  or dx1 eq '298'  or dx1 eq '2980' or dx1 eq '2981' or dx1 eq '2982'
  or dx1 eq '2983' or dx1 eq '2984' or dx1 eq '2988' or dx1 eq '2989'
  or dx1 eq '303'  or dx1 eq '309'  or dx1 eq '3090' or dx1 eq '3091'
  or dx1 eq '2992' or dx1 eq '3093' or dx1 eq '3094' or dx1 eq '3098'
  or dx1 eq '3099')

aggregate outfile=menlos.sav
        /break=hh subj
        /los2 'Number of LOS for Mental Diseas'=sum(los)
```

```
temporary
select if (dx1 eq '490' or dx1 eq '491' or dx1 eq '4910' or dx1 eq '4911' or
        dx1 eq '4912' or dx1 eq '4918' or dx1 eq '4919' or dx1 eq '492' or
        dx1 eq '493' or dx1 eq '4930' or dx1 eq '4931' or dx1 eq '4939' or
        dx1 eq '494' or dx1 eq '496')
aggregate outfile=reslos.sav
        /break=hh subj
        /los3 'Number of LOS for Respiratory'=sum(los)

match files file=strlos.sav
        /file=menlos.sav
        /file=reslos.sav/by hh subj
save outfile=dislos.sav

get file 'dislos.sav'
display sorted names/vars=all
freq los1 los2 los3
save outfile=dislos.sav

* The above is to calculate all length of stay for cardiovascular disease, mental disease,
and chronic respiratory conditions separately *
***************************************************************************
get file 'hosp.sav'
sort cases by hh subj

temporary
select if (dx1 eq '410' or dx1 eq '411' or dx1 eq '412' or dx1 eq '413' or
  dx1 eq '414' or dx1 eq '4140' or dx1 eq '4141' or dx1 eq '4148' or dx1 eq
  '4149' or dx1 eq '431' or dx1 eq '434' or dx1 eq '4340' or dx1 eq '4341' or
  dx1 eq '4349' or dx1 eq '436')

aggregate outfile=strhos.sav
        /break=hh subj
        /hosp1 'Number of Hospitalizations for Stroke'=sum(util4yr)

temporary
select if (dx1 eq '295' or dx1 eq '2950' or dx1 eq '2951' or dx1 eq '2952'
  or dx1 eq '2953' or dx1 eq '2954' or dx1 eq '2955' or dx1 eq '2956'
  or dx1 eq '2957' or dx1 eq '2958' or dx1 eq '2959' or dx1 eq '296'
  or dx1 eq '2960' or dx1 eq '2961' or dx1 eq '2962' or dx1 eq '2963'
  or dx1 eq '2964' or dx1 eq '2965' or dx1 eq '2966' or dx1 eq '2968'
```

```
 or dx1 eq '2969' or dx1 eq '297'  or dx1 eq '2970' or dx1 eq '2971'
 or dx1 eq '2972' or dx1 eq '2973' or dx1 eq '2978' or dx1 eq '2979'
 or dx1 eq '298'  or dx1 eq '2980' or dx1 eq '2981' or dx1 eq '2982'
 or dx1 eq '2983' or dx1 eq '2984' or dx1 eq '2988' or dx1 eq '2989'
 or dx1 eq '303'  or dx1 eq '309'  or dx1 eq '3090' or dx1 eq '3091'
 or dx1 eq '2992' or dx1 eq '3093' or dx1 eq '3094' or dx1 eq '3098'
 or dx1 eq '3099')

aggregate outfile=menhos.sav
        /break=hh subj
        /hosp2 'Number of Hospitalizations for Mental Diseas'=sum(util4yr)

temporary
select if (dx1 eq '490' or dx1 eq '491' or dx1 eq '4910' or dx1 eq '4911' or
        dx1 eq '4912' or dx1 eq '4918' or dx1 eq '4919' or dx1 eq '492' or
        dx1 eq '493' or dx1 eq '4930' or dx1 eq '4931' or dx1 eq '4939' or
        dx1 eq '494' or dx1 eq '496')

aggregate outfile=reshos.sav
        /break=hh subj
        /hosp3 'Number of Hospitalizations for Respiratory'=sum(util4yr)

match files file=strhos.sav
         /file=menhos.sav
         /file=reshos.sav/by hh subj
save outfile=dishos.sav

get file 'dishos.sav'
display sorted names/vars=all
freq hosp1 hosp2 hosp3
save outfile=dishos.sav

* The above is to calculate the total numbers of hospitalizations for cardiovascular
disease, mental disease, and chronic respiratory conditions separately. *
*********************************************************************

***Explanation of variables: diagcode: diagnosis code; gp1: number of GP visits for
cardiovascular disease; gp2: number of GP visits for mental disease; gp3: number of GP
visits for chronic respiratory conditions. ***

get file 'mcp.sav'
sort cases by hh subj
```

```
select if spec eq 1  /only select GP visits

temporary
select if (diagcode eq 410 or diagcode eq 411 or diagcode eq 412
        or diagcode eq 413 or diagcode eq 414 or diagcode eq 4140
        or diagcode eq 4141 or diagcode eq 4148 or diagcode eq 4149
        or diagcode eq 431 or diagcode eq 434 or diagcode eq 4340
        or diagcode eq 4341 or diagcode eq 4349 or diagcode eq 436)

aggregate outfile=strgp.sav
        /break=hh subj
        /gp1 'Number of GP visits for cardiovascular disease'=sum(util4yr)

temporary
select if (diagcode eq 295 or diagcode eq 2950 or diagcode eq 2951 or
        diagcode eq 2952 or diagcode eq 2953 or diagcode eq 2954 or
        diagcode eq 2955 or diagcode eq 2956 or diagcode eq 2957 or
        diagcode eq 2958 or diagcode eq 2959 or diagcode eq 296 or
        diagcode eq 2960 or diagcode eq 2961 or diagcode eq 2962 or
        diagcode eq 2963 or diagcode eq 2964 or diagcode eq 2965 or
        diagcode eq 2966 or diagcode eq 2968 or diagcode eq 2969 or
        diagcode eq 297  or diagcode eq 2970 or diagcode eq 2971 or
        diagcode eq 2972 or diagcode eq 2973 or diagcode eq 2978 or
        diagcode eq 2979 or diagcode eq 298  or diagcode eq 2980 or
        diagcode eq 2981 or diagcode eq 2982 or diagcode eq 2983 or
        diagcode eq 2984 or diagcode eq 2988 or diagcode eq 2989 or
        diagcode eq 303  or diagcode eq 309  or diagcode eq 3090 or
        diagcode eq 3091 or diagcode eq 2992 or diagcode eq 3093 or
        diagcode eq 3094 or diagcode eq 3098 or diagcode eq 3099)

aggregate outfile=mengp.sav
        /break=hh subj
        /gp2 'Number of GP visits for Mental Disease'=sum(util4yr)

temporary
select if (diagcode eq 490 or diagcode eq 491 or diagcode eq 4910 or
        diagcode eq 4911 or diagcode eq 4912 or diagcode eq 4918 or
        diagcode eq 4919 or diagcode eq 492 or diagcode eq 493 or
        diagcode eq 4930 or diagcode eq 4931 or diagcode eq 4939 or
        diagcode eq 494 or diagcode eq 496)

aggregate outfile=resgp.sav
```

```
        /break=hh subj
        /gp3 'Number of GP visits for Respiratory conditions'=sum(util4yr)

match files file=strgp.sav
        /file=mengp.sav
        /file=resgp.sav/by hh subj
save outfile=disgp.sav

get file 'disgp.sav'
display sorted names/vars=all
freq gp1 gp2 gp3
save outfile=disgp.sav
* The above is to calculate the total GP visits for cardiovascular disease, mental disease,
and chronic respiratory conditions separately. *
**************************************************************************

get file 'mcp.sav'
sort cases by hh subj
select if spec eq 2   /Only select specialist visits

temporary
select if (diagcode eq 410 or diagcode eq 411 or diagcode eq 412 or
        diagcode eq 413 or diagcode eq 414 or diagcode eq 4140 or
        diagcode eq 4141 or diagcode eq 4148 or diagcode eq 4149 or
        diagcode eq 431 or diagcode eq 434 or diagcode eq 4340 or
        diagcode eq 4341 or diagcode eq 4349 or diagcode eq 436)

aggregate outfile=strsp.sav
        /break=hh subj
        /sp1 'Number of SP visits for cardiovascular disease' = sum(util4yr)

temporary
select if (diagcode eq 295 or diagcode eq 2950 or diagcode eq 2951 or
        diagcode eq 2952 or diagcode eq 2953 or diagcode eq 2954 or
        diagcode eq 2955 or diagcode eq 2956 or diagcode eq 2957 or
        diagcode eq 2958 or diagcode eq 2959 or diagcode eq 296 or
        diagcode eq 2960 or diagcode eq 2961 or diagcode eq 2962 or
        diagcode eq 2963 or diagcode eq 2964 or diagcode eq 2965 or
        diagcode eq 2966 or diagcode eq 2968 or diagcode eq 2969 or
        diagcode eq 297  or diagcode eq 2970 or diagcode eq 2971 or
        diagcode eq 2972 or diagcode eq 2973 or diagcode eq 2978 or
        diagcode eq 2979 or diagcode eq 298  or diagcode eq 2980 or
```

```
        diagcode eq 2981 or diagcode eq 2982 or diagcode eq 2983 or
        diagcode eq 2984 or diagcode eq 2988 or diagcode eq 2989 or
        diagcode eq 303  or diagcode eq 309 or diagcode eq 3090 or
        diagcode eq 3091 or diagcode eq 2992 or diagcode eq 3093 or
        diagcode eq 3094 or diagcode eq 3098 or diagcode eq 3099)

aggregate outfile=mensp.sav
        /break=hh subj
        /sp2 'Number of SP visits for Mental Disease' = sum(util4yr)

temporary
select if (diagcode eq 490 or diagcode eq 491 or diagcode eq 4910 or
        diagcode eq 4911 or diagcode eq 4912 or diagcode eq 4918 or
        diagcode eq 4919 or diagcode eq 492 or diagcode eq 493 or
        diagcode eq 4930 or diagcode eq 4931 or diagcode eq 4939 or
        diagcode eq 494 or diagcode eq 496)

aggregate outfile=ressp.sav
        /break=hh subj
        /sp3 'Number of SP visits for Respiratory conditions'=sum(util4yr)

match files file=strsp.sav
        /file=mensp.sav
        /file=ressp.sav/by hh subj
save outfile=dissp.sav

get file 'dissp.sav'
display sorted names/vars=all
freq sp1 sp2 sp3
save outfile=dissp.sav

match files file=dishos.sav
        /file=dislos.sav
        /file=disgp.sav
        /file=dissp.sav/by hh subj
save outfile=distot.sav

* The above is to calculate the total specialist visits for cardiovascular disease, mental
disease, and chronic respiratory conditions separately. *
**********************************************************************
get file 'distot.sav'
display sorted names/vars=all
```

```
get file 'panel.sav'
   /keep=hh subj agegp6 sex marital hstat hstatgp2 cc1 ccgp4 disab
        sescore1 age urban incad5 educgp5
match files file=*
        /file=distot.sav/by hh subj
save outfile=disampl.sav
get file 'disampl.sav'
display sorted names/vars=all
```

* The above part is to get file disampl.sav which combines GP visits, specialist visits, and hospitalizations together. *

```
get file 'disampl.sav'
recode hosp1 hosp2 hosp3 gp1 gp2 gp3 sp1 sp2 sp3 (missing=0)

temporary
select if hosp1 ne 0 or gp1 ne 0 or sp1 ne 0
save outfile=stroke.sav

temporary
select if hosp2 ne 0 or gp2 ne 0 or sp2 ne 0
save outfile=mental.sav

temporary
select if hosp3 ne 0 or gp3 ne 0 or sp3 ne 0
save outfile=respir.sav
finish
```

* The above is to get a file containing GP visits, specialist visits, and hospitalization for cardiovascular disease, mental disease, and chronic respiratory conditions separately. *

****

Program-7.sps

Goal: To get nine data files of hospitalizations, GP and SP visits for cardiovascular disease, mental disease and respiratory conditions separately.

****

```
***Explanation of variables: admdate: admit date; disdate: discharge date;
agegp6: age in 6 groups; dx1: ICD9-code; los: length of stay;
SES9: education and income;
urban: residence area ***

get file 'hosp.sav'
  /keep=hh subj admdate disdate agegp6 dx1 los SES9 sex urban

sort cases by hh subj

print format admdate(adate9)
print format disdate(adate9)

temporary
select if (dx1 eq '410' or dx1 eq '411' or dx1 eq '412' or dx1 eq '413' or
 dx1 eq '414' or dx1 eq '4140' or dx1 eq '4141' or dx1 eq '4148' or dx1 eq
 '4149' or dx1 eq '431' or dx1 eq '434' or dx1 eq '4340' or dx1 eq '4341' or
 dx1 eq '4349' or dx1 eq '436')
export outfile=strhos.por

temporary
select if (dx1 eq '295' or dx1 eq '2950' or dx1 eq '2951' or dx1 eq '2952'
 or dx1 eq '2953' or dx1 eq '2954' or dx1 eq '2955' or dx1 eq '2956'
 or dx1 eq '2957' or dx1 eq '2958' or dx1 eq '2959' or dx1 eq '296'
 or dx1 eq '2960' or dx1 eq '2961' or dx1 eq '2962' or dx1 eq '2963'
 or dx1 eq '2964' or dx1 eq '2965' or dx1 eq '2966' or dx1 eq '2968'
 or dx1 eq '2969' or dx1 eq '297'  or dx1 eq '2970' or dx1 eq '2971'
 or dx1 eq '2972' or dx1 eq '2973' or dx1 eq '2978' or dx1 eq '2979'
 or dx1 eq '298'  or dx1 eq '2980' or dx1 eq '2981' or dx1 eq '2982'
 or dx1 eq '2983' or dx1 eq '2984' or dx1 eq '2988' or dx1 eq '2989'
 or dx1 eq '303'  or dx1 eq '309'  or dx1 eq '3090' or dx1 eq '3091'
 or dx1 eq '2992' or dx1 eq '3093' or dx1 eq '3094' or dx1 eq '3098'
 or dx1 eq '3099')
export outfile=menhos.por

temporary
```

```
select if (dx1 eq '490' or dx1 eq '491' or dx1 eq '4910' or dx1 eq '4911' or
       dx1 eq '4912' or dx1 eq '4918' or dx1 eq '4919' or dx1 eq '492' or
       dx1 eq '493' or dx1 eq '4930' or dx1 eq '4931' or dx1 eq '4939' or
       dx1 eq '494' or dx1 eq '496')
export outfile=reshos.por

************************************************************************

get file 'mcp.sav'
  /keep=hh subj agegp6 diagcode servdate SES9 sex spec urban

sort cases by hh subj
select if spec eq 1
print format servdate(adate9)

temporary
select if (diagcode eq 410 or diagcode eq 411 or diagcode eq 412
       or diagcode eq 413 or diagcode eq 414 or diagcode eq 4140
       or diagcode eq 4141 or diagcode eq 4148 or diagcode eq 4149
       or diagcode eq 431 or diagcode eq 434 or diagcode eq 4340
       or diagcode eq 4341 or diagcode eq 4349 or diagcode eq 436)
export outfile=strgp.por

temporary
select if (diagcode eq 295 or diagcode eq 2950 or diagcode eq 2951 or
       diagcode eq 2952 or diagcode eq 2953 or diagcode eq 2954 or
       diagcode eq 2955 or diagcode eq 2956 or diagcode eq 2957 or
       diagcode eq 2958 or diagcode eq 2959 or diagcode eq 296 or
       diagcode eq 2960 or diagcode eq 2961 or diagcode eq 2962 or
       diagcode eq 2963 or diagcode eq 2964 or diagcode eq 2965 or
       diagcode eq 2966 or diagcode eq 2968 or diagcode eq 2969 or
       diagcode eq 297  or diagcode eq 2970 or diagcode eq 2971 or
       diagcode eq 2972 or diagcode eq 2973 or diagcode eq 2978 or
       diagcode eq 2979 or diagcode eq 298  or diagcode eq 2980 or
       diagcode eq 2981 or diagcode eq 2982 or diagcode eq 2983 or
       diagcode eq 2984 or diagcode eq 2988 or diagcode eq 2989 or
       diagcode eq 303  or diagcode eq 309  or diagcode eq 3090 or
       diagcode eq 3091 or diagcode eq 2992 or diagcode eq 3093 or
       diagcode eq 3094 or diagcode eq 3098 or diagcode eq 3099)
export outfile=mengp.por

temporary
```

```
select if (diagcode eq 490 or diagcode eq 491 or diagcode eq 4910 or
        diagcode eq 4911 or diagcode eq 4912 or diagcode eq 4918 or
        diagcode eq 4919 or diagcode eq 492 or diagcode eq 493 or
        diagcode eq 4930 or diagcode eq 4931 or diagcode eq 4939 or
        diagcode eq 494 or diagcode eq 496)
export outfile=resgp.por

get file 'mcp.sav'
sort cases by hh subj
select if spec eq 2
print format servdate(adate9)

temporary
select if (diagcode eq 410 or diagcode eq 411 or diagcode eq 412 or
        diagcode eq 413 or diagcode eq 414 or diagcode eq 4140 or
        diagcode eq 4141 or diagcode eq 4148 or diagcode eq 4149 or
        diagcode eq 431 or diagcode eq 434 or diagcode eq 4340 or
        diagcode eq 4341 or diagcode eq 4349 or diagcode eq 436)
export outfile=strsp.por

temporary
select if (diagcode eq 295 or diagcode eq 2950 or diagcode eq 2951 or
        diagcode eq 2952 or diagcode eq 2953 or diagcode eq 2954 or
        diagcode eq 2955 or diagcode eq 2956 or diagcode eq 2957 or
        diagcode eq 2958 or diagcode eq 2959 or diagcode eq 296 or
        diagcode eq 2960 or diagcode eq 2961 or diagcode eq 2962 or
        diagcode eq 2963 or diagcode eq 2964 or diagcode eq 2965 or
        diagcode eq 2966 or diagcode eq 2968 or diagcode eq 2969 or
        diagcode eq 297  or diagcode eq 2970 or diagcode eq 2971 or
        diagcode eq 2972 or diagcode eq 2973 or diagcode eq 2978 or
        diagcode eq 2979 or diagcode eq 298  or diagcode eq 2980 or
        diagcode eq 2981 or diagcode eq 2982 or diagcode eq 2983 or
        diagcode eq 2984 or diagcode eq 2988 or diagcode eq 2989 or
        diagcode eq 303  or diagcode eq 309 or diagcode eq 3090 or
        diagcode eq 3091 or diagcode eq 2992 or diagcode eq 3093 or
        diagcode eq 3094 or diagcode eq 3098 or diagcode eq 3099)
export outfile=mensp.por

temporary
select if (diagcode eq 490 or diagcode eq 491 or diagcode eq 4910 or
        diagcode eq 4911 or diagcode eq 4912 or diagcode eq 4918 or
        diagcode eq 4919 or diagcode eq 492 or diagcode eq 493 or
```

```
        diagcode eq 4930 or diagcode eq 4931 or diagcode eq 4939 or
        diagcode eq 494 or diagcode eq 496)
export outfile=ressp.por

finish
```

****

Program-8.sas

Goal: To create three SAS files for cardiovascular disease, mental disease and respiratory condition separately. And each file contains the variables like: hh, subj, type(1=gp, 2=sp, 3=hosp),los, date (combine servdate and admdate into one variable), diagnosis.

****

For cardiovascular disease

```
libname xyz '/fs/d0/fs/n74nl';

data xyz.cardio;
set xyz.strhos(keep=hh subj admdate dx1 los urban sex SES9 agegp6) xyz.strgp(keep=hh
subj servdate diagcode spec urban sex SES9 agegp6) xyz.strsp(keep=hh subj servdate
diagcode spec urban sex SES9 agegp6);
format date date7.;

if admdate=. then do;
  date=servdate;
  if spec=1 then type=1;
  if spec=2 then type=2;
end;
else if servdate=. then do;
  type=3;
  date=admdate;
end;
output;
run;

proc sort data=xyz.stroke;
by hh subj;
run;

proc freq data=xyz.stroke;
  tables type;
run;
```

For mental disease

```
data xyz.mental;
set xyz.menhos(keep=hh subj admdate dx1 los urban sex SES9 agegp6)
xyz.mengp(keep=hh subj servdate diagcode spec urban sex SES9 agegp6)
```

```
xyz.mensp(keep=hh subj servdate diagcode spec urban sex SES9 agegp6);

format date date7.;

if admdate=. then do;
  date=servdate;
  if spec=1 then type=1;
  if spec=2 then type=2;
end;
else if servdate=. then do;
  type=3;
  date=admdate;
end;
output;
run;

proc sort data=xyz.mental;
by hh subj;
run;

proc freq data=xyz.mental;
  tables type;
run;
```

For chronic respiratory condition

```
data xyz.respir;
set xyz.reshos(keep=hh subj admdate dx1 los urban sex SES9 agegp6) xyz.resgp(keep=hh
subj servdate diagcode spec urban sex SES9 agegp6) xyz.ressp(keep=hh subj servdate
diagcode spec urban sex SES9 agegp6);

format date date7.;

if admdate=. then do;
  date=servdate;
  if spec=1 then type=1;
  if spec=2 then type=2;
end;
else if servdate=. then do;
  type=3;
  date=admdate;
end;
```

```
output;
run;

proc sort data=xyz.respir;
by hh subj;
run;

proc freq data=xyz.respir;
  tables type;
run;
```