University of Alberta

An Economic Evaluation of Home Care as an Alternative to Institutionalization

by

Konrad Fassbender ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Departments of Public Health Sciences and Economics

Edmonton, Alberta

Fall 1998

0-612-34760-5

University of Alberta

Library Release Form

Name of Author: Konrad Fassbender

Title of Thesis: An Economic Evaluation of Home Care as an Alternative to Institutionalization

Degree: Doctor of Philosophy

Year this Degree Granted: 1998

Konrad Fassbender

Konrad Fassbender
9717 111 Street, Apt.1808
Edmonton, Alberta T5K 2M6

OCT 02, 1998.

Date Submitted to the Faculty of Graduate Studies and Research

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommended to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled An Economic Evaluation of Home Care as an Alternative to Institutionalization submitted by Konrad Fassbender in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Philip Jacobs
P. Jacobs

L.D. Saunders

R.H.M. Plain

L. Laing
L.M. Laing

Peter Calder
P. Calder

L.J. Wei

Oct 2, 1998
Date Approved by Committee

ABSTRACT

INTRODUCTION: Over the last decade, home care expenditures have increased by more than 10% per year. Governments generally accept that home care is cost effective with relatively little formal analysis of the economic consequences of this paradigm shift.

OBJECTIVES: The objective was to identify and measure determinants of home care utilization incorporating service intensity and duration. The role of formal and informal service intensity in delaying institutionalization (nursing home admission) was explored.

FRAMEWORK: A modified Andersen predisposing-enabling-need framework was used to identify variables for a novel econometric method combining competing Cox Proportional Hazards (CPH) modelling with a two-stage procedure to correct self-selection bias.

DATA SOURCE / STUDY SETTING: An Alberta health authority provided anonymous, individual administrative data (May 1991 to Dec 1995) on client demographics, functional ability, provider characteristics and detailed service records for 4,962 individuals.

STUDY DESIGN: Econometric analysis of longitudinal data was performed. Formal service intensity was regressed on health care system variables and informal service intensity was regressed on environmental variables using ordinary least squares. Length of stay in the home care program was regressed (using CPH) on observed plus predicted formal and informal home care intensity controlling for predisposing and need variables.

RESULTS: Language, marital status, family size and income did not predict levels of informal service. Living alone was however a significant environmental variable. Health care system variables did not help to predict formal service intensity. This model provided evidence that formal and informal services are complementary goods. Age, gender, physical function, mental function, and comorbidity significantly predicted time spent in home care. Although formal and informal service intensity did not predict home care length of stay, correcting for self-selection bias resulted in higher significance levels. CONCLUSIONS: The Andersen framework identified determinants of home care utilization. Economic consequences of home care services were successfully quantified with a novel econometric model applied to an administrative database. Full correction of self-selection bias however requires better data. As a result of the methodological limitations and data quality issues, conclusions regarding the relationship between enabling resources and length of stay in a home care program would be premature.

I dedicate this book to my parents, Konrad and Emma, who instilled in me an appreciation and reverence for the value of education.

"What you eat and what you learn, nobody can take from you."

Acknowledgement

I thank the Capital Health Authority for providing me access to the Home Care database. In particular, I express gratitude to Lori Lindquist for providing technical expertise and to Kathleen Gubbins for her initiative and knowledge of the home care dataset. I appreciate the generous support that the Department of Public Health Sciences and the Institute of Pharmaco-Economics have provided me this past year. I express my deepest gratitude to Philip Jacobs for his trust, advice and mentorship. Above all, I would like to express my heartfelt admiration for endless patience and understanding of my wife, Regina Michetti.

Table of Contents

List of Tables

List of Figures

I. Introduction

Continuing care expenditures are the third largest in the Canadian health care sector behind acute-care hospitals and physician services (Hollander 1994). Public home care expenditures totalled $1,815 M. in 1995/96 (Health Canada 1998). In the last decade public home care expenditures have more than doubled, averaging annual rates of increase exceeding 10%. Although spending as a percentage of total public health expenditures was only 3.45% in 1995/96, it has increased from 1.51% one decade ago and from 0.65% two decades ago. In the province of Alberta, expenditures in 1995/96 were $132 M or 2.64% of total public health spending. It becomes apparent that home care expenditures continue to increase in an era of health care reform characterized by fiscal curtailment in other areas of health care spending.

Despite the current emphasis on curtailment of other public health expenditures, there is little literature in the Canadian context that addresses the economic implications of home or continuing care. The primary reason for this observation stems from the historical context out of which continuing care evolved and was financed. Secondary reasons owe primarily to the inherent complexity of the problem. As a result of this complexity, ambiguity in the literature is also evident in the US context as evidenced by the inconclusiveness of results stemming from the National Long Term Care Channeling and other demonstration projects.

In Canada, the Hospital and Diagnostic Services Act of 1957 and the Medical Care Act of 1966 prompted the provinces to provide comprehensive hospital and medical coverage for all individuals. Following the introduction of this legislation, the federal government matched provincial expenditures on hospital and medical care. However, health care expenditure inflation caused the federal government to transfer responsibility for health care expenditures to the provinces with passage of the Established Programs Financing initiative in 1977. Now the provinces had an incentive to allocate funds to health care sectors other than the hospitals and medical care. The provinces *prima facia* accepted the idea that community based provision of services was cost effective, and analyses of the economic implications of this paradigm shift were not forthcoming.

The inherent complexity of providing and financing continuing care services formed the second reason for the absence of substantive research in this field. Each jurisdiction implemented continuing care differently. Even the principle of portability outlined in the Canada Health Act of 1984[1], did not apply to continuing care due to the exclusion of non-insured services. Analyses of continuing care and/or any of its components should first require a description of the program. The following section provides a description of the program to be analysed in this dissertation.

As an extension to the issue of definition is nomenclature: continuing care, long-term care, and home care. Even terms like 'comprehensive care' or 'continuum of care' (Fisk 1983) and 'community care' (Grundy 1987) encompass a set of services and patient population which overlap services included in other classifications. These terms overlap

[1]The Canadian health care system was characterised by five principles in the Canada Health Act of 1984: accessibility, comprehensiveness, portability, universality and public administration. Portability refers to the fact that an individual should expect equal treatment in each of Canada's provinces.

and their meanings are often jurisdiction specific. As a result, some exposition is devoted to differentiating among these terms and to further delineating the different interpretations and meanings of home care itself.

This dissertation identifies and evaluates the determinants of home care as an alternative to institutional care (skilled nursing facilities or nursing homes). As a general statement, this means the consideration of all factors that are relevant to the delaying of institutionalisation. This includes the roles of the patients themselves, their caregivers, the physicians, the case managers, and the type and mode of service delivery. Specifically however, the primary objective of the dissertation is to evaluate the relationship between the intensity (or level) of home care expenditures and the utilisation (or quantity) of home care services consumed. Utilisation of home care is measured by the length in time that a patient remains free of institutional care.

A. Edmonton's Home Care Program

This section will briefly outline the financing and provision of services and the changes made to the Edmonton Home Care Program. The *Edmonton Board of Health Home Care Program* was established in 1973 through joint administration of the Edmonton Local Board of Health and Edmonton Social Services. Home Care Regulations, a provincial program introduced in 1978 as an appendix of the Health Unit Act, resulted in a commensurate increase in the funding of medical services. The provincial regulations allowed each local health unit to select its means of delivering home care. The total budget was $310,896.

The years 1978-1982 marked the developmental stage for the provincial co-ordinated Home Care Program now fully funded with provincial moneys. Funding was increased in each of these four years. In 1980-1983 a pilot project was introduced to assess the provision of support to disabled adults. A review was performed in 1981 resulting in the consolidation into three unit offices. Introduction of clearly defined roles for case co-ordinators emerged from this review.

After a year of no additional funding, the budget was substantially increased in 1984. With this increase, support services for the elderly, expansion of the palliative home care, and support services to physically disabled in designated facilities were provided. Revisions to the Co-ordinated Home Care Program Regulation (1985) expanded the admission criteria from a strictly medical need to a more general requirement for health care or support services. This resulted in a rapid increase in admissions resulting in the first ever-waiting list.

Categorization of client care into Acute, Chronic, and Palliative Care took place in 1986 with a budget that year of $6,673,158. This year also marked the computerisation of the home care information system database (HCIS).

Transfer of professional services from contracted providers to in-house staff took place in 1987. Case management was defined for the first time in the following year. In 1990, The home care program became the single point of entry for long term care services including nursing homes and auxiliary hospitals. The emphasis at this time was targeting of service expansion to those at high risk of institutionalization. In 1992 clients under the age of 65 were able to receive services. Finally, on April 1, 1995 management of the program was transferred from the Edmonton Board of Health to the Capital Health Authority.

Health services provided to home care clients without charge are home nursing services, occupational therapy, physiotherapy, respiratory therapy and speech therapy. Additional services may include: dressing, medication and other related preparations, and temporary use of health aids not provided under the Alberta Aids to Daily Living and Extended Health Benefits Program. Personal care and home support services are provided without charge for the first two weeks on the program. Home support may include heavy housework service, handyman services, meals-on-wheels, wheels-to-meals, and transportation services.

Currently, a $3,000/month service limit applies unless the individual is waiting admission to a long-term care facility, or the high service needs are required for less than three months. The Home Care Program does not provide 24-hour professional services on an ongoing basis.

Clients do not pay for health or personal care services. The fee for home support however is limited to $5 per hour to a maximum of $300/month. The monthly charge is determined on a sliding scale, which takes both family income and family size into consideration. Low-income clients qualifying under the following four alternative programs are exempt from payment: (1) SFI - Supports for Independence, (2) Widow's

Pension, (3) GIS - Guaranteed Income Supplement, or (4) AISH - Assured Income for the Severely Handicapped.

In Edmonton, a uniform assessment form is used as required criteria for admission: the Alberta Assessment and Placement Instrument for long-term care (AAPI, McKenzie et al 1989). In summary a client is admitted if: a health condition is present which limits independent function, the client requires and wishes to receive care at home, the home is suitable for providing care, home care is the most suitable method of care, adequate resources are available, and the client is eligible for Alberta Health Care Insurance.

B. Continuing Care, Long-term Care, and Home Care

Continuing care, long-term care and home care comprise a few of the terms that are encountered in both practise and the literature. Their use and misuse has contributed to the difficulty in analysis of home care. Although federal and provincial governments have attempted to define these terms, and even though others accept these definitions, an incomplete definition can only result in an incomplete analysis. A conceptual framework of these terms is developed through contrast with and exclusion of episodic care. Definition of these terms is accomplished through reference to services, location, financing, provision, and decision making. This allows for an unequivocal definition amenable to economic analysis.

These terms are introduced and illustrated in Figures 1.1 and 1.2. Each node describes a type of patient and type of service. Demand for health care services are classified as either episodic or continuing or long-term care as determined by the disease, patient and physician. Continuing or long-term care is differentiated from episodic care by location and can be provided institutionally or non-institutionally (Scanlon 1984). In turn, non-institutional care can be financed formally or informally. Lastly, formal provision of services can be centralised or decentralised. It is this latter point that most often results in confusion as many home care programs offer an increasing number of services outside the home (Gillick 1989, Kane 1995).

Figure 1.2 illustrates the decision-makers or characteristics that differentiate each patient - service combination. Although illustrated as 'Decisions 1-4', the diagram is not meant to portray a sequential process of decision making. At decision point 3 for example, informal caregivers help to determine whether a patient requiring long-term care will receive formal care. The purpose of these diagrams is to help clarify the terminology so that each node describes a unique health care service - patient combination. As discussed in this section a given term, can encompass multiple patient-service combinations or nodes in these two diagrams. An important point to remember is that although patient characteristics influence the decision as to which services and setting are chosen, the patient may not have have the final word in the decision (Sherwood et al 1986).

1. Continuing Care

Continuing care has been defined widely in Canada as encompassing both the service and the system of service delivery itself. A national/regional-sponsored committee has recently embodied this definition as follows:

> Continuing-care is multifaceted and combines aspects of both health and social services. Unlike hospital care or physician services, varied as they may be, continuing care is an amalgamation of diverse categories of service. These different categories of service are integrated by an overall "system" of service delivery. Thus ... it is important to remember that *continuing care is not a type of service, but a system of service delivery.* The efficiency and effectiveness of that system is based not only on its constituent parts, but also on the nature of the system itself.[2]

This definition, however, only partially alleviates the misunderstanding of the term. Based on an historical paradigm, the report implicitly defines continuing care through contrast with hospital care and physician services. However, home care may require intermittent hospital and physician care. Continuing care may include informal caregiver services and a community based provision such as seniors' drop-in centres.

Continuing care is also characterised along four dimensions: service, location, financing, and provision. A full description of these four attributes leads to the characterisation of "products" that are distinct and can be analysed with traditional and health economic methods. To characterise continuing care through emphasis on service

---

[2]Federal/Provincial/Territorial Subcommittee on Continuing Care, *Future Directions in Continuing Care,* Ottawa: Health and Welfare Canada, September 1992, pp. 2-3.

delivery constrains analysis to program evaluation. Evaluation of program or service delivery is typically descriptive and does not lend itself to analytical discourse.

Service refers to a description of the constituent activities that comprise a program. They may include items such as home nursing, home support or housekeeping, rehabilitative services, and physician services. Financing refers to the mode of compensation for the services described. They include ministry of health, private insurance, out-of-pocket, volunteer and caregiver support. Continuing care may be administered at various sites: home, community, hospital, seniors' lodges or seniors' residences. Finally, provision of services or service delivery differentiates the decentralised case management and co-ordination that has taken place historically from the co-ordinated, community based models that are in place today. Together these four dimensions can be used to delineate the other subsets of continuing care: long-term care and home care.

2. Long-term Care

Long-term care has several different meanings depending on where and when it has been used. For example, its meaning can be synonymous with continuing care as defined by another national/regional committee:

> Long-term care represents a range of services that addresses the health, social and personal care needs of individuals who, for one reason or another, have developed or have lost some capacity for self-care. Services may be continuous or intermittent, but is generally presumed that they will

> be delivered for the 'long term' that is, indefinitely to individuals who have demonstrated need, usually by some index of functional incapacity.[3]

This definition does not contain any limitations with respect to service, financing, site or provision. In fact, the definition is slightly more broad than that developed for continuing care in that intermittent care can be provided with the proviso that the services are expected to resume in the event of a temporary "improvement" in health status.

Whereas the commonly accepted concept of continuing care was based on provision, long-term care is historically associated with the site or location of services. In fact, the Canadian historical usage is synonymous with institutionalised care. As alluded to earlier, prior to incentives existing for the provision of community based provision of home care services, the majority of continuing care was provided in residential facilities also known as nursing homes.

Long-term care has yet a third meaning. It is used in Edmonton (and possibly other jurisdictions) to delineate home care provided to a subset of patients. In this context, long term care defines:

> Individuals who are expected to require Health Unit services on a continuing basis for greater than three months to gradually improve or maintain health status, functional status, level of independence, or to delay deterioration.[4]

---

[3]Subcommittee on Institutional Program Guidelines, *Assessment and Placement for Adult Long-Term Care: A Single-Entry Model*, Ottawa Health and Welfare Canada, 1988, p. 2.

[4]Home Care Information System, Data Standards Manual, Alberta Health, October 6, 1994.

A primary distinction between both continuing plus long-term care and episodic care as the primary driver of health care consumption is made in the context of expectations. Episodic care can be characterised by an expectation of improvement as a result or a limited duration of treatment. Continuing and long-term care, by default, is characterised by the absence of these expectations and in particular is directed toward individuals whose health is expected to deteriorate despite medical intervention.

3. Home Care

The conceptual model (Figure 1.1) defines home care as formal or informal services which take place in a non-institutional, decentralized setting. The formal provision of services can take place in a decentralised fashion whereby several individuals are involved in referral, intake, and provision of services. This node is characterised the Edmonton Home Care Program in its early years (1973-1982). In addition, provision of services can take place in a case-managed, community-based setting, as is currently the case. For the purposes of the dissertation, home care consists of formal and informal services.

The centralised provision of services takes place in the community such as adult day care as an alternative to institutionalization (Arling et al 1984). Therefore, home care *per se* can have two meanings. Firstly, home care can refer to the provision of services that enable the patient to remain at home. In this case services are provided in either the home or in a community-based setting. Secondly, home care can refer to services that are provided in the home (Council on Scientific Affairs, 1990). Although this interpretation is narrower, it is the one utilized in the dissertation due to nature of the Edmonton Home Care Program and data availability. Finally, informal, non-institutional services are included in home care and include services provided by caregivers and volunteers.

As with long-term care, the exact meaning of home care will vary from jurisdiction to jurisdiction. In Canada however, a national/regional committee has delineated three models of home care: the maintenance and preventative model, the long-term care substitution model, and the acute care substitution model.[5] In addition, as is the case in Edmonton and the conceptual model, the palliative care model falls under home care.

a) Acute Care Substitution

Home care in this context substitutes for acute care services. It refers to the provision of services in the home following early discharge from an acute care facility. Strictly speaking, these services do not qualify as home care since the nature of the illness is episodic as opposed to long-term. Inclusion of these services into the conceptual definition of home care would require episodic care to be divided into institutional vs non-institutional etc., or to subdivide institutional care into residential versus acute care facilities etc..

[5]Federal/Provincial/Territorial Subcommittee on Long Term Care, *Report on Home Care,*

b) Maintenance and Prevention

This model of home care describes services that enable the individual to remain at home or to maintain independence.

c) Long-term Care Substitution

This model of home care describes services that enable the individual to remain at home and out of an institution. The functional impairment of this group of individuals is greater than that which characterizes the maintenance and prevention group.

d) Palliation

This model describes palliative services that are typically provided in the home with the goal of allowing people to die with dignity.

Ottawa: Health and Welfare Canada, 1990, pp. v,1.

C. Home Care as an Alternative to Institutional Care

This dissertation is restricted to the consideration of home care as an alternative to institutional or nursing home care. As the demographic trends shift toward aging of the population, a higher percentage of the population will become high risk for institutionalisation. Hence, the importance and validity of the general problem will increase over time.

The potential contribution of the study is that both patient characteristics and program characteristics are simultaneously considered as determinants of home care utilisation. Patient characteristics are important in that they define the risk of institutionalisation in the absence of home care service provision. Patient and program characteristics are important in that they describe the range of services that are available and consistent with avoiding institutionalisation.

D. Home Care Expenditures Delay Onset of Institutional Care

The specific problem refers to the allocation of financing of formal services and the impact on utilisation of home care as measured by time spent out of an institution. An absence of clear findings in other studies results from the inadequate description of services and imprecise measures of utilisation. For example, the National Long Term Care Demonstration study merely considered whether an individual received case management or additional financial services. Since case management is a small percentage of the services provided to the typical home care case, a significant finding would have been surprising. Furthermore, additional home care services could not be practically withheld from the control group.

The rationale for this study arises from an absence of a clearly defined conceptual framework, widespread use of conflicting definitions, inadequate description of services, and weak analytical methodology. Testing for a significant association between the allocation of home care resources and utilisation of home care services without the limitations imposed by previous studies will advance knowledge in the are of home care.

It is hypothesised that individuals at risk for institutionalisation can remain at home longer with a greater intensity of service provision. The resulting null hypothesis is that expenditure on home care services has no effect on the length of time spent at home out of an institution. This hypothesis is the primary question that the dissertation attempts to answer.

E. Andersen's Behavioural Model of Health Services Utilization

Andersen's Behavioural Model (1968, 1973) is presented in Chapter II and is the predominant model in the literature used to explain utilisation of health services. Societal, health services system, and individual determinants interact to explain utilisation. To date, models developed in the literature are restricted to the estimation of individual determinants. While this corresponds to the focus of the dissertation (the consumer sovereignty assumption), the implications of incorporating societal and health system determinants will be briefly examined.

A simple model is derived on the basis of this framework. The behavioural model of utilisation is a function of predisposing, enabling and need variables. Predisposing variables include age, sex, marital status and family structure. Enabling variables include income, insurance and home care program expenditures. Finally, need variables capture the individual perception of difference between actual and desired health status and include diagnosis and HCFN (home care functional need score). This framework encompasses both individual and societal influences and resembles models based on alternate paradigms (eg. economic) as discussed in the next chapter.

F. Review of Literature

A brief literature review is provided in Chapter III. Each section will include a critique and evaluation of the validity of theory and/or empirical methods. Also, each section will illustrate the relationship within each article to conceptual and theoretical concepts introduced in Chapter II, and will show how the existing literature points to need for this dissertation. Using this approach the experimental, quasi-experimental and non-experimental literature are examined.

G. Statistical Methodology, Description and Analysis of Data

Chapter IV states the research problem and provides the tools necessary to solve the problem. To begin with, all of the variables used in the study are defined. The variables are grouped according to the theoretical concepts introduced earlier. Kaplan-Meier and Cox Proportional Hazard modelling is introduced. As well, all of the auxiliary models, variable construction methods, and modelling assumptions are detailed.

Limitations and delimitations of the methodology are listed and discussed. Specifically, validity of administrative data, definition of an episode of long-term care, unit of measurement of utilization, valuation of utilization, competing hazards and self selection bias due to unobservable variables. Finally, procedures for the collection and manipulation of databases are presented. This section also discusses the sampling methodology and selection criteria.

Chapter V describes the description, analysis and interpretation of data results. This includes presentation of descriptive statistics and regression results. Econometric testing of underlying assumptions is performed as required. A brief description of the issues accompanies each section.

H. Summary and Conclusions

Finally, in Chapter VI, generalization from research results to theoretical framework takes place. Serendipitous findings and health policy implications of research results are discussed. Limitations of the study are reviewed. And finally, implications for further research and practice are introduced.

Figure 1.1: Conceptual Framework for the Characterization of Home Care Services

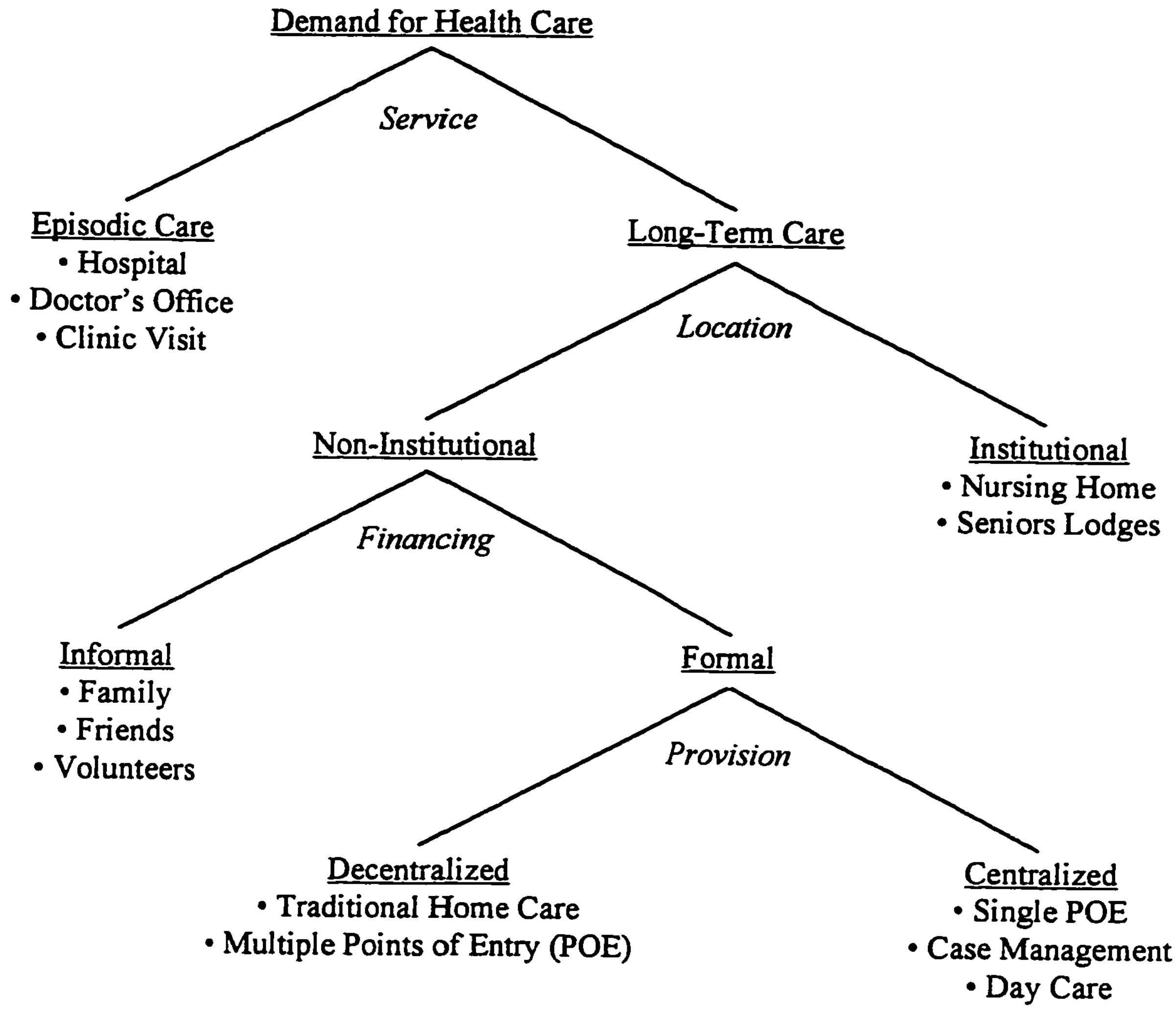
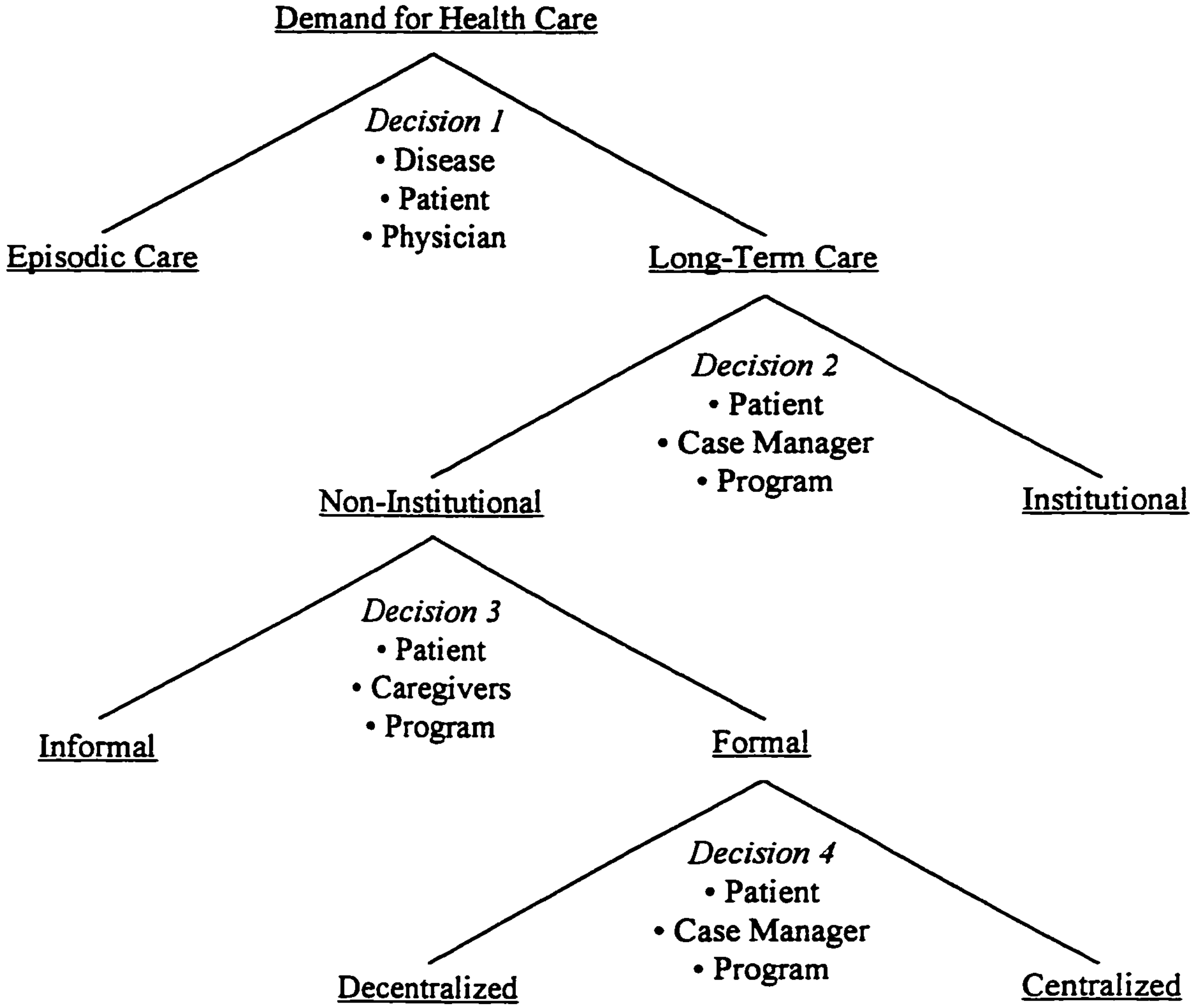


Figure 1.2: Decisions and Characteristics Influencing Use of Home Care Services



Bibliography

Andersen, R.M. (1968) Behavioral Model of Families' Use of Health Services. Research Series No. 25. Chicago, IL.: Center for Health Administration Studies, University of Chicago.

Andersen, R.M. and Newman, J.F. (1973): Societal and Individual Determinants of Medical Care Utilization in the United States. Milbank Memorial Fund Quarterly Journal, 51:95-124.

Arling, G., Harkins, E., B., and Romaniuk, M. (1984): Adult day care and the nursing home: The appropriateness of care in alternative settings. *Research on Aging*, 6:225-242.

Council on Scientific Affairs (1990): Home care in the 1990s. *JAMA*, 263:1241-1244.

Fisk, A.A. (1983): Comprehensive health care for the elderly. *JAMA*, 249:230-236.

Gillick, M.R. (1989): Long-term care options for the frail elderly. *Journal of the American Geriatrics Society*, 37:1198-1203.

Grundy, E. (1994): Community care for the elderly 1976-84. *British Medical Journal Clinical Research Ed*, 1987:626-629.

Health Canada. Pubic Home Care Expenditures in Canada, 1975-76 to 1997-98. March 9, 1998.

Hollander MJ. The Costs, and Cost Effectiveness of Continuing-Care Services in Canada. Queen's - University of Ottawa Economic Projects #94-10, June 1994, page 1.

Kane, R.A. (1995): Expanding the home care concept: blurring distinctions among home care, institutional care, and other long-term-care services. *Milbank Quarterly*, 73:161-186.

McKenzie DA, Capuzzi CF, Will SJ. Alberta assessment and placement instrument: description and interrater reliability. *Medical Care* 1989;27(10):937-941.

Scanlon, W.J. and Feder, J. (1984): The long-term care marketplace: an overview. *Healthcare Financial Management*, 38:18-19.

Sherwood, S., Morris, J.N., and Ruchlin, H.S. (1986): Alternative paths to long-term care: nursing home, geriatric day hospital, senior center, and domiciliary care options. *American Journal of Public Health*, 76:38-44.

II. Theoretical Framework

Selecting a theoretical framework for utilization of home care services first requires a discussion of utilization itself. Utilization is a multidimensional concept and can be measured using different methods. Identifying determinants of utilization is a multidisciplinary undertaking that requires a brief review of the advantages and disadvantages of economic, sociological, psychological, and pure mathematical methods. On the other hand, behavioral models comprise methods and variables from each of these disciplines to form a unified framework.

A literature review of behavioral models reveals that the Andersen model is the most comprehensive and dominant framework in the literature for utilization modelling. Historical developments of the Andersen model are reviewed and reflect a model that is continuously changing in the health services research literature. Its primary advantage is its flexibility to the problem addressed.

Although the Andersen model dominates the literature, it becomes apparent that there is a relationship to concepts elucidated by others such as Donabedian (1974). The framework originally espoused by Andersen (1968, 1973) also encompasses the traditional economic analysis in that it resembles the reduced form equations resulting from solving for equilibrium between demand and supply functions. Parallels between these various concepts will drawn later in this chapter.

The last sections of this chapter outline the essential elements of the Andersen Behavioral model that relate to utilization of health care services. In general, this section addresses the types of predictions that can be made by the Andersen model. This is a preliminary step to the derivation of the specific hypotheses (later in Chapter 4) regarding home care service utilization.

A. Definition of Utilization

Utilization of health services refers to the identification and measurement of determinants, variations, or trends in the consumption of health-directed goods or services. Before the phenomenon of health care utilization can be measured, it must be defined. This first requires addressing the definitions and dimensions of "health", "health services" and "utilization." Utilization must be discussed in relation to the purpose of a study, units of analysis, and the role of explanatory variables.

Health may be defined variously as state of well being or level of physical and mental functioning. Alternative definitions abound and further distinction requires philosophical discussions of existence and epistemology. Services covers both physical goods and "services." Services may be cognitive in nature (eg. counseling) or physical (eg. massage). An individual or a group of individuals may provide them. Lastly, they may relate to prevention, diagnosis, treatment, or palliation.

Health may be perceived or evaluated. That is to say that health may be subjectively assessed by some form of introspection or objectively assessed by a health care practitioner. This has implications for measurement of health care services since perceptions and evaluations may be of the same phenomena but may not be in agreement.

For example, a doctor and patient may not agree on concepts of care seeking, compliance, or technical aspects of services. However, subjective measures may be better explained by individual behavior while objective measures may be better explained by technical variables. Furthermore, it may be precisely the difference between perceived and evaluated health that may drive utilization of health care services.

Health services refer to those physical goods or services that are directed toward human health. A good or service may vary by individual, time or place that a good was provided. For example, physicians vary by experience and the cognitive services they provide are expected to vary in quality. Likewise, time and place of provision are important attributes associated with a good or service. Lastly, goods and services may be bundled with non-health goods or services. For example, a child receiving treatment for cancer in a hospital may watch as many movies as (s)he wants. A hospital stay may/would then be said to be comprised of both health and entertainment components.

In general, there is a consensus as to what constitutes a health service. However, there are exceptions that must be dealt with separately as the need arises. Suppose that a treatment, or counseling by a doctor resulted in increased food expenditures for an individual. Eating healthy foods requires greater cash outlays and these differences should technically be considered as a health service. Non-traditional medications are another example where they may be purchased with the intention of increasing or maintaining health but have not been proven to do so by scientific means. Cosmetics as well may have a health component. For example, looking better may result in increased self-esteem in turn leading to increased immune function and so forth. However, the practicality of separating and measuring these health services is limited.

Utilization refers to the use (or consumption) of a service or good. In terms of economic theory, it refers to the equilibrium quantity of that good or service supplied and consumed. In terms of the Behavioral model, it simply refers to the amount desired and consumed subject to external conditions (social constraints for example). The purpose of measuring utilization is primarily to explain or predict. These dimensions of utilization determine the method of study but are less likely to affect the measure of utilization. The measure of utilization is the dependent variable that requires a unit of analysis. Generally speaking, the unit of analysis is either expressed as expenditure or physical units. Expenditure data is not ideal since the price of health care service does not equal its marginal cost and decisions made on this basis have undesirable welfare implications. Physical measurements however are typically incomplete. A doctor's office visit, for example, does not address service intensity. Service intensity is necessary dimension and a refinement as to the quantity of resources consumed. For example, the number of diagnostic tests ordered is an appropriate intensity variable for explaining physician visits.

Explanation and prediction of resources require identification and measurement of determinants, variations, and/or trends of health services. In other words, selection of independent variables is a logical starting point. Determinants and variations also have implications for the type or purpose of the study, choice and unit of measurement of the dependent variable. Choice of explanatory factors is a tradeoff between hypothesized factors or variables and the constraint of data availability. A trend of health service use is, by definition, a function of time. This does not preclude the inclusion of other explanatory factors that may or may not themselves vary with time. Appropriate choice of independent variables is covered in the next sections of this chapter.

Identification and measurement of determinants is a preliminary and necessary step in health services research studies. In health services research, study objectives include the determination of availability, appropriateness, efficacy, effectiveness, efficiency, and equity. Availability asks the question whether a service is reaching those who need it? Appropriateness asks whether a service should be provided. Efficacy asks whether is can work while effective use of resources refers to whether is does work. Effectiveness incorporates physician and patient compliance plus other confounding factors in the consideration of non-experimental data as opposed to data arising from clinical trials. Efficient use of resources occurs when utilization is considered to be worthwhile. Lastly, equitable use of resources can only be determined by judgement and refers to the relative levels of utilization of one individual or group of individuals to another.

Each of these study objectives can be achieved using an experimental or non-experimental design. Experimental designs include randomized clinical trials and typically do not require adjustment for confounding or independent variables. On the other hand, non-experimental studies are performed when experimentation is expensive, unethical or impractical. In these cases the utilization studies are needed to identify the necessary confounding variables.

The precise definition of health care utilization, purpose of analysis, units of analysis, and choice of explanatory factors have important implications for the conclusions that arise from reviewing utilization of health care services. Even use of a resource may be described with the neutral terms "use" or "utilization," or with economic terms like demand, consumption or supply of health services. These and other aspects of utilization differ according to the discipline(s) and perspectives associated with study authors and are discussed in the next sub-section.

B. Measurement of Utilization

There are numerous approaches to modelling utilization of health services. Models are constructed according to the principles of the predominant disciplines explaining human behavior: economics, sociology, psychology, and statistics/mathematics. Although the goals of each discipline are the same, the approaches emphasize different elements of the analysis. A theoretical framework should clearly state (1) variables consistent with a theory of utilization behavior, and (2) the mathematical relationships between those variables.

Economic models tend to emphasize both the variables and their statistical relationships. However, the traditional neoclassical economic framework is inconsistent with behavior in the market for health services. As summarized in Culyer's (1971) article on "health care" as a commodity, health care differs from the neoclassical economic goods by virtue of consumer irrationality, uncertainty, and externalities. The informational asymmetries between providers and consumers of health care, the uncertainty of illness incidence and externalities can be summarized a market failure. The predominant economic model that attempts to deal with the market failure is primarily limited to the inclusion of health care insurance. Other economic models typically include the market for hospital services, reimbursement incentives, and physician induced demand. Economic models are however not sufficiently developed to deal with the complex nature of providing home care services to the elderly.

Sociological and psychological models of health services utilization on the other hand primarily focus on variables consistent with individual, social, plus social interaction behavior. As a result, these models tend to include more variables. A tradeoff emerges in the depth versus breadth of variables included in the various models. Models that cover fewer constructs or categories of variables typically examine more variables within that category.

Finally, pure mathematical models emphasize relationships between variables over the selection of variables. Examples of these models include time series modelling and Markov chains. Inclusion of variables in these models is ad hoc and of secondary importance. Selected economic models are characterized by sophisticated econometric techniques at the expense of relevant explanatory variables. These models are in themselves not of much interest and will not be considered in this dissertation.

One important consideration in favor of psychosocial over economic and models deals with the availability of data. Strictly speaking, economic models invariably require price and income observations. As a result, implementation of economic models results in the use of proxy variables instead of prices. An example of one of the few studies that used true prices was the RAND Health Insurance Experiment (Newhouse 1981). Of course, the reason for the lack of true price variables (or opportunity costs) stems from market failure, the same limitation as the theory. As alluded to earlier, behavior in the absence of price signals is not well developed in the health economic literature. The provision of health services is better described by a complete lack of a market and therefore requires more development in the behavior of individuals and institutions.

Psychosocial models can also be characterized by an inability to accurately measure the constructs resulting in a lack of data. Because these theories focus on individual characteristics, social influences or social interactions, the data requirements may be quite specific. Analysis of administrative data for example does not lend itself well to answering questions within these frameworks. Pure psychological or sociological models typically require a prospective cohort study design or administration of a health survey. Both designs require the collection of data by a questionnaire.

In summary, there appears to be a tradeoff between theoretical foundations for selection of variables, the extent to which the hypothesized relationship between variables can be portrayed and actual data that is available to analyze a given model. A compromise arises in that behavioral models actually encompass both economic and psycho-social variables. Behavioral models are the result of interdisciplinary collaboration and provide a reasonable linkage between theory and choice of variables. These class of models are also adept at hypothesizing the nature of the statistical relationship between variables. They will be considered in the next section.

C. *Behavioral Models of Health Services Utilization*

In order to choose a conceptual framework that is consistent with the objectives of this dissertation, a review of conceptual models of utilization behavior was performed. An extensive search for models, frameworks, approaches and/or paradigms was carried out in the health services literature.[1] The limitation of this search strategy is that it only identifies those conceptual models that are named. In fact, it is a reality that many of the interdisciplinary studies employ multiple approaches but do not identify or acknowledge the underlying theory.

An electronic search strategy was used to identify 547 utilization articles that made reference to models, paradigms, approaches, or frameworks (see Appendix 2.1.a). Abstraction of titles and abstracts of these articles resulted in the identification of 42 distinct behavioral theories. Using subjective criteria, the list of 42 theories was reduced to capture models that are influential in the explanation of health services utilization. This meant that references to the frameworks had to be made more than once and by separate sets of authors. Self-referenced theories were excluded. Theories that may have been more prevalent in the psychological or sociological literature would also be excluded by this search. Eight behavioral models were selected to provide the bulk of the conceptual frameworks employed in explaining utilization phenomena in the health services literature.

These eight selected conceptual models of health services utilization behavior are listed in Table 2.1 and are tabulated by decade. Model constructs and/or variables

[1] As of November 1996, there were 37,170 (English language) utilization papers identified in the Medline (1966 - present) and HealthSTAR (1975 - present) databases.

typically used in these models were assigned to one of four groups: environmental, sociological, psychological, and outcomes. Environmental variables are those that describe the health care system and external environment. Sociological variables include predisposing characteristics, enabling resources, and need. Psychological variables encompass personal health practices and personal use of health services. Finally, perceived health status, evaluated health status, and consumer satisfaction comprise outcome variables.

The Andersen model is the framework selected for use in this dissertation since it is an evolving theory, it encompasses interdisciplinary approaches, and it dominates the literature over the other behavioral models. The Andersen model was listed for each of the four decades to reflect the substantial evolution that has occurred over the last thirty years. Although models such as the Health Beliefs models have also changed, the changes have not been as significant. Secondly, the Andersen framework is encompassing. The Andersen model includes economic, institutional, psychological and sociological variables. Lastly (using a search strategy outlined in Appendix 2.2.b) the popularity of the Andersen framework has increased over time as illustrated in Figure 2.1. Reference to the framework has doubled over the last decade.[2]

[2] In fact the framework may be more prevalent as the need to acknowledge Andersen may decrease as the behavioral model becomes more generic.

D. The Andersen Framework

As tabulated in Table 2.1, Andersen's Behavioral model or framework has evolved substantially over the years. The Andersen framework consists of a set of variables and the relationships between those variables. As summarized by Andersen himself, Figure 2.2 (1960's) through Figure 2.6 (1990's) demonstrate that this model includes more variables with each passing decade. In addition to additional variables, the model has progressed to include a greater number of causal pathways (designated by arrows). Arrows in these diagrams designate the hypothesized directions of causality. Although these additional pathways require more sophisticated statistical techniques, evidence for their increased use has not been apparent.

Andersen was concerned with utilization as it related to access to medical services. Organization of variables in his models explicitly incorporated this concept. Access was characterized as potential versus realized. In the transition between these two states, both individuals and providers differ in their characteristics. Individual, family, social, or health care system characteristics that predispose or enable an individual that needs health care services are referred to as potential access. Realized access occurs when health care services have actually been consumed. This refers to use of health care services, consumer satisfaction, or health outcomes. Potential access characteristics can further be described as structural or process indicators. Structural indicators include the health care system and environmental variables. Process indicators refer to individual or population characteristics or health behavior.

Outcomes or realized access indicators can be further divided into subjective or objective measures. Subjective measures include consumer satisfaction of perceived health status. These measures vary from individual to individual and include concepts such as convenience, availability, and quality of care. Objective measures include evaluated health status and the measurements of health care usage. As discussed earlier, they include type, site, purpose, and time interval.

Selections of specific variables are driven by data availability. Variables are grouped according to hypothesized mechanisms of generation. That is to say that community, family and personal resources are listed under enabling resources since they comprise the budgetary constraints in individual and aggregate decision making. Environmental characteristics are grouped together since they are exogenous to the system as can be seen in Figure 2.5 (no feedback arrows).

Over the last four decades, these groupings of variables reflected the status quo in health services research. Introduction of consumer satisfaction in the 1970's and health outcome's in the 1980's are two good examples. Causal relationships are emphasized in a cyclical fashion. Introduction of new concepts is followed by attention in the relationships between the variables. This explanation is consistent with more arrows in Figures 2.3 (1970's) and 2.5 (1990's).

Statistical methods and study design should be guided by hypothesized causal relationships between sets of variables. Path analysis and hierarchical regressions refer to the stepwise addition of blocks of variables to explain the contribution of variance to each block of variables. This is particularly true for the predisposing, enabling and need variable models. Alternatively, econometric techniques of systems of equations and tests for endogeneity could be used. However, in practice single equation techniques dominate the literature.

E. The Andersen Framework, Donabedian and Economic Models

One reason for selecting the Andersen framework is that it is a compromise between availability of data and model construct validity. It is worth noting that the compromise may not be a compromise at all when a comparison of Andersen's model is made against two other dominant frameworks. Figure 2.6 compares the analytical frameworks of Andersen (1968, 1973), Donabedian (1974), and various economists (Arrow 1963, Wirick 1966, Feldstein 1966, Grossman 1972, M Feldstein 1973, Fuchs 1978, and Newhouse 1981). Donabedian's (1974) structure corresponds to Andersen's environmental and population characteristics. Process describes health behaviour and the consumption or utilization of health services. Finally, the consumption of these resources leads to an improvement in the health outcomes.

It becomes evident that a resemblance exists between Andersen's model and a reduced form equation resulting from solving a demand-supply equilibrium problem. Consumption of resources and the equilibrium quantity consumed represents the net health behaviour that Andersen describes. Andersen's environmental influences include the health care system variables that are typically found in a supply equation. Likewise, the predisposing-enabling-need paradigm encompasses the demand variables that are typically described in a demand model. Finally the health outcome has certain parallels to the quality paradigm in economics. It should be noted that in economics quality feeds back on the equilibrium quantity.

In conclusion, adoption of the Andersen framework does not necessarily equate to the abandonment of fundamental economic concepts or renunciation of Donabedian model and its simplicity.

F. Predictions Generated by the Andersen Framework

The purpose of modeling utilization of health services is to identify determinants, variations, or trends in the use of health directed resources. In particular the nature of the relationship between the dependent and independent variables is of importance. Figures 2.2 through 2.5 contain arrows that identify the relationships between sets of variables but do not indicate the magnitude or direction of influence.

Utilization studies primarily are used to identify determinants or predictors. Figure 2.5, for example, depicts a bi-directional relationship between health behavior and health outcomes. Further to this hypothesis is the contention that positive health practices improve health outcomes. In the reverse direction, improved health outcomes may reinforce the positive behavior or alleviate the impetus for the health behavior. The specific hypothesis depends on the context (the specific behavior and the specific outcome). In general, researchers measure the direction and significance of these relationships but are not concerned with the magnitude. Identification of determinants may be performed with various statistical models, the most common of which is the logistical model.

Measuring variations in utilization refers to the exercise of determining the relative importance of variables or classes of variables. For example, Andersen's original model as depicted in Figure 2.2, illustrates utilization as a function of predisposing, enabling and need variables. Hierarchical regression analysis is the most common statistical technique. The purpose of which is to determine the relative importance of each group of variables. Results are expressed in terms of percentage contributions of explaining variances in the dependent variable for each group of variables. Finally, these studies rely on the R^2 (explanatory power) statistic to explain the overall variation in utilization explained by the model. These studies compose the bulk of applications of the Andersen model in the literature.

Trends in health services utilization refer to temporal patterns or significance of time variables. Changes in time can be modeled as discrete (structural break) or continuous variable. These models do not imply that time is itself the causative variable, but rather that time is correlated with a change in another variable. For example, using the Andersen model as depicted in Figure 2.3, changes in health care policy such as a decrease in financing may lead to a decrease in health services use over the relevant time period.

The specific model and hypotheses to be tested by this dissertation are developed and presented in Chapter 4. In summary, this study will identify the ability of each variable to predict utilization of home care services. Variations in these variables will be important and time plays an important role as seen later.

Table 2.1: Selected Behavioral Models of Health Services Utilization

	Model	Theoretical Constructs			
		Environment	Social	Individual	Outcome
1960's	**Andersen Model** (Andersen 1968)		<u>Predisposing Characteristics</u> (Demographic, Social Structure, Health Beliefs), <u>Enabling Resources</u> (Personal / Family, Community), <u>Need</u> (Perceived, Evaluated)		
1970's	**Andersen Model** (Andersen, Smedby, and Anderson 1970; Andersen and Newman 1973; Aday and Andersen 1974; Andersen, Kravits and Anderson 1975)	<u>Health Care System</u> (Policy, Resources, Organization)	<u>Predisposing Characteristics</u> (Demographic, Social Structure, Health Beliefs), <u>Enabling Resources</u> (Personal / Family, Community), <u>Need</u> (Perceived, Evaluated)		<u>Consumer Satisfaction</u> (Convenience, Availability, Financing, Provider Characteristics, Quality)
	Health Belief Model (Rosenstock 1974)			Susceptibility, Benefits, barriers, Self-efficacy	
	Social Learning Theory (Bandura 1977)		Social Desirability	Self-Efficacy	Outcome Expectations
1980's	**Andersen Model** (Aday, Andersen, and Fleming 1980; Aday, Andersen, Loevy and Kremer 1985; Fleming and Andersen 1986)	<u>Health Care System</u> (Policy, Resources, Organization), <u>External Environment</u>	<u>Predisposing Characteristics</u> (Demographic, Social Structure, Health Beliefs), <u>Enabling Resources</u> (Personal / Family, Community), <u>Need</u> (Perceived, Evaluated)	<u>Personal Health Practices</u>	<u>Consumer Satisfaction</u> (Convenience, Availability, Financing, Provider Characteristics, Quality), <u>Perceived Health Status</u>, <u>Evaluated Health Status</u>
	Theory of Reasoned Action (Ajzen and Fishbein 1980)		External Social Influences (Subjective Norms)	Personal Attitudes	
	Health Promotion Model (Pender 1982)			Perceived Control of Health, Definition of Health	
	Theory of Planned Behavior (Ajzen 1985)		Subjective Norms	Attitudes, Percieved Behavioural Control	
	Self-Care Theory (Orem 1985)	Social Economic Status, Education		Motivation	
	Social Cognitive Theory (Bandura 1986)			Self-efficacy	Outcome Expectations
1990's	**Andersen Model** (Andersen, Marcus and Mashigian 1995; Andersen, Davidson, and Ganz 1994)	<u>Health Care System</u> (Policy, Resources, Organization)	<u>Predisposing Characteristics</u> (Demographic, Social Structure, Health Beliefs), <u>Enabling Resources</u> (Personal/Family, Community), <u>Need</u> (Perceived, Evaluated)	<u>Personal Health Practices</u>	<u>Consumer Satisfaction</u> (Convenience, Availability, Financing, Provider Characteristics, Quality), <u>Perceived Health Status</u>, <u>Evaluated Health Status</u>

Figure 2.1: Use of the Andersen Framework in the Health Services Literature

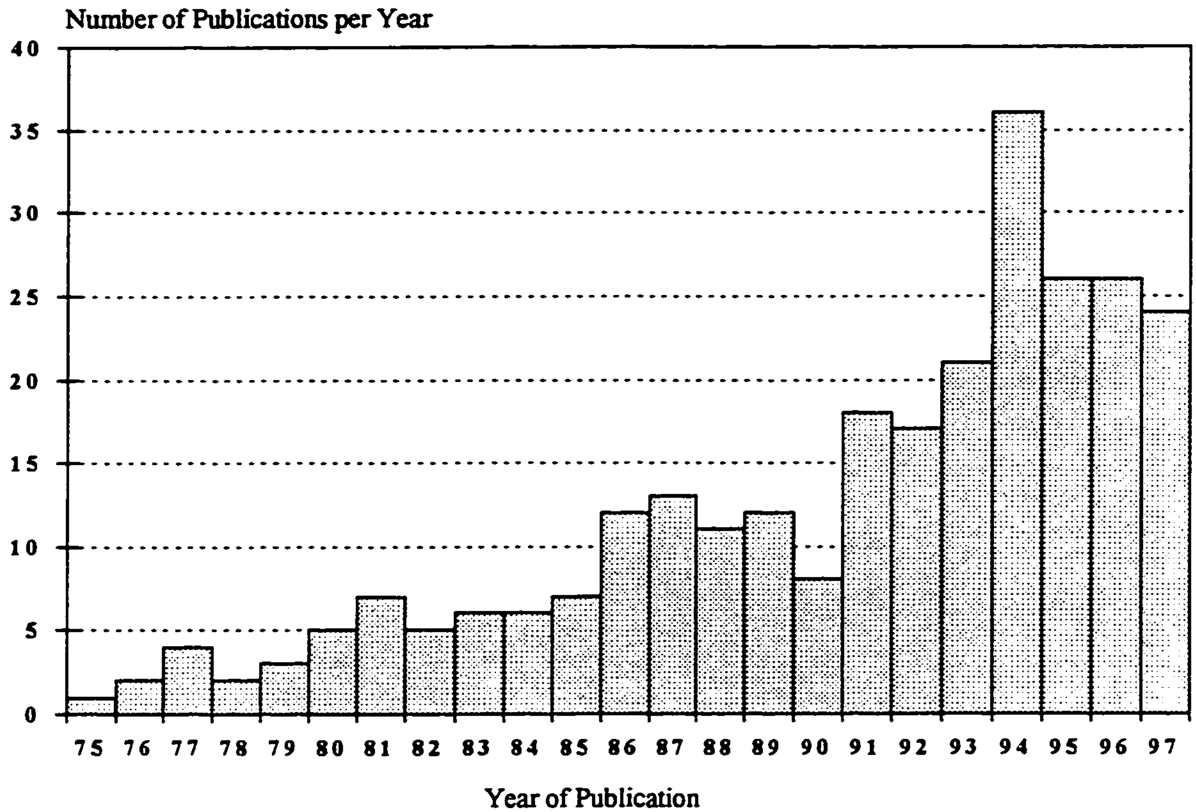


Figure 2.2: Andersen's Behavioral Model (1960s)

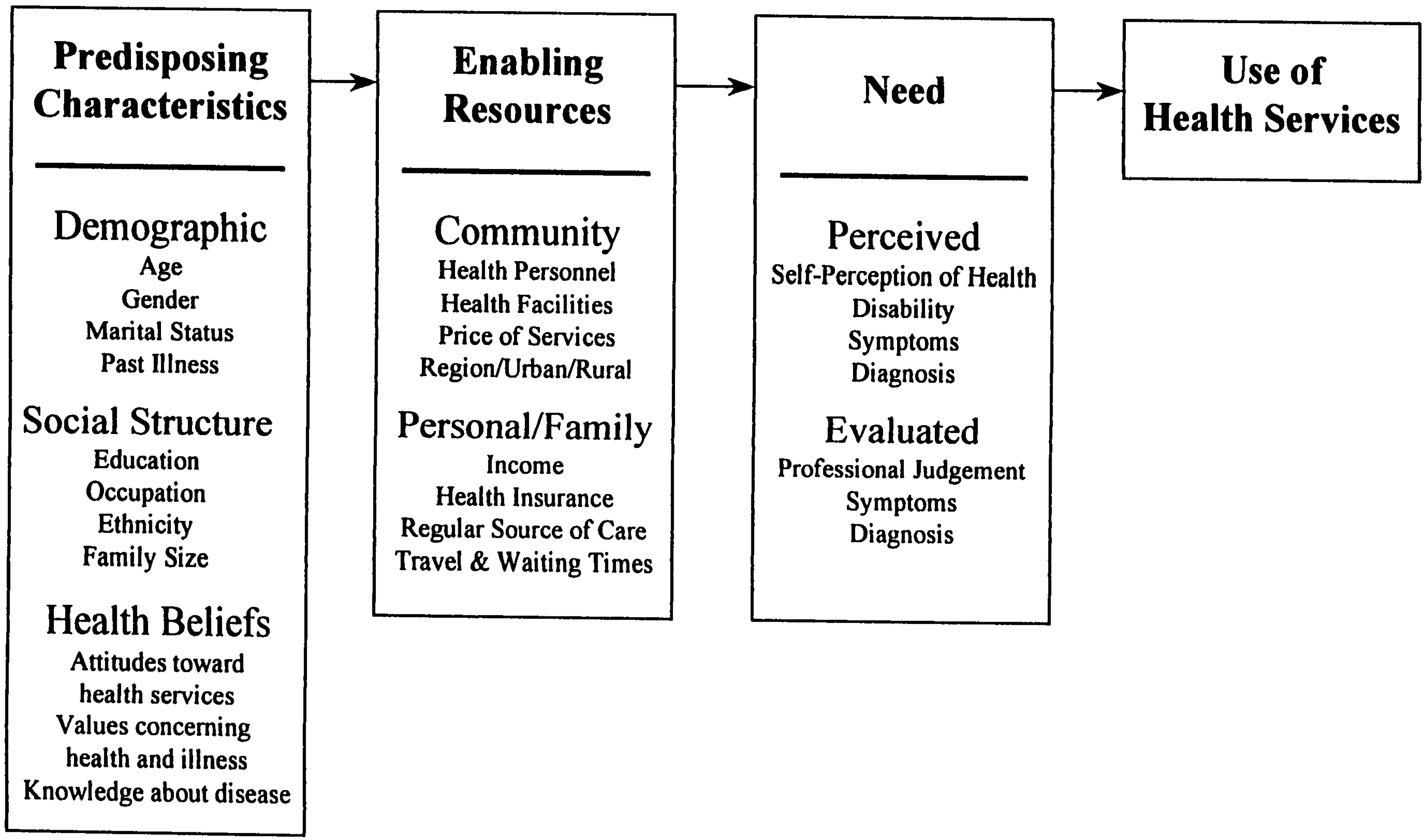


Figure 2.3: Andersen's Behavioral Model (1970s)

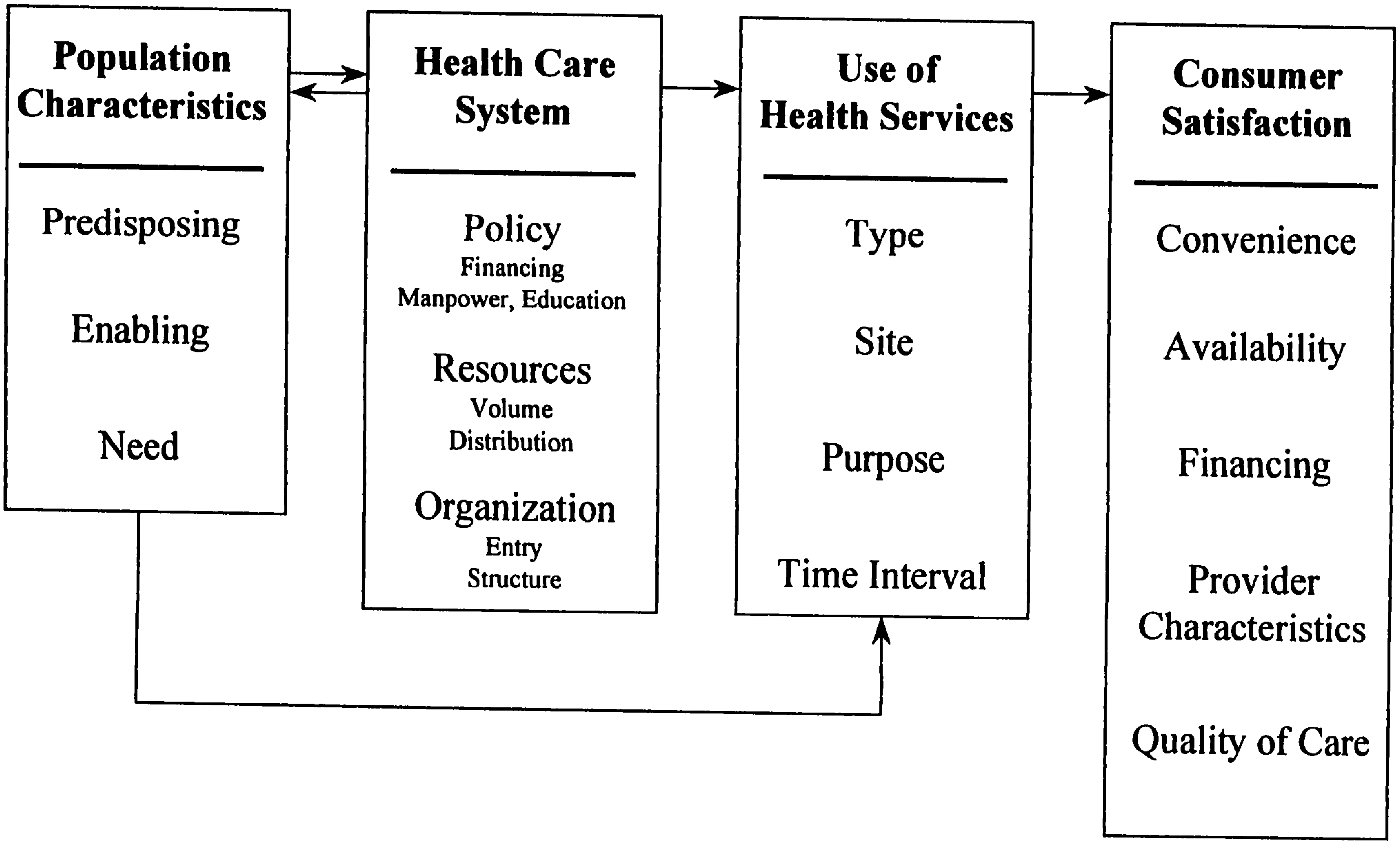


Figure 2.4: Andersen's Behavioral Model (1980s)

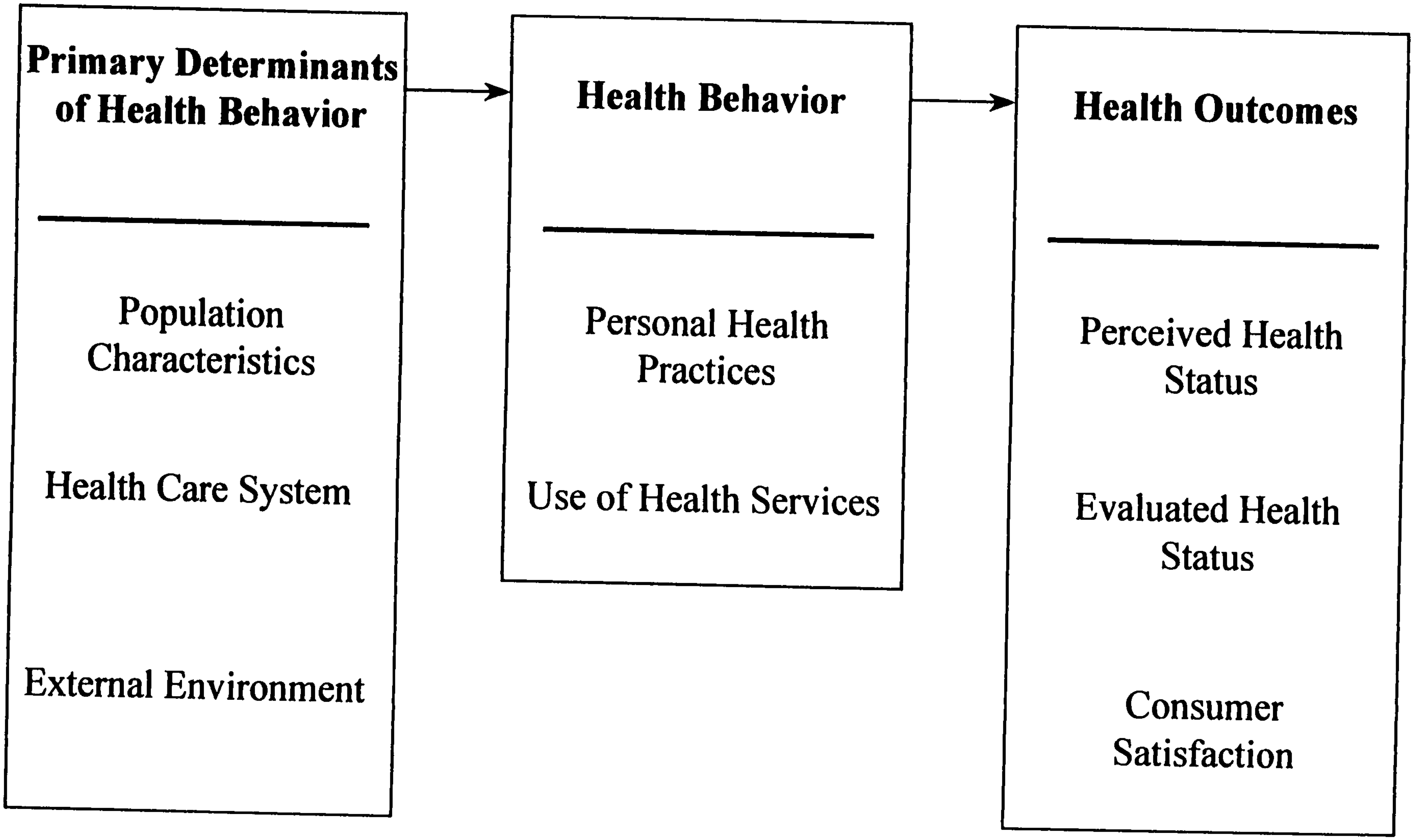


Figure 2.5: Andersen's Behavioral Model (1990s)

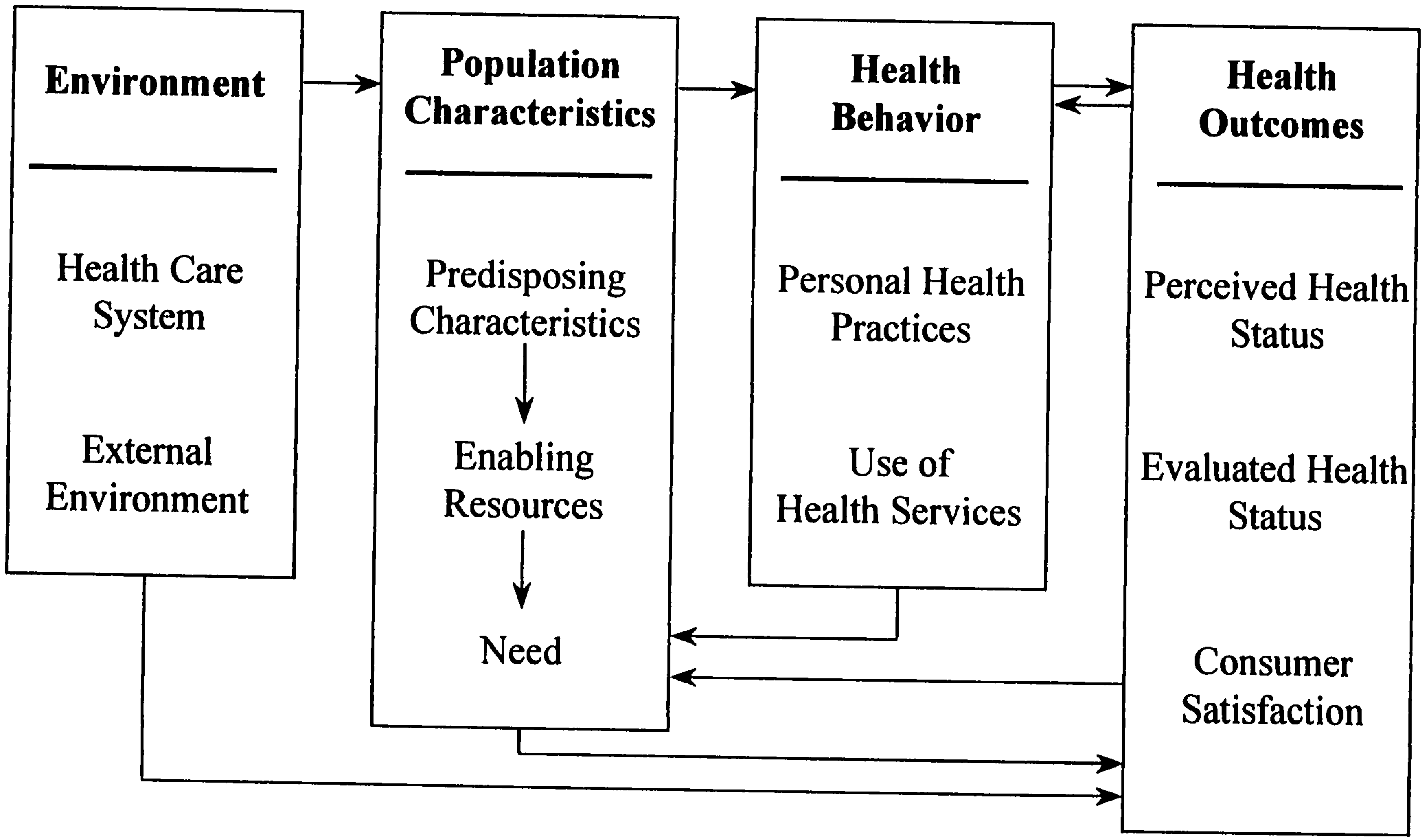
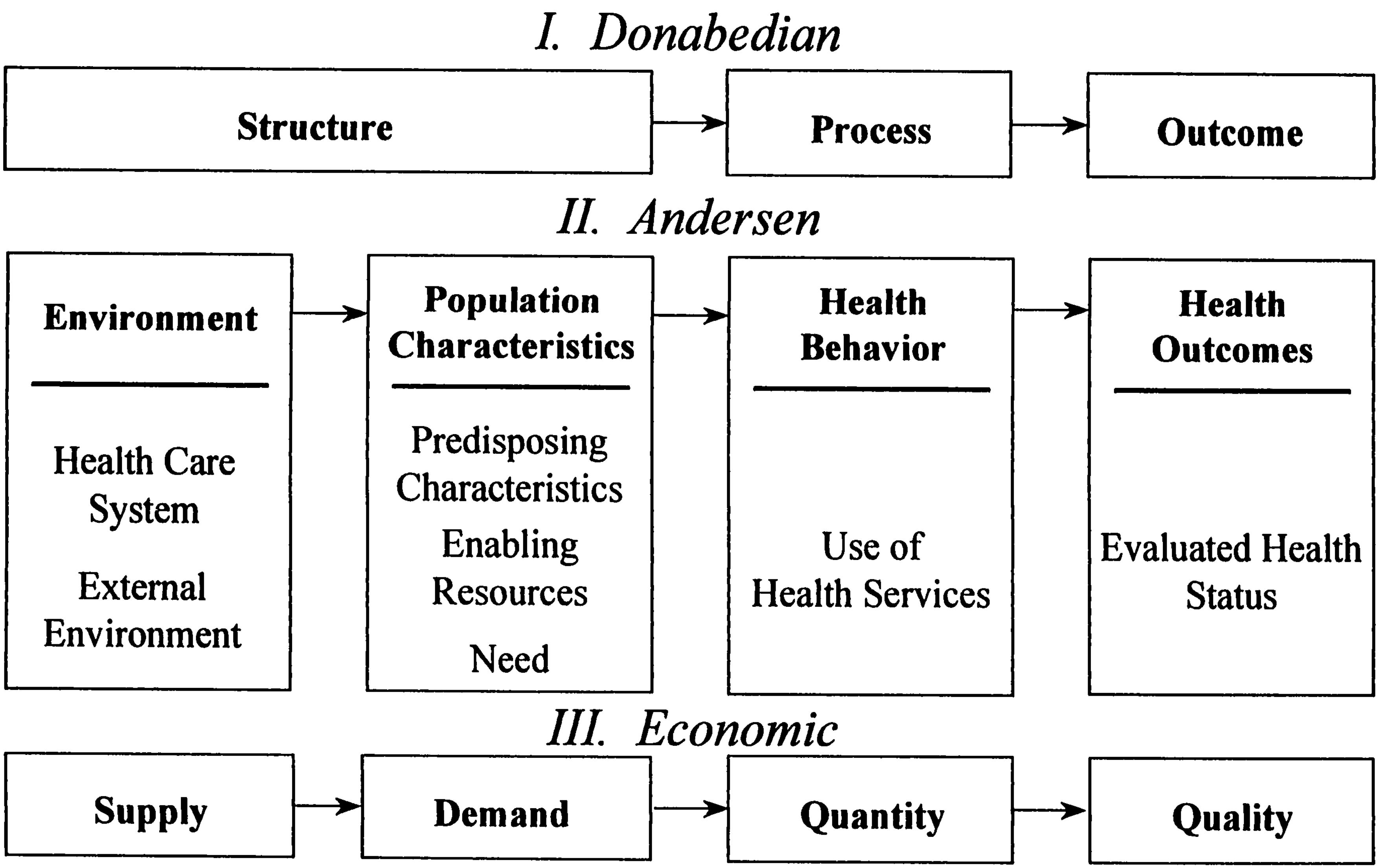


Figure 2.6: Comparative Analytical Frameworks



Bibliography

Aday, L.A. and Andersen, R. (1974): A framework for the study of access to medical care. *Health Services Research*, 9:208-220.

Aday, L.A., Andersen, R.M., and Fleming, G.V. (1980): Health Care in the U.S.: Equitable for Whom? Beverly Hills, CA.: Sage.

Aday, L.A., Andersen, R.M., Loevy, S., and Kremer, B. (1985): Hospital-Physician Sponsored Primary Care: Marketing and Impact. Ann Arbor, MI: Health Administration Press.

Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In J. Kuhl and J. Beckmann (Eds.), *Action-control: From Cognition to Behavior*. Heidelberg: Springer.

Ajzen, Icek and Fishbein, M. (1980), *Understanding Attitudes and Predicting Social Behavior*, Englewood-Cliffs, NJ: Prentice-Hall.

Andersen, R.M. (1968) Behavioral Model of Families' Use of Health Services. Research Series No. 25. Chicago, IL.: Center for Health Administration Studies, University of Chicago.

Andersen, R.M. and Newman, J.F. (1973): Societal and Individual Determinants of Medical Care Utilization in the United States. Milbank Memorial Fund Quarterly Journal, 51:95-124.

Andersen, R.M., Davidson, P.L., and Ganz, P.A. (1994): Symbiotic relationships of quality of life, health services research and other health research. Quality of Life Research, 365-371.

Andersen, R.M., Kravits, J., and Anderson, O.W. (1975): Equity in Health Services: Empirical Anlyses in Social Policy. Boston, MA.: Ballinger Publishing Company.

Andersen, R.M., Marcus, M., and Mashigian, M. (1995): "A Comparative Systems Perspective on Oral Health Promotion and Disease Prevention." In Oral Health Promotion: Socio-Dental Sciences in Action, edited by L.K. Cohen and H.C. Gift. Copenhagen, Denmark: Munsksgaard.

Arrow, K. (1963): "Uncertainty and the welfare economics of medical care." *American Economic Review*, 53(5): 941-73.

Bandura, A. (1977): Self-efficacy: toward a unifying theory of behavioral change. Psychological Review, 84:191-215.

Bandura, A. (1986): Social Foundations of Thought and Action: A Social Cognitive Theory. Englewood Cliffs, NJ: Prentice-Hall Inc..

Donabedian, A. (1974) Benefits in Medical Care Programs. Cambridge: Harvard University Press.

Feldstein, M.S. (1973) "The welfare loss of excess health insurance." *Journal of Political Economy.* 81(5): 853-72.

Feldstein, P.J. (1966): "Research on the Demand for Health Services," *Milbank Memorial Fund Ouarterlv,* 128-163.

Fuchs, V. (1978) "The supply of physicians and the demand for operations. *Journal of Human Resources.* 13(Supp): 35-56.

Newhouse, J.P. (1981) The demand for medical care services: a retrospect and prospect. In J. Van der Gaag and M. Perlman, eds. *Health, Economics, and Health Economics.* Amsterdam: North-Holland.

Nyman, J.A. (1989): The private demand for nursing home care. *Journal of Health Economics,* 8:209-231.

Orem, D.E. (1985): A concept of self-care for the rehabilitation client. Rehabilitation Nursing, 10:33-36.

Pender N (1987) Health Promotion in Nursing Practice.Stamford CT, Appleton and Lange.

Rosenstock, I.M. (1974). Historical origins of the health belief model. *Health Education Monograph,* 2, 328-335.

Wirick, G.C.J. (1966): A multiple equation model of demand for health care. *Health Services Research,* 301-346.

Appendix 2.1: Literature Search Strategies for Utilization Models

2.1.a Search Strategy for Behavioral Models of Health Services Utilization

1. economic model.tw. and ut.fs.
2. economic paradigm.tw. and ut.fs.
3. economic approach.tw. and ut.fs.
4. economic framework.tw. and ut.fs.
5. behavio?ral model.tw. and ut.fs.
6. behavio?ral paradigm.tw. and ut.fs.
7. behavio?ral approach.tw. and ut.fs.
8. behavio?ral framework.tw. and ut.fs.
9. utilization model.tw. and ut.fs.
10. utilization paradigm.tw. and ut.fs.
11. utilization approach.tw. and ut.fs.
12. utilization framework.tw. and ut.fs.
13. theoretic$ model.tw. and ut.fs.
14. theoretic$ paradigm.tw. and ut.fs.
15. theoretic$ approach.tw. and ut.fs.
16. theoretic$ framework.tw. and ut.fs.
17. conceptual model.tw. and ut.fs.
18. conceptual paradigm.tw. and ut.fs.
19. conceptual approach.tw. and ut.fs.
20. conceptual framework.tw. and ut.fs.
21. empiric$ model.tw. and ut.fs.
22. empiric$ paradigm.tw. and ut.fs.
23. empiric$ approach.tw. and ut.fs.
24. empiric$ framework.tw. and ut.fs.
25. 1 or 2 or 3 or 4
26. 5 or 6 or 7 or 8
27. 9 or 10 or 11 or 12
28. 13 or 14 or 15 or 16
29. 17 or 18 or 19 or 20
30. 21 or 22 or 23 or 24
31. 25 or 26 or 27 or 28 or 29 or 30
32. limit 31 to english language

2.1.b Search Strategy for Trends in Andersen Behavioral Models

1. andersen?.tw.
2. (framework or conceptual approach).tw.
3. (aday? or newman?).tw.
4. 1 and (2 or 3)
5. andersen? model.tw.
6. predisposing factor?.tw.
7. predisposing variable?.tw.
8. predisposing characteristic?.tw.
9. predisposing determinant?.tw.
10. enabling factor?.tw.
11. enabling variable?.tw.
12. enabling characteristic?.tw.
13. enabling determinant?.tw.
14. need factor?.tw.
15. need variable?.tw.
16. need characteristic?.tw.
17. need determinant?.tw.
18. 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17
19. (behavio?ral model).tw and ut.fs.
20. 4 or 5 or 18 or 19
21. limit 20 to english language

III. Literature Review

In this chapter the home care literature is reviewed in relation to three objectives. First, the literature is critiqued and evaluated for construct validity and empirical methods. The interpretation of each article is then related to the conceptual and theoretical concepts developed in the previous chapter. Finally, the deficiencies in the literature are shown to demonstrate the need for this dissertation.

Before the validity and methodology of the literature can be evaluated a brief description of the literature search strategy is provided and discussed. The second section then provides an overall discussion of the review, experimental and non-experimental studies. The third section relates the findings of the articles to the Andersen model as presented in the second chapter. Finally, the deficiencies of the literature are summarized and discussed.

A. Description of Literature Search Strategy

Home care studies were identified through an initial electronic database search and iteratively through the bibliographies of selected articles. The initial, electronic search strategy was specifically developed to identify articles where home care services were provided as an alternative to nursing home care or institutionalization. An iterative electronic search strategy was used to identify 3,389 unique citations in the Medline (1966-) and HealthSTAR (1975-)

databases. A combination of MeSH subheadings and keywords were used. The MeSH terms included home care services, community health nursing, community health services, residential facilities, deinstitutionalization, and institutionalization. Keywords included home care, home nursing or deinstitutionalization, institutionalization, community health, community care, home health, home visit, nursing home, and institutional care.

The initial 3,389 citations were abstracted by types of studies, unit of observation, study design and by patient population. Descriptive study types were excluded. The studies included in the review were the case-control, cohort, demonstration (a type of randomized study), and randomized trial. Analytical articles were selected, as these study designs are generally associated with higher quality of evidence. The unit of observation was limited to the individual.

As well, patient eligibility was restricted to clients or patients already enrolled in home care. Patient populations were identified by the transition from eligibility to outcome status or setting. For example, patients can be classified as existing in one of five states: community (or home or self-care); home care (or skilled nursing care or adult day care); acute care (or hospital); nursing home (or residential or institutional); and death (or morbidity). Studies that considered the transition from home care to community, institutional care or death were considered. The grey literature was not reviewed due to the difficulty associated with obtaining selected studies and the fact that these articles are generally considered to be of lower quality.

The resulting articles which met these criteria after abstracting the entire article (as opposed to the title and abstract) and those identified through subsequent searches of selected bibliographies were then grouped into four categories: review, experimental, quasi-experimental, and non-experimental studies. Some of the articles did not meet the criteria listed above but inclusion in the discussion was determined on the basis of merit. In the following three sections, the findings of these articles are tabulated and discussed.

B. Literature Overview

In this section of the chapter, I summarize and discuss the literature that resulted from the search strategy described above. Firstly, eleven review articles are summarized (8, 10, 16, 25, 26, 29, 34, 39, 45, 66, 67). Twenty-six experimental articles (1,2,6,7,9,11,12,13,14,17,22, 37,38,46,49,51,52, 53,54,55,60,63,64,65,68,69,70,71), 7 quasi-experimental (27, 28, 30, 31, 32, 33, 50), are reviewed followed by 24 observational or non-experimental analyses (3, 4, 5, 15, 18, 19, 20, 21, 23, 24, 35, 36, 40, 41, 42, 43, 44, 47, 48, 56, 57, 58, 59, 61, 62) meeting the criteria described above. It should be noted that some datasets are reported in more than one article. In the following sections, those articles representing a common dataset but which differ substantially will be summarized separately.

1. Review Articles

Evaluation of home care programs is difficult due to challenges imposed by analysis of a complex organizational entity that may change during the evaluation (Manton et al 1995). Following this logic, conclusions from any one study should therefore be interpreted with caution. However, such changes are not synchronized across studies and therefore conclusions of multiple program evaluations should be of higher quality. The purpose of this section is to briefly evaluate the review literature. In fact these review articles compare substantially fewer articles than the number analyzed in this chapter and therefore are not representative of the cumulative body of knowledge.

The following table summarizes seven separate studies that review the evidence for the effects of home care service provision on delaying institutionalization. Three review articles (10, 34,67) summarize one experimental study, the National Long Term Care Channeling Demonstration Project, and will be discussed in the next section. When the number of studies reviewed column is blank, the paper is a discussion paper and not a formal review *per se*.

Table 3.1: Literature Review Articles

Authors	**Number of Studies Reviewed**	**Summary of Findings**
Chappell (1994)		Review of literature. Efficacy studies have not concluded that home care is less costly or necessarily better for the quality of life when compared with institutional care
De Klerk, Huijsman, and Rutten (1995).		Review of demonstration projects in the Netherlands.
Hedrick and Inui (1986)	12 Studies	No effect on anything.
Hedrick, Keopsell, and Inui (1989)	13 Studies	Small benefit on mortality.
Hughes (1985)	13 Studies	Poor study designs. Contradictory findings.
Kemper, Applebaum, and Harrigan (1987)		Reviews of community care demonstrations. Home care leads to increased costs and quality of life without reductions in family caregiving.
Weissert, Cready, and Pawelak (1989)	27 Studies	Utilization of all services and cost measures increased. Limited health status improvements

Only Hedrick, Keopsell, and Inui (1989) attempted a meta-analysis of the literature. With the exception of Weissert, Cready, and Pawelak (1988), the number of studies has been relatively small. In addition, many of the references referred to are only available as consultants reports. As a result, the findings may not be well known.

In general, the review literature seems to suggest that evidence for the effectiveness of home care is sparse and inconsistent. Utilization studies suggest that home care utilization goes up and that other health resource consumption also increases. In part this effect can be explained by virtue of the fact that home care may prolong life and thereby allow patients to consume more resources. Alternatively, increased consumption of non-home care health resources may be responsible for increase of benefits. These articles are not reviewed further due to the fact that a greater number of articles are covered in the following sections within the analytic framework set up in the previous chapter.

2. Experimental Studies

Experimental studies of home care and other social programs are referred to as demonstrations. The largest such study in home care is the National Long-Term Care Channeling Demonstration project (NLTCCD). The study design and results have been reported in a number of articles (1, 2, 6, 7, 9, 11, 38, 51, 60, 70). In addition, eight sub-analyses have also been conducted on this dataset (12, 13, 22, 37, 46, 52, 53, 54). The NLTCCD sub-analyses are discussed separately since sub-analysis can introduce selection bias. Finally, the NLTCCD study has been summarized in three articles (10,34,67).

Patient recruitment for the NLTCCD project began in 1982 with a final study size of 6,326 patients referred from the community. Some of the patients may already receive home care services while others are newly admitted. The study in fact consists of two experiments, one for enhanced financial services (primarily direct services) and the other for enhanced case management. Limited regression analysis was performed to control for the effects of possible confounders but did not change the basic results. Institutionalization, survival, formal and informal service utilization, total costs, and functional status were not significantly affected by the interventions. There was evidence however to suggest that patient and caregiver satisfaction had increased.

In summary, although the NLTCCD project has been the most ambitious and the one that adhered to the principles of RCT design as closely as possible, the results were nonetheless inconclusive. The arguments for this phenomenon are basically twofold. First, the study design did not prevent the control group receiving the services given to the experimental group. Secondly, the treatment variables were binary. They did not consider the intensity of resource provision, the duration or frequency, nor did they consider the breadth of service provision.

Three articles reviewing the apparent failure of the NLTCCD were published in 1988 (10,34,67). Christianson et al (1988) primarily review the management of the project itself and suggest that case manager workload was excessive and that the potential for self selection bias was present. Self selection bias can be eliminated by the separation of screening and case management functions within a study. The authors acknowledge however that another problem is introduced in that the patient would need to be visited by two individuals (rather than one). Weissert (1988) argues along similar lines in that the primary failure of the study was not to better target services to those at highest risk of institutionalization. He furthermore asserts that treatment costs should have been capped and that treatments should be tailored to the individual. Finally, Kane (1988) makes the most important observation that no continuous measure of treatment intensity was taken. In fact, there is nothing that prevents controls from obtaining increased financial or case management services. This is facilitated by the Hawthorne effect - where providers improve their patterns of practice by virtue of participating in an experimental setting.

Table 3.2 summarizes the eight NLTCCD sub-analyses, tabulating the sample size, statistical methodology, dependent variable, and results. In general, the analyses are more sophisticated (with the exception #13) reflecting the fact that these studies were largely conducted in the 1990's. As well, there tends to be some evidence that home care services have desirable outcomes (37, 46). In terms of predictive variables 3 studies (12,22,37) report evidence that physical function is a predictor of institutionalization, or home care service intensity. Likewise two studies (12, 22) demonstrate evidence that environmental variables are significant predictors for institutionalization.

Table 3.2: NLTCCD Sub-Analyses

Reference	Sample	Methodology	Dependent Variable	Results
12. Coughlin and McBride (1990)	632	Multi-nomial Logistic Regression	Nursing Home Admission	No evidence for age, sex, formal or informal home care services as predictors. Some evidence for functional status, disease burden, and system variables. Environmental variables are significant predictors.
13. Coughlin, and Liu. (1989)	5626	Univariate (No formal tests)	Total Cost	Total costs have increased for cognitively impaired patients.
22. Greene, Lovely, Miller, Ondrich. (1994)	3503	Markov Models; Optimization Analysis	Transition Probability (Insitutionalization)	No evidence for age and sex. Some evidence that formal services explain transition probabilities. Functional status and environmental variables are shown to mediate institutional transitions.
37. Kemper. (1992)	1039	Duan's 2-part Model (Probit, OLS)	Formal and Informal Care (hours/week)	Some evidence for age, mental function, system and environmental variables. Evidence for gender and physical function status.
46. Manton, Vertrees, and Clark. (1993.)	?	Grade of Membership Procedure, Lifetable Analysis	LOS (Home Care, Nursing Home, Institution)	Some evidence for impact of NLTCCD interventions.
52. Rabiner. (1992)	3387	2SLS, Ordered Probit	Patient Satisfaction	No evidence for formal services. Some evidence for informal service provision. Evidence for age, sex and physical function status.
53. Rabiner, Mutran, and Stearns. (1995)	1726	Structural Equation Model (LISREL)	Patient Satisfaction	Evidence that formal services indirectly affect patient satisfaction. Regression coefficients not reported
54. Rabiner, Stearns, and Mutran. (1994)	2109	Structural Equation Model (LISREL)	Nursing Home Care Utilization (s)	Some evidence that formal service intensity is mediator for patient satisfaction. Regression coefficients not reported

Table 3.3 lists the characteristics of 19 quasi- and experimental, non-NLTCCD articles. The sample size, initial year, length of follow-up, referral source and statistical methods employed are listed for each article. The sample sizes vary from 100 (49) to 1,871 (64). One study was conducted in the 1960's (49) nine were conducted in the 1970's (28, 30, 31, 33, 44, 55, 64, 65, 68) and eight were conducted in the 1980's (14, 17, 27, 32, 50, 63, 69, 71). Follow-up varied from 6 months (14,17,32) to 48 months (30,33,44). Eligibility criteria resulted in patients being admitted from acute care facilities (14, 28, 30, 31, 32, 33, 44, 49, 50, 55, 68), community (17, 27, 28, 55, 63, 64, 65, 68, 69, 71), home care (64,65,68), and nursing homes (55,68). Finally, in addition to univariate analysis, a variety of statistical methods were utilized in these articles. Multiple classification analysis (MCA) was used in three articles (64,65,68); multivariate methods in two (14,27); Kaplan-Meier in two (17,50) and Markov models in one article (44).

Interventions and results for 10 quasi-experimental and experimental studies are summarized in table 3.4 as some study datasets were referred to in more than one article.[1] With the exception of institutionalization, the results tend to support or refute certain outcomes. The Benjamin Rose, Chicago and Day Care/Homemaker studies support the notion that institutionalization can be delayed. However four studies do not support this notion (Copenhagen, Maryland, Monroe County, and Netherlands). The evidence would suggest that home care does not increase survival or functional status. Five studies show no support for survival, while three show an effect. Although the Copenhagen and Georgia studies shows an increase in survival, the Day Care/Homemaker study provides evidence of a decrease in survival. Out of seven studies, only the Chicago shows some evidence that home care improves functional status while the other six show no effect.

Table 3.3: Experimental and Quasi-Experimental Study Characteristics

Reference	Sample	Year	Follow-Up	Referral Source	Method
Cummings, Hughes, Weaver Manheim, Conrad, et al. (1990)	419	1983	6 months	Hospital	Multivariate
Eggert, Zimmer, Hall, and Friedman. (1991)	476	1983	6 months	Community	Kaplan-Meier
Hendriksen, Lund, and Stromgard. (1994)	572	1980	36 months	Community	Log-Linear, Markov Models
Hicks, Raisz, Segal, and Doherty. (1981)	502	1976	24 months	Community, Hospital	Univariate
Hughes, Conrad, Manheim, , and Edelman. (1988)	313	1977	48 months	Hospital	Univariate
Hughes, Cordray, and Spiker. (1984)	245	1977	9 months	Hospital	Univariate
Hughes, Cummings, Weaver, Manheim, Conrad, and Nash. (1990)	233	1984	6 months	Hospital	Univariate
Hughes, Manheim, Edelman, and Conrad. (1987)	313	1977	48 months	Hospital	Univariate
Manheim and Hughes. (1986)	313	1977	48 months	Hospital	Markov Model
Nielsen, Blenkner, Bloom, Downs and Beggs. (1972)	100	1966	?	Hospital	Univariate
Oktay and Volland. (1990)	191	1983	12 months	Hospital	Univariate, Kaplan-Meier
Skellie, Mobley, and Coan. (1982)	747	1976	12 months	Community, Hospital, Nursing Home	Univariate
van Rossum, Frederiks, Philipsen, Portengen, Wiskerke and Knipschild. (1993)	580	?	?	Community	Univariate
Wan, Weissert, and Livieratos. (1980)	1871	1975	12 months	Community, Home Care	Univariate, MCA
Weissert, Wan, Livieratos and Katz. (1980)	384	1975	12 months	Community, Home Care	Univariate, MCA
Weissert, Wan, Livieratos and Pellegrino. (1980)	884	1975	12 months	Home Care, Hospital, Nursing Home	Univariate, MCA
Williams, Williams, Zimmer, et al. (1987)	117	1983	12 months	Community	Univariate
Zimmer, Groth-Juncker and McCusker. (1985)	167	1983	6 months	Community	Univariate

[1] In fact some of the articles are almost identical while others take a slightly different perspective. Nonetheless, the results are aggregated for each dataset.

Other important outcomes include satisfaction (patient and caregiver), hospitalization, total health care costs, and quality of life. Increases in both satisfaction and quality of life are supported by studies without counter-evidence (in these studies). Hospitalizations, drug and total costs are shown to increase however. Only Monroe and the VA Hospital studies provide some evidence that total costs decrease. In part the decrease in total costs is due to the fact that early discharge is probably cost saving compared to long-term home care. In other words, the increase of home care costs generally is not offset by decreases in other health care expenditures.

One explanation for lack of conclusive evidence that home care delays institutionalization, increases survival and improves functional status lies in the interventions themselves. Only the Chicago, Georgia, Maryland, and VA Hospital studies include direct services in their intervention. Others consider home care aides who provide homemaker services and personal care or assessment and case management (Connecticut, Copenhagen, Monroe County, and Netherlands). As well, not all studies are asking these important questions. Out of ten studies, 7 address institutionalization, 8 address survival, and 7 address changes in functional status. Lastly, imprecise measurement of interventions and variables cause imprecision in the results (eg. dichotomous service intensity or a single visit in the last year).

Table 3.4: Experimental and Quasi-Experimental Study Interventions and Results

Study (References)	Intervention	Results
Benjamin Rose Hospital (49)	Home Care Aide Service	Institutionalization is delayed, satisfaction is increased and no evidence that survival is affected.
Chicago: Five Hospital Homebound Elderly Program (30,31,33,44)	Comprehensive Home Care	Institutionalization is delayed, total cost and quality of life is increased, hospitalizations and survival are not affected, and some evidence that functional status is decreased.
Connecticut: Triage (28)	Assessment, Case Management	Total cost is significantly increased an no evidence for alteration of functional status.
Copenhagen (27)	Assessment	Hospitalization is decreased and survival is increased but institutionalization is not affected.
Day Care/Homemaker (64,65,68)	Geriatric Day Care and Homemaker Services	Some evidence that institutionalization is delayed, survival is decreased however while total cost is increased and no evidence for physical function.
Georgia: Alternative health Services Project (55)	Comprehensive Home Care	Survival, drug and total costs are shown to have increased with home care.
Maryland Post-Hospital Support Program (50)	Comprehensive Home Care	Caregiver stress is reduced, but no evidence that institutionalization, survival, or functional status are affected.
Monroe County (17,69,71)	Team-oriented Geriatric Assessment Clinic	Total costs have decreased, some increase in satisfaction and some decrease in total costs but no evidence for institutionalization, survival or functional status.
Netherlands (63)	Case Management	Total costs have increased, but no evidence for institutionalization rates, survival or functional status.
VA Hospital-Based Home Care (14,32)	Hospital-Based Home Care	Satisfaction has increased, hospitalization has decreased, some evidence for reduction in total costs, and no evidence for change in functional status.

3. Non-Experimental Studies

Non-experimental or observational studies differ from experimental studies in that they require statistical methodology to adjust for confounding variables. When patients are not randomized to treatment and control groups, the choice of treatment is usually considered to be correlated with patient or provider characteristics. The purpose of statistical analysis is to remove or adjust for these effects. As a benefit the relationship between dependent and independent variables are usually recorded. Although in experimental study designs the effect of the intervention is of primary importance, the researcher is not obligated to specify which of independent variable(s) is(are) of primary interest in observational studies.

In this section of the chapter, I will summarize 14 observational studies encompassing the 24 articles identified in the literature search strategy described previously. As in the previous section, some articles overlap with respect to data sources and can be organized into larger "studies."

In Table 3.5, the study, reference number, author(s), sample size, initial year, length of follow-up and referral sources are summarized. Sample sizes vary enormously from 101 (36) to 59,721 (59) and unkown in three articles (19,35,61). Only 6 articles or two studies: B.C. (40, 41,56,57) and Massachusetts (5,20) were conducted primarily in the 1970's, the remaining articles and studies were initiated (and completed) in the 1980's.

Follow-up varies from 6 months (42) to 60 months for the B.C. studies (40,41,56,57) while 10 articles do not provide sufficient information (3,4,21,23,35,36,47,58,61,62). Finally, patients have been referred to these programs from the community in 14 articles, home care in 13, nursing homes in 10, and hospitals in only one article.

Table 3.5: Summary of Observational Studies

Study	No	Reference	Sample	Year	Follow-Up (Months)	Referral Source
Arizona	4	Bauer. (1996)	2923	1989	?	Community
	23	Greene. (1983)	124	1980	?	Home Care
Boston	61	Trisolini, Thomas, and Cashman, Payne. (1994)	273 (visits)	1989	?	Home Care
B.C.	40	Lane, Uyeno, Stark, Gutman and McCashin. (1987)	9483	1978	60	Home Care, Nursing Home
	41	Lane, Uyeno, Stark, Kliewer and Gutman. (1985)	1653	1978	60	Home Care, Nursing Home
	56	Stark, Gutman and McCashin. (1982)	3518	1978	60	Nursing Home
	57	Stark and Gutman. (1986)	1653	1978	60	Home Care, Nursing Home
Bundoora	42	Lazarus and Gray. (1988)	116	1985	6	Community, Home Care, Hospital, Nursing Home
Lucas County	58	Starrett, Rogers, Walters. (1988)	400	1984	?	Community, Home Care
Maryland	62	Tsuji, Whalen, and Finucane. (1995)	334	1986	?	Community
Massachusetts	5	Branch, Jette, Evashwick, Polansky, Rowe and Diehr. (1981)	1625	1974	12	Community, Home Care
	20	Evashwick, Rowe, Diehr and Branch. (1984)	1317	1974	15	
Minnesota PAS/ACG Program	15	Davidson, Moscovice and McCaffrey. (1989)	239	1984	12	Community
	48	Moscovice, Davidson and McCaffrey. (1988)	214	1984	12	Community
Monroe County	59	Temkin-Greenerand Meiners. (1995)	59,721	1984	48	Community, Nursing Home
Netherland	35	Kempen and Suurmeijer. (1991a)	?	1987	?	Community
	36	Kempen and Suurmeijer. (1991b)	101	1987	?	Community
New York: NHWW	21	Gaumer, Birnbaum, Pratter, Burke, Franklin and Ellingson-Otto. (1986)	724	1980	?	Community
NHIS Longitudinal Study	47	Mohr. (1994)	4335	1984	?	Community, Home Care, Nursing Home
NLTC Survey	18	Ettner. (1993)	380	1982	24	Nursing Home
	19	Ettner. (1995)	?	1982	24	Home Care
	24	Headen. (19??)	5581	1982	24	Community, Home Care, Nursing Home
	43	Liu, Coughlin and McBride. (1991)	5795	1982	24	Community, Home Care, Nursing Home
Ohio	3	Bass and Noelker. (1987)	586	?	?	Home Care

C. Conceptual and Theoretical Overview of Observational Studies

In this section of the paper the observational studies are summarized and compared to the results and predictions generated by the Andersen framework.. The three sections will deal with the statistical methods, dependent variables and independent variables. Independent variables are organized according to the Andersen framework discussed in the last chapter. Table 3.6 summarizes the statistical methods and dependent variables for the 24 articles.

1. Statistical Models

The evolution of the Andersen of Behavioral model resulted in a complex relationship between the variables hypothesized to be significant explanatory predictors. For example, Figure 2.5 illustrates a complicated set of equations by way of a path diagram. Each arrow represents a separate relationship that should be modelled with a separate equation. In econometrics there are two methods that allow for the simultaneous estimation of relationships: two-stage least squares or three-stage least squares. In the field of psychological measurement, causal modelling is more common and is also able to estimate these relationships.

Table 3.6: Summary of Observational Study Methods and Dependent Variables

Study	No	Statistical Method	Dependent Variable
Arizona	4	Cox Proportional Hazards	LOS (home care)
	23	Causal Modelling (2 simultaneous equations)	Formal Support (levels), Informal Support (levels)
Boston	61	Multivariate Regression (stepwise)	Formal Services (nursing visit time)
British Columbia	40	Markov Chain Model	Transition Probability (insitutionalization)
	41	Markov Chain Model, Moving Average Growth, and Log-Linear Regression	Transition Probability (insitutionalization)
	56	Univariate (no formal tests)	
	57	Univariate (no formal tests)	
Bundoora	42	Cox Proportional Hazards	LOS (home care)
Lucas County	58	Path Model, Hierarchical Regression Analysis	Home Care Utilization (form not specified)
Maryland	62	Cox Proportional Hazards (stepwise)	LOS (home care)
Massachusetts	5	Hierarchical Regression Analysis	Home Care Utilization
	20	Hierarchical Regression Analysis (stepwise)	Home Care Utilization (dichotomous, used in past 15 months)
Minnesota PAS/ACG Program	15	Cox Proportional Hazards	LOS (Home Care)
	48	Causal Modelling (2 simultaneous equations), 2SLS	Formal Care, Informal Care
Monroe County	59	Discrete Time Hazard Function (competing hazards)	LOS (Home Care)
Netherlands	35	Mokken Scale Analysis	Home Care Utilization (y/n)
	36	Hierarchical Regression Analysis	Home Care Casemix (logarithm)
New York: Nursing Homes Without Walls	21	Univariate	
NHIS Longitudinal Study on Aging	47	Structural Equation Model (path analysis), Logistic Regression	Functional Status, Nursing Home Admission
NLTC Survey	18	Univariate, Probit	Waiting List Status (probit model)
	19	Probit	Home Care Utilization (days/week)
	24	Cox Proportional Hazards (competing hazards)	LOS (home care)
	43	Weibull Hazard Function (single equation, competing hazards)	LOS (home care)
Ohio	3	Hierarchical Regression Analysis	Home Care Utilization (dichotomous)

There are limitations though in that the type of model is constrained by the nature of the dependent variable. For example, length of stay is a "time-to-event" variable and requires lifetable (actuarial) or survival analysis (Kaplan-Meier, parametric or semi-parametric hazard models). These dependent variables are also referred to as censored and therefore require these special techniques. More will be said about survival models in the next chapter. The limitation imposed by these requirements is that there is no model that can deal with "time-to-event" and causal modelling simultaneously.

Lastly, an important consideration is that observational data are subject to selection bias and special techniques have been developed to deal with them. In the case of experimental data, the randomization procedure ensures an equal distribution of covariates between the treatment and control groups. In this case univariate analysis is sufficient. In the presence of selection bias however, models such as those by Heckman or Duan are required (further discussed on these models is postponed until the next chapter).

Therefore the three articles that report univariate statistical methods (21,56,57) ignore selection bias and are incorrectly specified. As these studies are quite old (1980, 1978, 1978 respectively) no harm is done by ignoring them. Three additional papers employ techniques that will not be further considered in this dissertation (35,40,41). Although it should be noted that the Markov models used in the B.C. study (40,41) have merit and should be further explored elswhere. Multivariate techniques (especially the single equation variety) are likewise useful and have been considered in three papers (18,19,61). In particular two (18,19) have attempted to correct for selection bias but include too few variables in favor of a more theoretical treatment.

Four articles (3,5,20,36) employ hierarchical regression analysis that calculates the incremental explanation of variation through the stepwise addition of blocks of variables. In particular this is the classical technique used to evaluate the Andersen model. Indeed these articles make explicit reference to the framework. One limitation of this method is the increased complexity of estimating direct versus indirect effects (more easily accomplished in causal modelling). More importantly, the dependent variable is constrained (eg. skewed, dichotomous, or censored dependent variables are inappropriate for this statistical model).

The estimation of simultaneous relationships is better accomplished with causal modelling. Only four articles (23,47,48,58) employ this technique. Sample sizes for three of the studies (23,48,58) vary from 124 to 400 and is judged to be insufficient for this type of model. One study (47) has a sufficient sample size of 4,335.

The next chapter will discuss survival analysis carefully as it is the main regression techniques used in this dissertation. There are 7 studies that employ this technique (4,15,24,42,43,59,62). One model employs stepwise selection of variables (62). Three others are (4,15,42) utilize the basic semi-parametric form. Three studies employ competing hazard analysis (24,43,59) and one estimation is with the parametric form (43). Competing hazard analysis is important since death and institutionalization for example "compete with each other." In other words they are mutually exclusive outcomes for a single patient. Finally, the dependent must meet fairly strict criteria in order to estimate a parametric survival function.

2. Dependent Variables

As alluded to earlier, the selection of the dependent variable and the statistical model are related. For the four articles (3,5,20,36) which employed hierarchical regression analysis one article did not specify the exact specification of home care utilization (5), two were dichotomous (3,20) and one was a logarithmic transformation of home care casemix (20). Linear regression techniques are not appropriate for the two dichotomous dependent variables and the technique does not allow for endogeneity (selection bias).

The estimation of simultaneous relationships is accomplished with causal modelling in four articles (23,47,48,58). In two of the cases (48,58) the specification of home care utilization is not provided. Levels of formal and informal care characterize another study (23). In one case functional status and nursing home admissions are the dependent variables (47). Seven studies employ the survival technique (4,15,24,42,43,59,62), all of which use length of stay in home care as a dependent variable. The primary limitation of causal modelling is the inability to incorporate duration of home care as a dimension into utilization of home care. The primary limitation of survival analysis is the lack of methods to deal with selection bias in the presence of censored variables. (In the next chapter, a novel solution is proposed.)

3. Independent Variables

Table 3.7 summarizes the evidence for independent variables for selected statistical models. The table lists the reference number, statistical model, and variables. These 9 categories of variables correspond to the Andersen framework.. Categories abstracted are age, sex, formal support, informal support, physical function, mental function, weighted comorbidity index, system variables, and environmental variables respectively. The evidence has been categorized into levels of evidence (described in the legend). Each category is discussed briefly following the table with a detailed discussion of these variables in the next two chapters.

Table 3.7: Evidence for the Andersen Model in Selected Observational Studies†

No	Statistical Method	Age	Sex	FS	IS	PFn	MFn	WCI	Sys	Env
23	Causal Modelling	0	-	**	**	**	*	-	-	0
47		**	0	-	*	**	-	-	0	0
48		0	**	0	0	*	0	-	0	0
58		**	0	-	-	**	-	-	-	*
3	Hierarchical	0	0	-	*	*	0	*	-	*
5	Regression	0	0	-	-	**	-	0	-	0
20		**	0	-	-	**	-	0	-	**
36		0	0	-	**	**	0	-	-	*
4	Survival Analysis	**	0	*	0	0	-	*	*	*
15		0	0	0	**	**	0	-	-	0
24		**	**	0	-	0	**	**	*	**
42		**	0	-	-	*	0	-	-	-
43		**	**	**	**	**	**	**	*	*
59		**	**	*	-	-	-	*	-	-
62		-	-	-	-	**	0	0	-	*

† 0 means no evidence, *means some evidence, ** means strong evidence, - not included.

a) Predisposing Characteristics

Table 3.7 summarizes the evidence for predisposing variables for selected statistical models. The two predisposing variables are age and sex. There is substantially more evidence that age is an important predictor of home care utilization as strong evidence exists for 8/14 as opposed to gender at 4/13 papers. The fact that there is less support in the hierarchical regression models is that need and enabling variables explain most of the variation leaving little (direct) explanatory power for predisposing variables. Both variables will be considered in the dissertation.

b) Enabling Resources

Table 3.7 summarizes the evidence for enabling variables for selected statistical models. The two categories of variables considered are formal and informal resources (usually resource intensity). There is more evidence in support of informal resources because the dependent variable is usually formal resource utilization. Also, selection bias makes the interpretation of formal resource intensity difficult in that formal resource intensity is typically endogenous. More will be said on endogeneity and selection bias in the next chapter.

Strong evidence exists ($p<0.05$) for 2/7 papers while some evidence exists ($0.1<p<0.05$) for another 2 papers (=4/7). More evidence exists for informal resources in a total of 6/8 articles.

c) Need Variables

Table 3.7 summarizes the evidence for need variables for selected statistical models. As is expected and widely reported (3,5,20,36), need variables explain the majority of variations in home care utilization. In particular physical functioning (usually expressed as ADL deficits) is significant in 12/14 of the models. Good measures of mental functioning are not as prevalent as given the higher exclusion rate (6/15 as opposed to 1/15 omissions for physical functioning). Weighted comorbidity, disease burden or general health conditions provided some additional evidence in 5/8 studies.

d) Health Care System

Table 3.7 summarizes the evidence for health care system variables for selected statistical models. Because of the difficulty in expressing variations in system wide variables at the individual level, not many studies have attempted any such measures. In fact only 5 papers attempted construction of health care system variables and only 3/5 provided some evidence.

e) External Environment

Table 3.7 summarizes the evidence for external environment variables for selected statistical models. These variables include constructs such as income, family arrangement, language, and race. As a result, more papers have attempted their inclusion (13/15). Of the 13, 8 have shown some evidence while 2 demonstrate strong evidence.

D. Deficiencies in the Home Care Literature

1. Inconclusive Results

The predominant conclusion of the studies is that the evidence for the effectiveness of home care as a deterrent for nursing home placement is inconclusive. Although this dissertation deals with explanation and prediction, and not effectiveness, the issue remains whether enabling resources are significant predictors of home care utilization. In other words, the distinction is subtle to the point of nonexistence (in this case). The effectiveness question is: (1) Does home care prevent nursing home admissions or institutionalization? Whereas two utilization questions one could ask are (2) Do home care services explain or predict nursing home admissions or institutionalization? And (3) does an increase in home care explain or predict a decrease in nursing home admissions or institutionalization? Given the body of evidence and the rationale for implementation of home care programs, question (3) is more appropriate than question (2). Furthermore, question (3) does not differ from question (1).

This inconclusiveness points to the requirement for understanding the relationship between the intensity of home care provision and the quantity of length of service over which the services are demanded. This issue addresses the construction of the variables themselves. Rather than express formal or informal care as a dichotomous variable it will be expressed as an intensity variable. Furthermore, both intensity and duration need to be simultaneously considered as their product equals total utilization.

2. Confounding Variables

The second overall conclusion is that there does not exist a unifying framework for the inclusion of confounding variables in studies of utilization. The results appear to be driven by data availability as opposed to theory.

One purpose of literature review was to facilitate or justify definition and construction of a minimum dataset for the analysis of homecare. Commensurate with the dataset is a theory and associated predictions. Although the Andersen model is but one of many frameworks which explain utilization or demand for health care services, it is an encompassing model with a strong academic tradition.

3. Self Selection Bias

Lastly, both the variables and the statistical procedures do not address the most fundamental limitation in studies of this kind - self-selection bias. Using the NLTCCD studies as an example, the controls were not prevented from obtaining enhanced home care services. Furthermore, this study design is not possible due to the substitution effect between formal and informal services that would result from limiting formal services. That means that unobservable patient characteristics partially explain variation in the intensity and duration of home care service utilization. The primary objective of this dissertation is to directly address this issue through the development of a statistical procedure addressing self-selection bias in the context of censored variables and incorporation of service intensity.

Bibliography

1. Applebaum, RA (1988) The evaluation of the National Long Term Care Demonstration. 3. Recruitment and characteristics of channeling clients. Health Services Research 23(1):51-66.

2. Applebaum, RA, Christianson, J.B., Harrigan, M., and Schore, J. (1988) The evaluation of the National Long Term Care Demonstration. 9. The effect of channeling on mortality, functioning, and well-being. Health Services Research 23(1):143-159.

3. Bass DM, Noelker LS. The influence of family caregivers on elder's use of in-home services: an expanded conceptual framework. Journal of Health & Social Behavior 1987;28(2):184-96.

4. Bauer, EJ (1996) Transitions from home to nursing home in a capitated long-term care program: the role of individual support systems. Health Services Research 31(3):309-326.

5. Branch L, Jette A, Evashwick C, Polansky M, Rowe G, Diehr P. Toward understanding elders' health service utilization. Journal of Community Health 1981;7(2):80-92.

6. Brown, R.S. (1988) The evaluation of the National Long Term Care Demonstration. 2. Estimation methodology. Health Services Research 23(1):23-49.

7. Carcagno, G.J. and Kemper, P. (1988) The evaluation of the National Long Term Care Demonstration. 1. An overview of the channeling demonstration and its evaluation. Health Services Research 23(1):1-22.

8. Chappell, N.L. (1994) Home care research: what does it tell us? Gerontologist 34(1):116-120.

9. Christianson, J.B. (1988) The evaluation of the National Long Term Care Demonstration. 6. The effect of channeling on informal caregiving. Health Services Research 23(1):99-117.

10. Christianson, J.B., Applebaum, R., Carcagno, G., and Phillips, B. (1988) Organizing and delivering case management services: lessons from the National Long Term Care Channeling Demonstration. Home Health Care Services Quarterly 9(1):7-27.

11. Corson, W., Grannemann, T., and Holden, N. (1988) The evaluation of the National Long Term Care Demonstration. 5. Formal community services under channeling. Health Services Research 23(1):83-98.

12. Coughlin TA, McBride TD, Liu K. Determinants of transitory and permanent nursing home admissions. Medical Care 1990;28(7):616-31.

13. Coughlin, T.A. and Liu, K. (1989) Health care costs of older persons with cognitive impairments. Gerontologist 29(2):173-182.

14.Cummings, J.E., Hughes, S.L., Weaver, F.M., Manheim, L.M., Conrad, K.J., et al. Cost-effectiveness of Veterans Administration hospital-based home care. A randomized clinical trial. Archives of Internal Medicine 150(6):1274-1280, 1990.

15.Davidson, G., Moscovice, I., and McCaffrey, D. Allocative efficiency of case managers for the elderly. Health Services Research 24(4):539-554, 1989.

16.de Klerk, M.M., Huijsman, R., and Rutten, F.F. (1995) New options in long-term care for the elderly: evaluation results of demonstration projects in The Netherlands. Home Health Care Services Quarterly 15(2):19-40.

17.Eggert, G.M., Zimmer, J.G., Hall, W.J., and Friedman, B. Case management: a randomized controlled study comparing a neighborhood team and a centralized individual model. Health Services Research 26(4):471-507, 1991.

18.Ettner SL. Do elderly Medicaid patients experience reduced access to nursing home care? Journal of Health Economics 1993;12(3):259-80.

19.Ettner SL. The Effect of the Medicaid Home Care Benefit on Long-Term Care Choices of the Elderly. Economic Inquiry 1995;32(1):103-27.

20.Evashwick C, Rowe G, Diehr P, Branch L. Factors explaining the use of health care services by the elderly. Health Services Research 1984;19(3):357-82.

21.Gaumer, G.L., Birnbaum, H., Pratter, F., Burke, R., Franklin, S., and Ellingson-Otto, K. (1986) Impact of the New York Long-Term Home Health Care Program. Medical Care 24(7):641-653.

22.Greene VL, Lovely ME, Miller MD, Ondrich JI. Reducing nursing home use through community long-term care: an optimization analysis. Journals of Gerontology 1994;(4):S259-68.

23.Greene VL. Substitution between formally and informally provided care for the impaired elderly in the community. Medical Care 1983;21(6):609-19.

24.Headen, AE. (19??) Economic disability and health determinants of the hazard of nursing home entry. The Journal of human Resources 28(1):80-110.

25.Hedrick, S.C. and Inui, T.S. (1986) The effectiveness and cost of home care: an information synthesis. Health Services Research 20(6:Pt 2):851-80.

26.Hedrick, S.C., Koepsell, T.D., and Inui, T. Meta-analysis of home-care effects on mortality and nursing-home placement. Medical Care 27(11):1015-1026, 1989.

27.Hendriksen, C., Lund, E., and Stromgard, E. Consequences of assessment and intervention among elderly people: a three year randomised controlled trial. British Medical Journal Clinical Research Ed 1984(6457):1522-1524, 1994.

28.Hicks, B., Raisz, H., Segal, J., and Doherty, N. The triage experiment in coordinated care for the elderly. American Journal of Public Health 71(9):991-1003, 1981.

29.Hughes, S.L. (1985) Apples and oranges? A review of evaluations of community-based long-term care. Health Services Research 20(4):461-488.

30.Hughes, S.L., Conrad, K.J., Manheim, L.M., and Edelman, P.L. Impact of long-term home care on mortality, functional status, and unmet needs. Health Services Research 23(2):269-294, 1988.

31.Hughes, S.L., Cordray, D.S., and Spiker, V.A. Evaluation of a long-term home care program. Medical Care 22(5):460-475, 1984.

32.Hughes, S.L., Cummings, J., Weaver, F., Manheim, L.M., Conrad, K.J., and Nash, K. A randomized trial of Veterans Administration home care for severely disabled veterans. Medical Care 28(2):135-145, 1990.

33.Hughes, S.L., Manheim, L.M., Edelman, P.L., and Conrad, K.J. Impact of long-term home care on hospital and nursing home use and cost. Health Services Research 22(1):19-47, 1987.

34.Kane, R.A. The noblest experiment of them all: learning from the national channeling evaluation. Health Services Research 23(1):189-198, 1988.

35.Kempen GI, Suurmeijer TP. Factors influencing professional home care utilization among the elderly. Social Science & Medicine 1991;32(1):77-81.

36.Kempen GI, Suurmeijer TP. Professional home care for the elderly: an application of the Andersen-Newman model in The Netherlands. Social Science & Medicine 1991;33(9):1081-9.

37.Kemper P. The use of formal and informal home care by the disabled elderly. Health Services Research 1992;27(4):421-51.

38.Kemper, P. The evaluation of the National Long Term Care Demonstration. 10. Overview of the findings. Health Services Research 23(1):161-174, 1988.

39.Kemper, P., Applebaum, R., and Harrigan, M. Community care demonstrations: what have we learned? Health Care Financing Review (4):87-100, 1987.

40.Lane D, Uyeno D, Stark A, Gutman G, McCashin B. Forecasting client transitions in British Columbia's Long-Term Care Program. Health Services Research 1987;22(5):671-706.

41.Lane D, Uyeno D, Stark A, Kliewer E, Gutman G. Forecasting demand for long-term care services. Health Services Research 1985;20(4):435-60.

42.Lazarus, R. and Gray, L. Clinical characteristics and outcomes in a cohort of patients who were assessed as in need of nursing-home care. Medical Journal of Australia 149(8):410-411, 1988.

43.Liu, K., Coughlin, T., and McBride, T. Predicting nursing-home admission and length of stay. A duration analysis. Medical Care 29(2):125-141, 1991.

44.Manheim LM, Hughes SL. Use of nursing homes by a high-risk long-term care population. Health Services Research 1986;21(2:Pt 1):161-76.

45.Manton KG, Newcomer R, Lowrimore GR, Vertrees JC, Harrington C. Evaluating long-term care demonstrations in real time with study design and plan performance interactions. Medical Care 1995;33(13):1210-27.

46.Manton, K.G., Vertrees, J.C., and Clark, R.F. A multivariate analysis of disability and health, and its change over time in the National Channeling Demonstration data. Gerontologist 33(5):610-618, 1993.

47.Mohr RA. A causal modelling approach to instituttionalization. Health Services Management Research 1994;7(2):79-90.

48.Moscovice I, Davidson G, McCaffrey D. Substitution of formal and informal care for the community-based elderly. Medical Care 1988;26(10):971-81.

49.Nielsen, M., Blenkner, M., Bloom, M., Downs, T., and Beggs, H. Older persons after hospitalization: a controlled study of home aide service. American Journal of Public Health 62(8):1094-1101, 1972.

50.Oktay, J.S. and Volland, P.J. Post-hospital support program for the frail elderly and their caregivers: a quasi-experimental evaluation. American Journal of Public Health 80(1):39-46, 1990.

51.Phillips, B.R., Kemper, P., and Applebaum, R.A. The evaluation of the National Long Term Care Demonstration. 4. Case management under channeling. Health Services Research 23(1):67-81, 1988.

52.Rabiner, D.J. The relationship between program participation, use of formal in-home care, and satisfaction with care in an elderly population. Gerontologist 32(6):805-812, 1992.

53.Rabiner, D.J., Mutran, E., and Stearns, S.C. The effect of channeling on home care utilization and satisfaction with care. Gerontologist 35(2):186-195, 1995.

54.Rabiner, D.J., Stearns, S.C., and Mutran, E. The effect of Channeling on in-home utilization and subsequent nursing home care: a simultaneous equation perspective. Health Services Research 29(5):605-622, 1994.

55.Skellie, F.A., Mobley, G.M., and Coan, R.E. Cost-effectiveness of community-based long-term care: current findings of Georgia's alternative health services project. American Journal of Public Health 72(4):353-358, 1982.

56.Stark AJ, Gutman GM, McCashin B. Acute-care hospitalizations and long-term care: an examination of transfers. Journal of the American Geriatrics Society 1982;30(8):509-15.

57.Stark AJ, Gutman GM. Client transfers in long-term care: five years' experience. American Journal of Public Health 1986;76(11):1312-6.

58.Starrett RA, Rogers D, Walters G. Home health care utilization: a causal model. Home Health Care Services Quarterly 1988;9(4):125-40.

59.Temkin-Greener H, Meiners MR. Transitions in long-term care. Gerontologist 1995;35(2):196-206.

60.Thornton, C., Dunstan, S.M., and Kemper, P. The evaluation of the National Long Term Care Demonstration. 8. The effect of channeling on health and long-term care costs. Health Services Research 23(1):129-142, 1988.

61.Trisolini MG, Thomas CP, Cashman SB, Payne SM. Resource utilization in home health care: results of a prospective study. Home Health Care Services Quarterly 1994;15(1):19-41.

62.Tsuji, I., Whalen, S., and Finucane, T.E. Predictors of nursing home placement in community-based long-term care. Journal of the American Geriatrics Society 43(7):761-766, 1995.

63.van Rossum, E., Frederiks, C.M., Philipsen, H., Portengen, K., Wiskerke, J., and Knipschild, P. Effects of preventive home visits to elderly people. BMJ 307(6895):27-32, 1993.

64.Wan, T.T., Weissert, W.G., and Livieratos, B.B. Geriatric day care and homemaker services: an experimental study. Journal of Gerontology 35(2):256-274, 1980.

65.Weissert, W.G, Wan, T., Livieratos, B., and Katz, S. Effects and costs of day-care services for the chronically ill: a randomized experiment. Medical Care 18(6):567-584, 1980.

66.Weissert, W.G. and Cready, C.M. A prospective budgeting model for home- and community-based long-term care. Inquiry 26(1):116-129, 1989.

67.Weissert, W.G. The national channeling demonstration: what we knew, know now, and still need to know. Health Services Research 23(1):175-187, 1988.

68.Weissert, W.G., Wan, T.T., Livieratos, B.B., and Pellegrino, J. Cost-effectiveness of homemaker services for the chronically ill. Inquiry 17(3):230-243, 1980.

69.Williams, ME, Williams, TF, Zimmer, et al. How does the team approach to outpatient geriatric evaluation compare with traditional care: a report of a randomized controlled trial. Journal of the American Geriatrics Society 35(12):1071-1078, 1987.

70.Wooldridge, J. and Schore, J. The evaluation of the National Long Term Care Demonstration. 7. The effect of channeling on the use of nursing homes, hospitals, and other medical services. Health Services Research 23(1):119-127, 1988.

71.Zimmer, J.G., Groth-Juncker, A., and McCusker, J. A randomized controlled study of a home health care team. American Journal of Public Health 75(2):134-141, 1985.

IV. Methodology

In this chapter variables used to analyze utilization of home care services are defined. To begin with, the selection and specification of each variable is discussed. In Section B, the following section, the statistical methods used to examine the relationships between the dependent and independent variables are described. This includes the Kaplan Meier univariate, stratified Kaplan Meier, and Cox Proportional Hazards multivariate statistical methods. Testing for underlying assumptions and hypothesis testing in these contexts will also be included. This section includes the statistical methods for unobserved variables. In section C, the limitations and delimitations of this analytical approach are summarized. This includes a discussion of the validity of observational data, definition of an episode of home care, unit of measurement of utilization, unobservable variables, self-selection bias, and competing hazards. Finally, in Section D, procedures for collecting the data, pre-processing of data, and sample selection are provided. This will include analysis of variance testing for generalizability of the sample and a descriptive analysis of the variables.

A. Specification of Variables

The determination of utilization of home care services in an observational or non-experimental context requires (1) specification of the confounding variables and relationships between them, and (2) addressing self-selection bias. The analysis will be constrained by the data and analytical methods currently available. In turn, the analytical method is constrained by the specification of the dependent variables and the hypothesized relationships between the dependent and independent variables. As this is a utilization study, the primary variable of interest is home care resource consumption.

In this study, resource consumption is measured by the total amount of home care services utilized as an alternative to institutionalization. Total resource consumption then is determined by predisposing, enabling and need characteristics. A well-known problem with this specification is that total resource consumption is a function of intensity of resource utilization and duration of utilization. Furthermore, the length of service primarily determines total resource consumption. The second problem is that resource consumption is censored. Between May 1991 to December 1995 total resource consumption is not observed due to (1) the fact that some individuals have not been discharged from the homecare program, and (2) the fact that some individuals have been discharged for other reasons (e.g. death and no longer requires service).

Both of these issues are addressed by specifying the length of service as the dependent variable and resource intensity as a covariate. In other words, total resource consumption is decomposed into its two components. In other words, total resource utilization is the product of length of service (censored) and <u>average</u> resource utilization intensity. It then follows that length of service can be written as a function of average resource utilization intensity. The Cox Proportional Hazards model is an appropriate statistical technique for this model specification. Censored data and statistical modelling issues will be discussed in the next section.

The independent variables were chosen on the basis of the literature review conducted in the previous chapter, and data availability as constrained by data quality. The final section of this chapter describes the homecare database structure and the resulting practical limitations in variable definitions. Grouping of the variables is in accordance with Andersen's model as discussed in the previous two chapters: predisposing, enabling, need, external environment, and the health care system.

1. Dependent Variables

Self-selection bias in the non-experimental context refers to the fact that the intensity of resource utilization is endogenously determined. In an experimental setting, the level of services are determined exogenously (i.e. independently of unobserved client characteristics). Ordinary least squares regression estimates are biased when a regressor is correlated with the error term. Hence the requirement for the specification of three endogenous (i.e. dependent) variables: (i) length of service, (ii) formal service intensity, and (iii) informal service intensity.

In the multiple linear regression model, two-stage least squares or the method of instrumental variables can be used to address this problem. In the Cox Proportional Hazards model, this is addressed through a fitted variable approach. First, separate linear regressions of formal and informal service intensity are performed on appropriate regressors. Secondly, predicted variables from these two equations act as covariates to explain variations in length of service. Hence formal service intensity is both a dependent and independent variable, while predicted formal service intensity is strictly an independent covariate in the Cox model. The same holds true for informal service intensity.

The dichotomization of service intensity reflects the fact total costs are comprised of direct and indirect costs. Direct costs refer to the expenditures for professional care of the client. Indirect costs for the home care population are mainly comprised of caregiver time. These are significant and must be taken into account as they are predicted to influence both the requirement for, and the duration of home care service consumption. Caregiver time reflects an opportunity cost of the caregiver's foregone activities.

a) Length of Service

Length of service (LOS) is defined as the number of days from admission to the homecare program until the time of discharge to a nursing home (institutionalization). Obviously, the definition of an episode of home care service is relevant for the calculation of this variable. The rationale for the adopted definition is discussed in section C.2 of this chapter.

b) Formal Service Intensity

Formal service intensity (FS) refers to the average observed or predicted level of homecare service provision over the time period in which services were actually provided (length of service). It may be specified in units of hours, dollars (1995 Canadian), or number of visits per day. Formal services may also be disaggregated by type of service or provider.

Predicted formal service intensity is obtained through regressing actual formal service intensity on health care system variables while controlling for predisposition, informal enabling and need variables. The exact functional form is specified in the next section and the results discussed in the next chapter. The purpose of this auxiliary equation is twofold. Firstly, interpretation of the significance, sign and magnitude of the coefficients is important as formal service intensity is itself a utilization variable. Secondly, within the context of the primary regression (explaining variations in length of stay, LOS, in the home care program), the level of services consumed by a client is not wholly explained by observed variables. This forms the basis for hypothesizing the existence of self-selection bias. It is assumed that the health care system variables are correlated with these unobserved variables and at the same time not correlated with the error term in the primary regression. Regressing LOS on the predicted formal service intensity should correct for the phenomenon of self-selection.

c) Informal Service Intensity

Informal service intensity (IS) is defined by two variables. This first is an observed, categorical variable specifying the category of informal service intensity. The second is calculated (later this chapter, section B.5) and expressed in the same units as formal service intensity. Informal services are defined as the level of indirect costs or resource consumption for which no payment is observed. This is limited to the level of provision or availability of caregiver support. Foregone client's time (labour and/or leisure losses) associated with variable home care service intensity is not included in indirect costs due to the absence of data.

Caregiver support levels and requirements are summarized by HCIS, the homecare informal support score. It is an ordered ordinal variable (1-5) which summarizes the level of informal support for thirteen informal support indicators (0-4 each). The five levels are defined in Table 4.1.

Table 4.1: Informal Support Categories and Definitions

Level	Definition
1	Informal support is not required.
2	All or almost all of the informal support is available.
3	Most (>½) informal support is available.
4	Some (<½) informal support is available.
5	None or very little informal support is available.

This categorization reflects both the requirement and availability of caregiver resources by and to the client. Requirement and availability are not directly comparable. However, if category 1 is interpreted as having all the informal support available, then these levels could be ordered. When entered in a regression as a continuous variable, an underlying assumption is made that the differences between each consecutive pair of categories is equal to the others.

This assumption would be supported if the assessor mentally identified increasing levels of informal support in similar increments of caregiver time.

Availability of informal support is made in reference to the level of need by the client in each of thirteen categories. If the goal of the homecare program is to provide an equitable level of support for each client by level of need, then it can be assumed that for each level of need, a unique level of formal plus informal services are provided. One may infer that given a level of need and formal resource intensity, that the remainder is met by the availability of informal services. This forms the basis for the calculation of informal service intensity weights. These weights quantify the differences in consecutive pairs of informal service levels and convert an interval scale into a ratio scale. This statistical procedure is outlined and discussed in the next section.

As with formal services, informal service levels are also subject to self-selection bias. In addition, informal services are postulated to be subject to measurement error. Both of these problems are addressed through the fitted variable approach. Again, informal service intensity is regressed on external environment variables (socio-economic variables) while controlling for predisposition, formal enabling, and need variables. LOS is then regressed on predicted informal service intensity.

2. Independent Variables

Determination of utilization of home care services requires identification of the determinants or independent variables through examination of their significance, sign and magnitude within the contextual and statistical frameworks. The contextual framework, the Andersen model, is discussed in Chapter 2. This framework sets out the nomenclature of independent variable groups and the purported causal relationships between them.

Figure 4.1 summarizes the entire model, variables, and causal relationships. The solid arrows indicate the direction of the hypothesized causal relationships within the model. The dashed lines identify independent variable groupings. All the independent variables in this diagram are discussed in this section of the paper.

Coefficient sign and magnitude are secondary considerations in a study of utilization.[1] Tables 4.2 and 4.3 summarize the study hypotheses as expected effects of the independent variables. Table 4.2 summarizes the effects on length of service. With LOS it is important to note that the coefficient signs are dependent on the competing hazard of interest. (Competing hazards were introduced in the last chapter and will be discussed in greater depth later.) The reason for this phenomenon is that the competing hazards are like apples and oranges. High-risk clients are institutionalized or die while low risk clients are often discharged to the community. In fact, the hazard "other discharges" is dominated by the two categories: can manage care and no longer requires services (see Table 4.4, IV.C.5.). For many covariates, the effect is indeterminate as indicated by a '0'-effect size.

[1] Sign and magnitude are secondary to levels of significance. However, as argued before, the hypothesis should reflect current knowledge (i.e. literature) and the purpose for which home care is provided.

Table 4.2: Expected Effect on Length of Service (LOS) of Independent Variables

VARIABLE	HAZARD			
	INSTITUTIONAL-IZATION	DEATH	OTHER DISCHARGES	ALL DISCHARGES
Age (AGE)	-	-	+	-
Gender (SEX)	+	+	0	+
Formal Service Intensity (FS)	+	+	0	+
Homecare Classification Score (HCS)	-	-	+	-
Homecare Informal Support Score (HCIS)	-	-	0	-
Homecare Functional Need Score (HCFN)	-	-	+	-
Informal Support Service Intensity (IS)	+	+	0	+
Physical Function (PFn)	+	+	-	+
Mental Function (MFn)	+	+	-	+
Weighted Comorbidity Index (WCI)	-	-	0	-

Table 4.3 summarizes the expectations for the two auxiliary, service intensity models. Following Table 4.3 is a description of the construction of each variable.

Table 4.3: Expected Effect on Service Intensity (FS and IS) of Independent Variables

VARIABLE	AUXILIARY MODELS	
	FORMAL SERVICES (FS)	INFORMAL SERVICES (IS)
Age (AGE)	+	-
Gender (SEX)	0	0
Formal Service Intensity (FS)	n/a	-
Informal Support Service Intensity (IS)	-	n/a
Physical Function (PFn)	-	0
Mental Function (MFn)	-	0
Weighted Comorbidity Index (WCI)	+	+
Capital Health Authority / Edmonton Board of Health (CHA)	0	n/a
Nursing Home Waitlist (WTL)	0	n/a
Ratio of Urgent to Total Referrals (RUT)	0	n/a
English Speaking (ENG)	n/a	-
Married or Common Law (MS)	n/a	+
Live With Spouse or Others (LIV)	n/a	+
Family Size (FAM)	n/a	+
Income (INC)	n/a	+

a) Predisposing Characteristics

There are two predisposing variables in this model: clients age at time of admission (AGE), and gender (SEX). Age and gender are hypothesized to predict variations in length of service and service intensity. Age is a continuous variable calculated as the difference in days between the date of birth and the date of first assessment divided by 365.25. Gender is a dichotomous variable with females coded as '1' with males coded as '0'.

b) Enabling Resources

There are two variables which summarize enabling resources in the model: formal support (FS) and informal support (IS). They are hypothesized to explain variations in the length of service. They are however endogenous or determined by other factors [see sections A.1.a (ii) and A.1.a (iii)]. Endogeneity is characterized by a correlation between the independent variable and the error term resulting in inconsistent estimates.

The most desirable method for reducing the probability of inconsistent estimates, the method of instrumental variables, is not possible in the context of a Cox proportional hazards model. Another candidate method that addresses this issue is called the fitted variable approach where the predicted value of the variable is used in the regression. Hence, LOS is regressed on predicted formal support ($\overline{FS}$) and predicted informal support ($\overline{IS}$), thereby dealing with the endogeneity issue.

c) Need Variables

Need is defined by the client's functional status (HCS, HCFN, PFn, MFn) and disease burden (WCI). These variables are hypothesized to explain variations in length of service, plus formal and informal service intensity. The functional status measures represent different levels of aggregation of the need concept.

The combined homecare classification score (HCS) is a casemix variable which controls for diagnosis, severity of illness, and other predictors of resource utilization. It is an ordered ordinal variable that goes from '0', the level associated with the lowest predicted resource consumption, to the highest, '9'. It comprises both enabling resources and level of need in that it is the sum of both HCIS and HCFN.

The homecare functional need score (HCFN) is an ordered ordinal variable summarizing the levels of need as captured by thirteen indicators of functional status. It also is an ordered ordinal scale going from lowest need = '1' to highest need ='5'.

A separate principal components analysis was performed on the thirteen indicator variables resulting in the construction of two variables: physical function (PFn) and mental function (MFn). These continuous variables have been normalized so that the average for the homecare clients is 0. A +1 indicates one standard deviation above average functional status and -1, one standard deviation below average functional status. The construction and weights of these two indices are provided in IV.B.4.

Disease burden is measured by a (modified) weighted Charlson index of disease comorbidity (WCI). The index is a ratio scale from 0 to 10. Its derivation is likewise presented later in IV.B.3.

d) External Environment

The external environment is characterized by sociological variables over which the individual or the health care system has little control. These are also referred to as socio-economic variables. Their influence on LOS in this model is hypothesized to be indirect in that they are determinants of informal service intensity (IS), which in turn influences length of service. The database allows for analysis including 5 external environmental variables: English speaking (ENG), marital status (MCL), living arrangement (LSO), family size (FSZ), and income (INC).

Upon admission to the homecare program, clients' preferred primary and secondary languages are recorded. English speaking (ENG) is coded as a binary variable where 1 signifies the affirmative, 0 is all other languages, and 9 signifies missing information.

Marital status is identified by one of six categories. For purposes of this study however, these have been recategorized by method of judgement into a binary variable as listed in Table 4.4. This re-categorization is also known as a "crosswalk.

Table 4.4: Marital Status Crosswalk

Marital Status (Original)	**Marital Status (Study)**
Married Common Law	MCL = 1
Single Separated Divorced Widowed	MCL = 0
Missing	MCL = 9

Likewise, the living arrangement (LSO) has been reclassified using judgement as listed in Table 4.5.

Table 4.5: Living Arrangement Crosswalk

Living Arrangement (Original)	**Living Arrangement (Study)**
Alone	LSO = 0
With Spouse Only With Spouse and Others With Others	LSO = 1
Missing	LSO = 9

Family size, FSZ, is treated as a continuous variable where the number reflects the combined number of adults and children in the family unit. They need not reside at the same location. There can be one or two adults and there is no distinction made above 6 children. For example, two adults and 7 children are still coded as an '8.'

Income (INC) has been recalculated as a continuous variable from discrete income categories using median values.

e) Health Care System

The health care system variables are those which indirectly influence length of service through formal service intensity. They act as constraints in the homecare managers' decision as to the appropriate level of service intensity provision. The hypothesized constraint is most likely to act in the diminishment of services due to temporary bottlenecks. These bottlenecks occur because of the uncertain timing of the demand for home care services. However, rapid expansion of services may also result in a higher than predicted level of services based on predisposing and need variables alone.

Health care system variables are typically not incorporated into an analysis where the individual is the unit of analysis. In this study however, three variables have been constructed which translate the effects of the health care system into individually observable influences. They are a variable signifying the effects of urgent to total institutional referrals (RUT), a nursing home waitlist variable (IWL), and a dummy variable signifying a change in management of the program (CHA).

Referrals to the nursing home care (continuing care program) are classified as either regular or urgent. Raw data and calculated variables are provided in Table 4.6 for the 21-month period beginning April 1994. The homecare database did not record individual discharges as regular or urgent. Neither was the intention to discharge recorded separately from the actual discharge date. As a result, the imputed influence of referrals to institutional facilities is through caseworker and office congestion. That is to say that if the ratio of urgent to total referrals (RUT) is high in the months immediately prior to the discharge of a client, then it is assumed that the client's discharge has been postponed relative to the intended date. In other words, office and caseworker congestion leads to an increased length of service relative to the ideal situation. Ratios of urgent to total referrals vary from 26% to 58% over this time period. Since the hypothesized effect takes place nearer the end of the episode, an individual RUT is calculated as a geometric average.

Institutional waitlists reflect the fact that individual clients receive homecare services for a longer period of time than intended. The continuing care community waitlist reflects the number of individuals who are waiting for entrance to a skilled nursing facility. Not all are from the homecare program, and some homecare clients may be on a waitlist outside the community. However, the waitlist reflects the disparity between demand and supply of nursing home beds. Dividing the actual waitlist number by a geometric-weighted 21-month average creates a monthly variable. Over this time period the waiting list has varied from 82% to 119%.

Table 4.6: Calculation and Assumptions for System Variables

	Regular Continuing Care Referrals	Urgent Continuing Care Referrals	Ratio of Urgent to Total Referrals RUT	Continuing Care Community Waitlist	Normalized Continuing Community Waitlist IWL	EBH/CHA Dummy CHA
Apr-94	27	15	0.36	179	0.99	0
May-94	37	13	0.26	157	0.87	0
Jun-94	24	11	0.31	164	0.91	0
Jul-94	15	16	0.52	178	0.98	0
Aug-94	26	26	0.50	210	1.16	0
Sep-94	28	15	0.35	216	1.19	0
Oct-94	21	23	0.52	184	1.02	0
Nov-94	31	26	0.46	165	0.91	0
Dec-94	26	36	0.58	148	0.82	0
Jan-95	23	29	0.56	166	0.92	0
Feb-95	37	26	0.41	161	0.89	0
Mar-95	55	40	0.42	162	0.90	0
Apr-95	46	34	0.43	182	1.01	1
May-95	67	37	0.36	183	1.01	1
Jun-95	44	28	0.39	184	1.02	1
Jul-95	63	36	0.36	187	1.03	1
Aug-95	48	38	0.44	189	1.04	1
Sep-95	63	33	0.34	194	1.07	1
Oct-95	71	36	0.34	183	1.01	1
Nov-95	58	36	0.38	207	1.14	1
Dec-95	53	25	0.32	201	1.11	1

Management of the homecare program changed during the study period. On May 1, 1995, the Capital Health Authority took over responsibility of the homecare program from the Edmonton Board of Health. Restructuring and consolidation of separate health care programs and jurisdictions into a single health authority may result in a different allocation and delivery of home care resources. The variable CHA is an arithmetic average, thereby taking a value of 1 when an episode of home care falls entirely within the CHA's management period. A value of 0 is assumed when the episode falls entirely under the EBH management. Values between 0 and 1 reflect varying proportions of time between the two regimes.

B. Statistical Methods

This section will describe the statistical methods used for variable manipulation and modelling. I discuss the statistical methods used to examine the relationships between the dependent and independent variables. Since the main dependent variable is censored, the Kaplan Meier univariate, stratified Kaplan Meier, and Cox Proportional Hazards multivariate statistical methods are used. Testing for underlying assumptions and hypothesis testing in these contexts will also be included. Finally, this section includes the statistical methods for self-selection bias, as the data are observational and not experimental.

1. Kaplan-Meier Survival Analysis

Kaplan and Meier (1958) describe a popular method to analyze time to event data such as length of service considered in this study. This technique is predicated on the fact that at the end of a study period, that a proportion of events have not occurred. A basic assumption of the Kaplan-Meier (KM) technique is that the censored individuals do not differ from those individuals for which an event has occurred. The primary output of the technique is the probability of survival at each point in time; in this case, the probability of receiving home care services. This is also known as the cumulative survival function and differs from lifetable analysis in that the time intervals are not fixed.

The Kaplan-Meier cumulative survival at time t is calculated as a product of survival at time t by the survival at time $t-1$ by the survival at time $t-2$ until the first period. By assumption, the survival at time 0 is equal to 1 (i.e. all individuals start out alive). For each time period t, if n_t individuals are at risk of an event and d_t events are observed during the time interval t, then cumulative survival at time t is given by the following: $S(t) = \prod_t \left(1 - \frac{d_t}{n_t}\right)$.

Mean survival time and 95% confidence intervals of this cumulative survival will be reported in the next chapter.[2] The mean survival time is calculated using special techniques in SPSS. Assuming that the survival estimates are normally distributed, confidence intervals are calculated by the formula S(t)+/-1.96×se[S(t)], where se[·] refers to the standard error. There are several methods by which the standard error can be estimated: Greenwood (1926), Peto

(1984), and Rothman (1978). Greenwood's estimate is accurate when S(t) is relatively large for the time period of interest (as is the case in this dissertation) and is the one used by the SPSS program. The standard error formula is given as follows: $se(t) = S(t)\sqrt{\sum_t \frac{d_t}{n_t(n_t - d_t)}}$.

Univariate Kaplan-Meier analysis is used to elucidate the influences of individual explanatory variables on the cumulative survival function, in this case - length of home care service. Separate survival curves are calculated for individual values of discrete variables. This necessitates the transformation of continuous into categorical variables. Although the choice of the numbers and definitions of categories is somewhat arbitrary, dichotomization of continuous variables was performed such that roughly equal numbers of events was observed for each group to ensure smooth survival curves.

The significance of individual explanatory variables in their ability to explain variations in survival times may be formally tested with the logrank (Mantel-Cox), the generalized Wilcoxon (Breslow), or the Tarone-Ware tests. The test basically depends on assessing the relationship between expected (E_t) and observed (O_t) events between k categories for each period in time. The generic test statistic may be written as $\sum_t w_t(O_t - E_t) \sim \chi^2_{k-1}$ where w_t refers to different weights used by the three test statistics. The logrank test statistic assumes that $w_t=1$ and is reported alongside the mean survival times and confidence intervals.

Stratified Kaplan-Meier analysis is used to elucidate the influences of the primary independent variable of interest, intensity of formal resource provision (cost per day), on

[2] The median survival times are not available in this study as the time which satisfies the expression

survival stratified by each of the other independent variables. Mean survival time, the standard error and the 95% confidence intervals will be reported for each individual stratum. The categorical definitions are the same used as in the univariate analysis.

2. Cox Proportional Hazards Model

Before model specification is addressed, some background concepts are discussed. Model specification and model selection procedures determine the final functional form used for analysis and discussion purposes. Finally, the most important assumption of proportionality is discussed and the procedures for testing this assumption are outlined.

a) Background

The Cox (1972) proportional hazards model allows for the simultaneous identification of multiple determinants of length of service. A second feature of the model is that both categorical and continuous variables can serve as independent variables. A cumulative survival function can also be generated (as with Kaplan-Meier analysis). However, this method also generates an instantaneous failure rate otherwise known as the hazard rate.

The hazard rate is the probability of an event at time t having survived until t. The relationship between cumulative survival and the instantaneous hazard may be illustrated with the introduction of a probability density function, $\phi(t)$, which is the instantaneous probability of

---

$S(t)=0.5$ exceeds the time period over which homecare clients are observed.

an event (between 0 and 1). The hazard is a rate per unit of time and therefore can exceed one. Although survival and hazard are related concepts, they are useful for the conveyance of different results. The cumulative survival function is useful for the comparison of one or more different groups of clients. Figure 4.2 illustrates this relationship. The hazard function is useful for communicating the rate of failure at any point in time (Parmar and Machin 1995).

According to the Cox proportional hazards model, specification of the survival, S(t), and hazard, h(t), functions incorporating multiple covariates. The survival function is given by the following expression: $S(t) = [S_0(t)]^{e^{\beta X}}$ where βX is a vector, i.e. $\beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k$. The hazard function can also be expressed as a function of βX, $h(t) = [h_0(t)]e^{\beta X}$. $S_0(t)$ and $h_0(t)$ are nuisance parameters referred to as baseline survival and hazard rates. They are not of interest as they are equal to the survival and hazard rates when all the covariates equal zero (typically a meaningless level). Since h(t) is simpler than S(t), the Cox model is usually expressed as a hazard equation. It is a semi-parametric model since there is no distributional assumptions placed on the "nuisance" parameters, $S_0(t)$ or $h_0(t)$.

b) Model Specification

Model specification is the process of determining which covariates are included in the model and the relationship between those variables. The relationship is constrained by censorship and the nature of the Cox model as discussed. The number of covariates is constrained by the nature of the Cox model and the number of observations. There are several rules for determination of the maximum covariates in a Cox model. The rule used here is that the number of covariates does not exceed the fourth root of the sample size (Parmar and Machin 1995). In this case this equates to 7 or 8 covariates.

There are four candidate models that are estimated as illustrated in Figure 4.3. They are not nested models, as they do not reflect merely the consideration of additional variables. Rather they represent different degrees to which additional model assumptions are incorporated. Model A encompasses the fewest modelling assumptions. The only calculations performed are those used to create a weighted comorbidity index. Progression to Model B merely reflects the fact that overall case-mix can be disaggregated into its two components.

The three arrows between Models B and C represent additional assumptions. The first arrow (from HCIS to Informal Support) represents an additional model and assumptions required for converting the categorical scale of informal support to the same units as formal support. The second and third arrows represent a factor analysis of the thirteen indicator variables that are used to construct the functional status variable (HCFN).

The two arrows between Models C and D represent the incorporation of fitted (i.e. predicted) variables to deal with self-selection bias. A structural equation diagram for Model D is illustrated in Figure 4.4. The arrows represent causal relationships between individual variables. The three error terms represent the incorporation of uncertainty into the model and the relationship between the three equations.

Selection bias is addressed in Models 'D' through a two-stage regression technique illustrated in Figure 4.4. Functional status and comorbidity in model 'D' are calculated using home care data based on observations at the beginning of the episode.

The three regression equations illustrated in Figure 4.4 are written as follows:

E1 $h_m(t) = [h_0(t)]e^{\beta X}$ where $\beta X = \beta_1 AGE + \beta_2 SEX + \beta_3 FS + \beta_4 IS + \beta_5 PFN + \beta_6 MFN + \beta_7 WCI + \varepsilon_1$ where $m = 1$ (institutionalization), 2 (death), or 3 (other discharges)

E2 $FS = \delta_0 + \delta_1 AGE + \delta_2 SEX + \delta_3 IS + \delta_4 PFN + \delta_5 MFN + \delta_6 WCI + \delta_7 RHA + \delta_8 WTL + \delta_9 RUT + \varepsilon_2$

E3 $IS = \gamma_0 + \gamma_1 AGE + \gamma_2 SEX + \gamma_3 IS + \gamma_4 PFN + \gamma_5 MFN + \gamma_6 WCI + \gamma_7 ENG + \gamma_8 MS + \gamma_9 LIV + \gamma_{10} FAM + \gamma_{11} INC + \varepsilon_3$

Equation E1 specifies the hazards of discharge from the home care program. Cox's proportional hazards model (Cox 1972) is used to estimate E1 as the dependent variable is time-to-event. For purposes of this study the four competing hazards are death, discharge to a nursing home, discharge due to other reasons, and total discharges (d, i, o, and t respectively). Equations E2 and E3 are average cost equations used only for model 'D'. These two equations will be estimated with ordinary least squares (OLS). Predicted average costs or service intensities from these two equations then become regressors in equation E1, also known as the fitted variable approach, a technique of instrumental variable estimation. Instrumental variable estimation is well known in econometrics and is increasingly becoming recognized in health services research (Zellner 1970, Pagan 1984, Zohoori et al 1997).

Construction of a weighted comorbidity index, factor analysis, construction of informal support, and self selection bias are each dealt with later in this chapter (B3-B.6).

c) Model Selection

Although the models are not nested, test statistics based on the likelihood function are used determine which of the four models explains variations in client utilization of home services the best. The likelihood ratio test statistic and the score (also known as the global chi-square) statistic are reported for each model. The minimum, maximum, median and mean statistics are then compared so that some comparisons can be made between the models.

d) Proportionality

The hazard rate illustrates a very important assumption, that of proportionality. By assumption and for any two individuals, the ratio of hazards is a constant (i.e. independent over time). Since this is an important assumption of the Cox model, it should be tested. The most common tests for proportionality are graphical in nature, however more formal tests exist as well.[3] Four graphical tests are employed to test this assumption: stratified survival, log-minus-log (LML) survival, Schoeldfeld's (1982) partial residual, and Therneau's (1990) martingale residual plots. All of these tests were performed due to the importance of the proportionality assumption and the fact that violation of this assumption would require another modelling strategy.

[3] Formal testing requires modelling covariates as varying with time. Due to the large database and/or model complexity, this was not possible despite using SPSS and a Pentium computer with 64Meg.

The stratified Cox model can be written as $h_i(t) = [h_{i,0}(t)]e^{\beta X}$, where i refers to the stratum. In essence, a separate survival function is estimated for each stratum. Both the cumulative survival and ln[-ln] plots are examined for parallelism. Partial residuals are plotted against time and proportionality results in an equal dispersion around zero. Martingale residuals are plotted against covariate values and Xβ with the equal expectation of equal dispersion around zero.

3. Creation of a Weighted Comorbidity Index

The Charlson (1987) weighted comorbidity index was developed to control for confounding illnesses. The original index was developed to refine a prognostic index for 30-day in-hospital mortality. It can be said that the primary illness in this client population is old age. Also, rather than mortality, the outcome measure for this study is utilization. Therefore the weights used in the original study are not generalizable to this population.

Some studies simply add up the number of comorbid conditions on the basis of simplicity. This method simply assumes that each condition affects utilization equally. The weights used in the comorbidity index developed here were developed using a Cox model and dummy variables for each of the disease conditions. No other covariates were included in the equation. The weights were taken as the ratios of the β's as opposed to ratios of hazard ratios as was performed in Charlson's initial paper. Due to the nature of the Cox regression β's are additive, hazard ratios are not.

The following table summarizes the diagnostic codes, prevalence, β values, p-values, and the derived modified weights. Insignificant diagnoses were assigned a weight of zero. Rounding of the weights was also conducted. For comparison purposes, the corresponding Charlson weights are also included.

Table 4.7: Derivation of Modified Charlson Weights

Diagnoses	Name	Prevalence (%)	Beta	p-value	Modified Weights	Charlson Weights
D01	Infectious and Parasitic Diseases	0.9	0.2194	0.3240	0	0
D02	Neoplasms	6.0	0.4935	0.0000	3	6
D03	Endocrine, Nutrition & Metabolism	13.7	0.0952	0.1563	0	0
D04	Blood & Blood Forming Organisms	1.2	0.1933	0.3298	0	0
D05	Mental Disorders	10.6	-0.0021	0.9801	0	1
D06	Nervous System & Sense Organs	28.3	0.1559	0.0089	1	1
D07	Circulatory System	37.6	0.1724	0.0011	1	1
D08	Respiratory System	11.2	0.2006	0.0055	1	1
D09	Digestive System	5.3	0.0005	0.9958	0	1
D10	Genito-Urinary System	5.4	0.0923	0.3627	0	0
D11	Pregnancy/Childbirth & Puerperium	0.0	0.9492	0.3432	0	0
D12	Skin & Subcutaneous Tissue	4.4	-0.0184	0.8759	0	0
D13	Musculoskeletal System	36.3	-0.3128	0.0000	-2	0
D14	Congenital Anomalies	0.6	-0.2487	0.4363	0	0
D15	Perinatal Period	0.0	-7.6811	0.8795	0	0
D16	Ill Defined Symptoms	1.2	0.1241	0.5600	0	0
D17	Injury and Poisoning	2.3	-0.4807	0.0114	-3	0
D18	Other Reasons	5.2	0.0980	0.3507	0	0

The index is then constructed as five plus the sum of the weights for each additional diagnosis (up to three). The reason for adding five is due to the fact that two diagnoses correspond to negative weights. Injury and muskuloskeletal disorders are in fact associated with longer lengths of stay than individuals that do not have any comorbid conditions. This means that the risk of discharge is less for these clients. Since there are only three diseases recorded (practical limitation), the index goes from a minimum of zero to a maximum of ten.

4. Factor Analysis and Functional Status

The categorical, functional status index (HCFN) is composed of thirteen constituent components. These are illustrated in the bottom left of Figure 4.4. Inclusion of all of these indicators would violate the rule for the maximum, recommended number of covariates. Factor analysis is a data reduction method that accomplishes two tasks. Firstly, the categorical variable may be converted into a continuous score. More importantly, the indicator variables can be used to create a more manageable number of variables and yet provide more information than HCFN alone. Factor analysis consists of two components: exploratory and confirmatory factor analysis.

Exploratory factor analysis consists of determining the "natural" number of factors that explain the correlation between the thirteen indicator variables. This is done with a scree plot that indicates that this dataset contain only two factors. Secondly, a test statistic indicating the merits of factor analysis is printed out by SPSS. The test statistic exceeds 0.9 and is deemed 'meritorious' meaning that factor analysis is supported by the sample. Both the scree plot and the rotated factor solutions are illustrated in figure 4.5a-b.

Confirmatory factor analysis is the calculation of weights for each indicator given the number of factors. Inspection of the weights indicates that the indicator variables are grouped into two distinct groups. The first group consists of memory, coping, and the potential for self-injury. This group, consisting of points 3, 8 and 9 in figure 4.5.b is then named mental function. The remaining ten indicators are physical in nature and hence named physical function.

5. Informal Support Calculations

Converting informal support from a categorical measure to a continuous variable is desirable since interpretation of the results will be simplified in resource consumption is denominated in the same units as formal support. Equally important is the fact that there are the problems with the scale itself, as discussed in section A.1.c of this chapter.

Firstly, personal care and home support, formal resource intensity was regressed against four informal service dummy variables: IS2, IS3, IS4, and IS5. Informal support was not deemed to be a substitute for assessment, case management or direct service provision. The OLS equation coefficients were then used to estimate the value of the differences between each corresponding category. Next the weights for IS1 were set to 1.0 and 0.0 for IS5. Intermediate weights were determined from the differences. Weights for each category was multiplied by the formal service value. The weights are provided in the Table 4.8.

Table 4.8: Informal Support Weights

	Personal Care	Home Support
IS1	1.0	1.0
IS2	0.9	0.9
IS3	0.8	0.7
IS4	0.6	0.5
IS5	0.0	0.0

For example, an individual identified as IS4 (Some, <½, informal support is available) with an average personal care intensity of 2 hours per day and 1 hour of home support per day would be deemed to have received $0.6\times2+0.5\times1=1.7$ hours per day of informal support.

Although the assumptions and method for the conversion are crude, this methodology was necessary, as the variable is an important one. Also, it is hoped that future collection of informal support takes these issues into account.

6. Self Selection Bias

Self-selection bias refers to inability to correctly identify the change in a dependent variable (length of service) associated with an independent variable (service intensity). In the context of a regression model, this is partly addressed through model specification. When significant explanatory variables are excluded, parameter estimates are biased. For example, in comparison to the Kaplan-Meier model, the Cox proportional hazards model corrects for the specification bias through the introduction of additional, confounding variables. Self-selection bias is however an additional problem normally associated with efficacy studies. It is nonetheless an important consideration in the identification of determinants of home care service utilization because service intensity acts like a treatment variable in a non-experimental context.

Regression parameter estimates have certain desirable properties. They should be unbiased, efficient and consistent. The fundamental problem associated with non-experimental data is that one or more independent variables are often correlated with the equation disturbance. As a result, regression estimates will suffer from consistency problems. Consistency refers to the fact that as the sample size increases, that the parameter estimates converge to their true (population) values. In econometric terminology, the probability limit of the parameter estimate, in the event of a regressor-disturbance correlation, does not equal the true value; $p\lim(\hat{\beta}) \neq \beta$. The bias is also referred to as regressor-disturbance correlations. When this happens, additional information is available that can be used to eliminate or reduce bias in the estimates of regression coefficients. This type of bias occurs most frequently in the context of self-selection.

There are two primary econometric strategies to introduce additional information: the control function and the instrumental variable estimators (Heckman et al 1985). In the econometric literature, the instrumental variables approach has a fairly long history. In the 1920's there appeared two articles (Working, 1927; Sewall and P.G. Wright, 1928) which incorporated instrumental variables in the estimation of a system of demand and supply equations. The term "instrumental variables" however was coined by Reiersol in 1941. Two years later, they were incorporated into a system of linear, simultaneous equations as we know them today. Sargan's (1958) article was also a substantial contribution to the estimation of structural relationships" through their use of instrumental variables. More recently, however, instrumental variables are finding a unique application in the health sciences literature.

Primarily in the context of multiple linear regression and the instrumental variable estimator, there are four sources of additional information: (1) fitted instruments, (2) proxies, (3) transformations of independent variables, and (4) ad hoc methods. Fitted instruments are variables constructed by a prior regression process. Proxies are variables that are correlated with the independent or right hand side (RHS) variables but not the error term. Transformations of independent variables may, for example, include variables such as X^2 and X^3. Finally, an index number is an example of ad hoc methods. Index numbers are constructed by sorting the variable(s) of interest and assigning an index consisting of consecutive integers. The primary requirement is that the instruments are highly correlated with service intensity but uncorrelated to the error term.

In the absence of evidence on econometric techniques for the correction of self-selection bias in the presence of censored data, the instrumental variable estimator approach is modified as follows.[4] Since the hypothesized source of the self selection bias are the two service intensity variables, they are regressed on a separate set of variables and these "fitted" or "predicted" variables are used in their place. It is hypothesized that the variables are "purged" of their correlation with the error term. As a result of this procedure, the usual tests of inconsistency (Wu-Hausman 1973,1978) and instrument admissibility (Hausman-Taylor 1981) are not available.

[4] A thorough search of the EconLit, Medline, and HealthSTAR databases yielded no techniques. I have also consulted with four statisticians at the University plus have posted the question to an internet discussion group with no additional advice as to what I have proposed here.

C. Limitations and Delimitations

1. Validity of Administrative Data

Use of data for research purposes requires that variables be defined and used unambiguously. This means that the clinician, data-entry personnel, systems analysts and investigators are in agreement as to how each variable was constructed and manipulated (LL Roos et al, 1996; Roos NP et al, 1988).

There is no threshold for quality assessment of data elements. Organizational constraints and practices require that the effort in collecting and maintaining data be in balance with the benefits of that data to the particular organization. Rather the investigator will determine how each factor in the collection and management of data affects the research question (LL Roos et al, 1989).

2. Definition of a Long-Term Care Episode

The episode definition plays an obvious but important role in defining the length of service (Hornbrook et al 1985). Less obvious perhaps is the fact that episode definition also plays a role in defining the resource intensity. Left censoring, the episode start date, and right censoring due to competing hazards are the three specific issues that require discussion.

Left censoring refers to the fact that some clients were already admitted to the homecare program and receiving services at the study start date: May 1, 1991. Three potential solutions are to exclude these cases completely, include these clients and assume that the May 1991 is their episode start date, or model the episode length for these clients using some form of discriminant analysis. Discriminant analysis is not feasible due to the large number of groups that would need to be defined and the fact that there are no observations in excess of 56 months. If May 1991 were the episode start date, then the probability of loss to follow up and other data quality issues is expected to increase when comparing clients admitted prior to May 1991. Lastly, the 1990's mark a period of rapid health care reform in Alberta. It is plausible that clients admitted prior to this period may differ or have received different services and/or service intensity. For these reasons, they have been excluded from the analysis. Further discussion on the inclusion and exclusion criteria is discussed in the last section of this chapter.

Candidates for the episode start date are referral dates to the homecare program, the date of admission, and the first assessment date (of functional status). This database does not include the referral data and therefore this definition could not be used. If this date were available, the waiting time to program admission and receipt of services could however have been used to explain concomitant resource utilization. The reason for consideration of this definition is that functional status is assumed to decline over time. Hence, the ability of the functional status to act as a covariate depended on its uniform measurement between clients. Ideally, functional status should be measured at time of admission and time of discharge with measurements taken at regular intervals of say 6-12 months. Only 18.7% of clients received more than one assessment, and of those clients, there was no clear trend in changes of functional status over time.[5] This problem is compounded if one postulates that the probability of a repeat assessment is correlated with either the initial level or the rate of change in functional status.

Measurement of functional status poses two additional problems. Firstly, the assessment tool has only been validated and in use from April 1, 1994. Secondly, not all individuals are assessed within one month of admission. The first problem overlaps with the issue of left censoring. At the time of first assessment, some individuals have been in the program for a long period of time as opposed to other clients who were assessed more readily. The second problem leads to the question as to why are some individuals assessed more quickly than others? If individuals higher at risk for institutionalization are assessed earlier, and/or more frequently, then a bias would result in that the higher risk individuals would have their length of service increased relative to their low risk counterparts.

Non-uniform service intensity during the course of an episode was an additional problem that precluded the first assessment date as episode start date. Until this time, the issues relating to functional status could be overlooked if one assumed that the decline in functional status occurred slowly enough. Service intensity however is not uniform and in fact is lower during the initial months. As seen in Figure 4.6, assessment and case management costs are higher during the first three months. Direct service provision, personal care and home support is lower for the first year. This means that resource utilization between the date of admission and the first assessment date does not reflect average resource consumption. Figure

[5] Of the 930/4962 clients who received more than one assessment during their episode, 62.8% experienced a decline in physical function and 46.2% experienced a decline in mental function. For both physical and

4.7 illustrates the overall trend in costs and demonstrates that it takes more than one year before equilibrium service intensity is achieved. This would bias the shorter episodes upward with respect to service intensity as compared to longer episodes.

The same point is illustrated with the participation rate. The participation rate is defined as the proportion of enrolled clients (in the home care program) that actually receive services during an episode month. This is illustrated in Figure 4.8. Participation in three first months exceeds 95% probably due to the high number of clients that are assessed during this time. Participation then drops to 84% for the next few months and slowly climbs back to equilibrium (90%). This diagram conjures two hypotheses. The first is that coordination of services and client schedules takes a long time to achieve. The second hypothesis is that clients are not admitted uniformly with respect to urgency of service requirements. If low participation were due to scheduling, then the date of admission does not bias the episode definition. If however, clients are not comparable on admission to an extent not documented with other variables, then neither the admission nor the first assessment date is a good candidate. Since the first hypothesis is more plausible, and an alternative definition isn't readily available for the second, the first assessment date is dropped as a candidate for the episode start date.

This leaves the date of admission as the episode start date. It has been uniformly defined for all homecare clients. It reflects the intended date of formal service provision. There are some services that may be provided prior to the date of admission related to provisional admission and preliminary assessments of functional status. This may occur for those clients

---

mental function, 13.2% remained the same. Physical and mental function improved 24.0% and 40.5% of who are at both high and low risk of institutionalization if appropriateness of home care is in question. These services were not captured but this was not considered to be a serious limitation.

Two additional issues arise from consideration of defining the episode start date as the date of admission. Firstly, individuals could be admitted to the short-term, long-term, or palliative homecare programs. Only long-term admissions were considered as they reflect the population at risk for institutionalization. Secondly, a number of individuals were admitted to the program more than once during the study period.[6] This may occur when the previous episode resulted in a discharge because services were no longer required. Only the last episode per client was used for the purposes of this study.[7]

3. Unit of Measurement of Utilization

The basic unit of observation is the number of hours each provider type spent for each type of service on each client. Formal service intensity is obtained by summing the number of hours across service and provider types for each client and dividing by each client's length of service (Equation 4).

---

the time respectively.

[6] Of 4962 clients, 3282 (66.1%) have experienced only a single episode, 1212 (24.4%) have experienced two episodes, with the remaining 468 (9.6%) experiencing up to 7 episodes per client.

[7] Therefore, the unit of observation and analysis is the individual client. Two alternative modelling strategies were considered. Firstly, the episode could characterize the unit of analysis. Secondly, a multiple-spell competing hazards framework that includes multiple admissions and discharges was possible. I felt that the additional complexity of either manipulating the database or considering another analytical method would overshadow the benefits of including the additional observations.

E4: $$FS_i = \frac{\sum_{j=1}^{5}\sum_{k=1}^{6} w_{jk} \times hr_{ijk}}{LOS_i},$$

where i = 1… 4,962 clients; j = assessment, case management, direct services, personal care, and home support; k = home care aide, licensed practical nurse, nurse, occupational therapist, respiratory therapist, and social worker; w_{jk} = a weighting factor; and hr_{ijk} = the number of hours. Setting $w_{jk} = 1 \ \forall\ j,k$ results in FS_i expressing the average hours of total services per day for each client.

4. Valuing Units of Utilization

I compiled a standard cost and visit lists on the basis of detailed observations. This standard cost list is in agreement with the standard cost list developed by Jacobs et al (1995). Formal service intensity is expressed in terms of average dollars, hours or number of visits per day. Standard costs and standard visits are expressed as the weights in the previous expression (w_{jk}). Table 4.9 lists all 16 provider-service combinations and the standard costs and hours/visit used for this study:

Weighting hours by dollars/hour reflects an opportunity cost expressed in terms of alternate activities by the service providers and is the primary measure used in economic analyses. Weighting hours by visits/hour is also standard practice in health economics and places increased emphasis on the distribution of resource consumption over time. For example, providing the client with all personal care services at once is impractical. Likewise, dividing the total allocation into very small units is also ineffective. A visit then, falls somewhere in-between and should reflect both economies of scale in terms of provision and be related to the level of the client's need.

Table 4.9 Formal Home Care Standard Cost and Visit List

Number	Provider Type	Service Type	Standard Cost ($/Hour)	Standard Visit (Hours/Visit)
1	Nurse	Assessment	$35.00	1.00
2	Nurse	Case Management	$35.00	0.25
3	Nurse	Direct Services	$35.00	0.50
4	LPN	Direct Services	$18.50	1.00
5	LPN	Personal Care	$18.50	4.00
6	Home Care Aide	Personal Care	$13.50	4.00
7	Home Care Aide	Home Support	$12.50	4.00
8	Occupational Therapist	Assessment	$35.00	0.50
9	Occupational Therapist	Case Management	$35.00	0.25
10	Occupational Therapist	Direct Services	$35.00	0.50
11	Respiratory Therapist	Assessment	$35.00	0.25
12	Respiratory Therapist	Case Management	$35.00	0.25
13	Respiratory Therapist	Direct Services	$35.00	0.50
14	Social Worker	Assessment	$35.00	1.50
15	Social Worker	Case Management	$35.00	0.25
16	Social Worker	Direct Services	$35.00	1.00

Expressing formal service intensity by the number of hours per day does not adequately reflect the underlying value of the resources consumed. That is to say that services provided by a nurse and services provided by the home care aide are not directly comparable. Expressing formal service intensity by the number of visits per day should not be relied upon as the number of visits is calculated (not directly observed). Even if the number of visits were observed, interpreting results in terms of visits is not intuitive. Hence the primary expression of formal service intensity is dollars per day.[8]

5. Competing Hazards

Competing hazards is an important consideration in defining the length of service in that end of an episode of home care is defined as the date of discharge to an institution (defined as a nursing home or skilled nursing facility). Discharges for any other reason are treated as a competing hazard (right censored data). A complete list of definitions for discharge reason, competing hazard categories and censorship status is provided below in 4.10. Note that almost 2/3 of discharges result in censored observations.

[8] Three other measures of service intensity exist: dollars/hour, hours/visit, and dollars/visit. They were not considered due to the fact that for 49 (1%) individuals, the denominator was zero. Hence these individuals would have been excluded from the analysis. Again, the second and third measures are based on calculated visits and are not intuitive. Although dollars/hour is intuitive in that it reflects the service mix of provider types, it was not used because it cannot be combined with LOS_i to obtain insight into total resource use.

Table 4.10: Discharge Reasons, Competing Hazards and Right Censorship Status

REASON FOR DISCHARGE	DEFINITION	NUMBER*	HAZARD (NUMBER)	STATUS
Admitted to a Facility	Admissions to either an acute care or long term care facility but no longer requires home care in the near future.	755	Institutional-ization (651=35.3%)	Event
Deceased	Discharged due to a death while in the homecare program.	446	Death (446=24.2%)	Censored
Moved	The client has moved out of the health region.	100	Other (746=40.5%)	Censored
Requires Services Not Provided	Services covered under the program but not available due to unique circumstances.	51		
Refused Services	Family/client refuses services that the health unit/region is willing to provide.	117		
Can Manage Care	The family/client can now manage care of the client.	338		
No Longer Requires Services	Due to improvement in functional status, services are no longer required.	229		
Insufficient Home Care Resources	The health unit/region has insufficient financial resources to provide the required service.	26		
Other	Reasons for discharge other than those provided.	36		
Total	All discharges.	1843	All Discharges	

*Does not add up to 1,843 as up to three reasons could be provided for a single discharge.

D. Procedures

1. Procedures for Data Collection

a) General Database Description

The Home Care Information System (HCIS) database contains all of the raw data used in this study. The organization of the data however does not allow for direct estimation of homecare utilization. The primary data were stored in five files. Table 4.11 provides the summary of file names, titles, the variables that uniquely describe the file record, and a brief description of the contents.

b) Description of Computer Software

All database manipulations and analysis was performed in SPSS version 7.0.

2. Database Manipulation

A new data model was developed to reorganize the raw data. The new files and units of observation are an episode, a classification, and a service. The variable definitions have been provided in section A of this chapter.

Table 4.11: HCIS Data File Description

File Name	Title	Index	Description
CC	Classification	An individual is classified on a specific date as Long Term, Short Term or Palliative.	Contains both the classification and assessment data. Assessments are recorded only for Long Term Care clients.
CM	Demographics	The unit of observation is the individual.	Contains demographic data entered at time of first contact with the home care program.
HS	Home Care Status	An individual may be admitted or discharged on a given date.	Contains the referral source and discharge destinations.
SM	Provider Services (Detailed)	The unit of observation is an service record. Some aggregation exists for contracted providers.	Contains information on whether a service provider is a nurse, licensed nurse practitioner, home care aide, social worker, or therapist. The services are subdivided into various categories. Contains both raw hours and costs.
SSUM	Provider Services (Archived)	Archived provider service data is aggregated to the level of the provider type, agency, and service type.	Same as above.

3. Sampling Methodology (Selection Criteria)

Two study populations consisting of 4,962 and 2,914 clients classified as long-term care clients were admitted to the home care program between April 1, 1994 and December 31, 1995. These are referred to as sample II and sample I respectively and differ with respect to increasingly stringent selection criteria. The cumulative selection criteria are summarized in Table 4.12.

Table 4.12: Sample Selection Criteria

Sample	Selection Criteria	Number Excluded	Percent Excluded
II	Admission On/After May 1, 1991 and Before Dec 31, 1995.	918	15.2%
	Valid episode. Each client must have a valid admission date and a valid discharge date if discharged.	35	0.6
	Valid assessment. Each client must have received at least one assessment during his or her last episode.	137	2.3
	Valid demographic data. Each client must have a valid date of birth, gender, and a valid diagnosis.	45	0.7
	All four criteria (13 individuals excluded for more than one reason).	1122	18.4%
I	Admission on or after April 1, 1994.	1414	23.2
	Assessment within one month of admission date.	634	10.4
Total (I)	All six criteria (13 individuals excluded for more than one reason).	3170	52%

As increasingly stringent criteria are applied to the initial population, the sample becomes more homogenous. This results in greater validity but less generalizability over the entire time period. To answer whether the results for the 4,962 clients are generalizable to the entire client base of 6,084 requires anova and chi-square tests for differences in the means and proportions of identifiable covariates is performed. The means and p-values are presented in the following table.

Table 4.13: Chi-Square Tests for Differences in Means

Covariate	**Mean or Proportion Admitted Before**	**Mean or Proportion Admitted After**	**p-value**
Age	73.89	73.32	.284
Gender (% Female)	0.78	0.63	.000
Informal Support 1	0.18	0.09	.000
Informal Support 2	0.35	0.26	.000
Informal Support 3	0.23	0.26	.048
Informal Support 4	0.14	0.23	.000
Informal Support 5	0.10	0.16	.000
Cost per Day	6.83	12.54	.000
Cost per Hour	16.54	19.82	.000
Cost per Visit	34.60	28.82	.000
Hours per Day	0.48	0.93	.000
Hours per Visit	2.37	1.86	.000
Mean Days / Visits	28.20	10.60	.000
# of Diagnoses		1.65	.000
Physical Function	0.23	-0.005	.000
Mental Function	0.36	-0.06	.000

The exact numbers of individuals in the two groups may differ for each individual test depending on the available covariates. On average more males have been admitted after May 1, 1991 with less informal support, a greater number of diagnoses and lower functional status. The treatment intensity is substantially greater due to the increased levels of funding in recent years.

Figure 4.1: Causal Relationships in Home Care Utilization

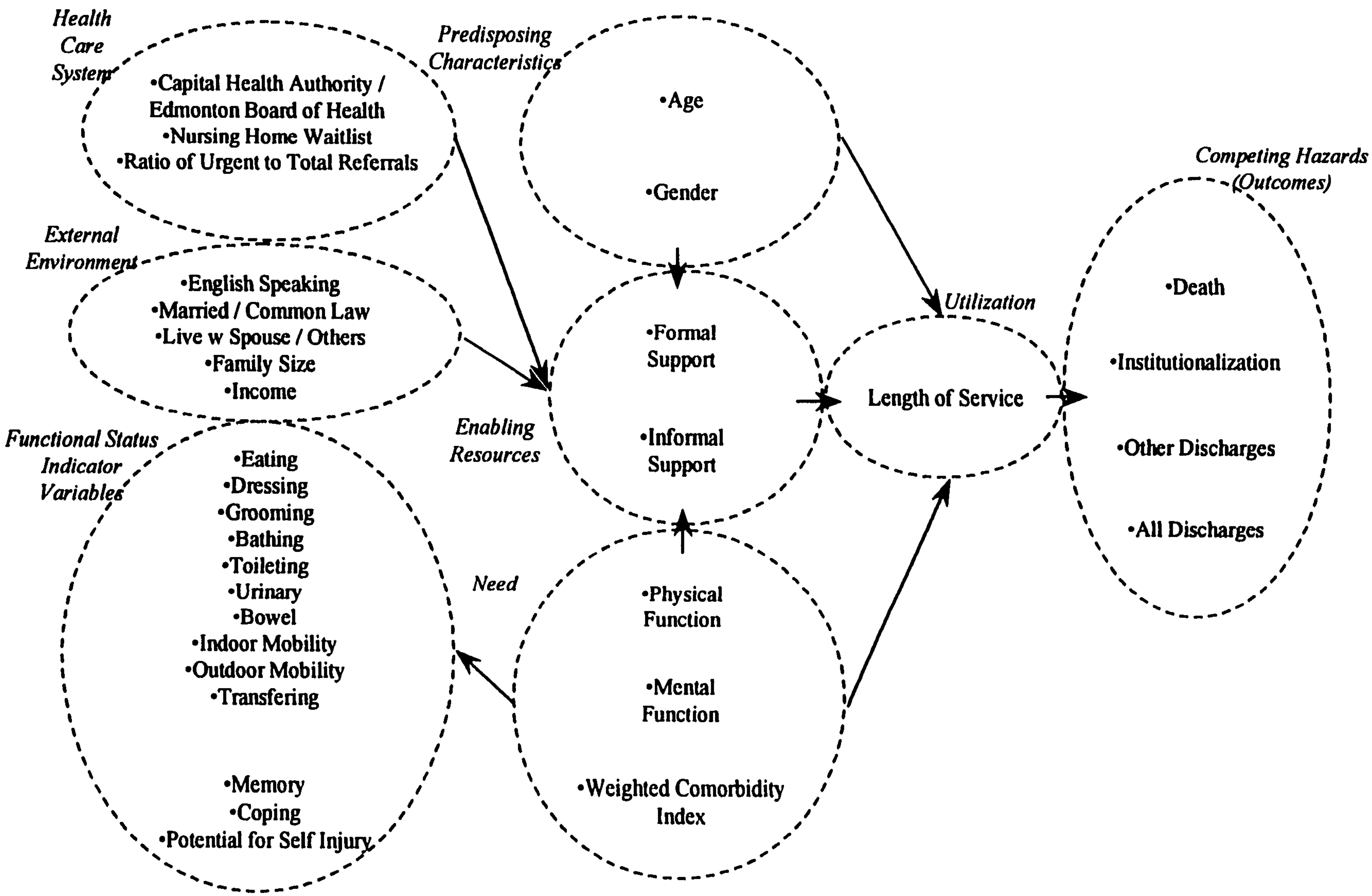


Figure 4.2 Distribution of Survival Times

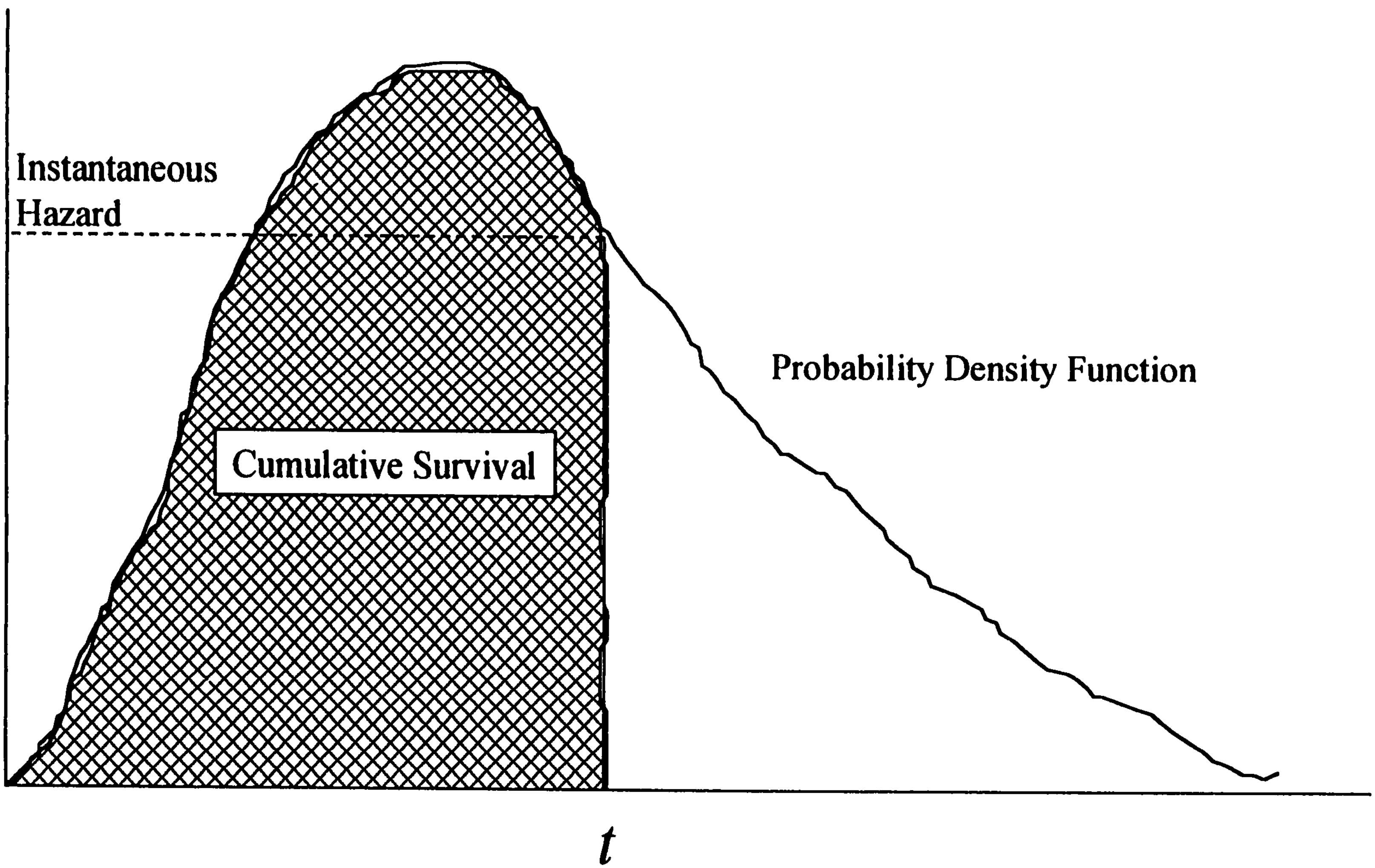


Figure 4.3: Cox Proportional Hazard Model Specifications

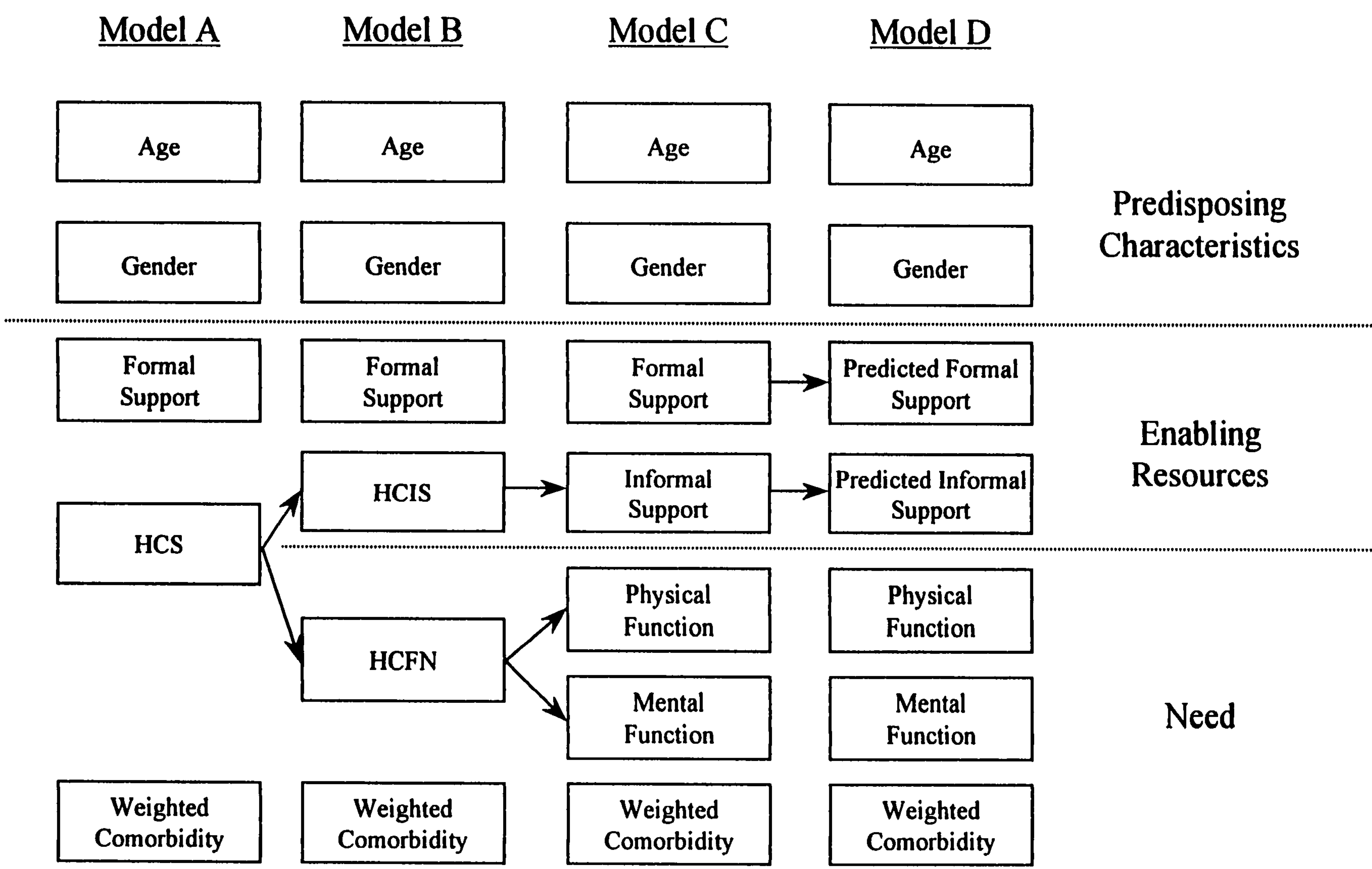


Figure 4.4: Structural Equation Diagram for Cox Model 'D'

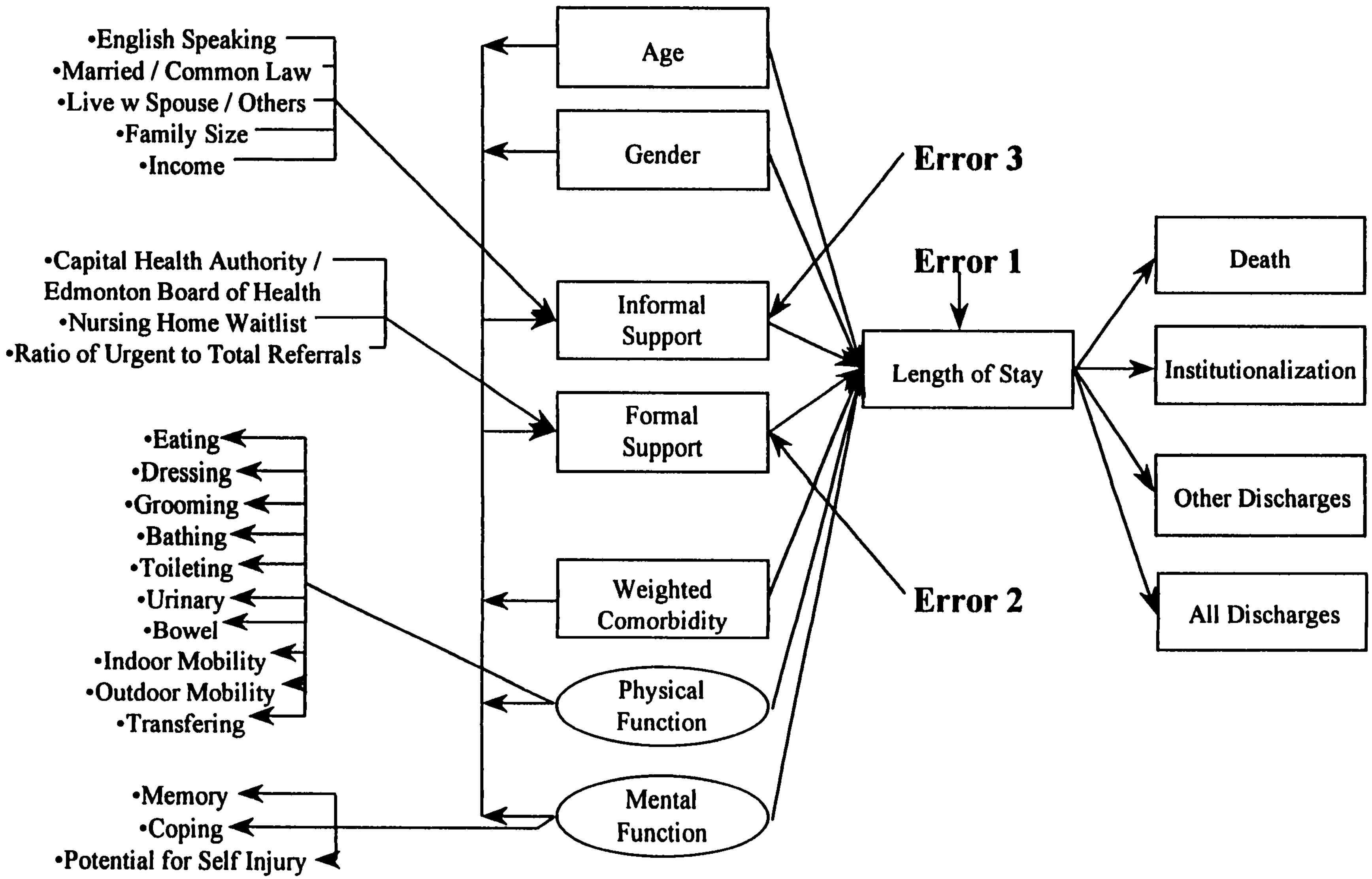
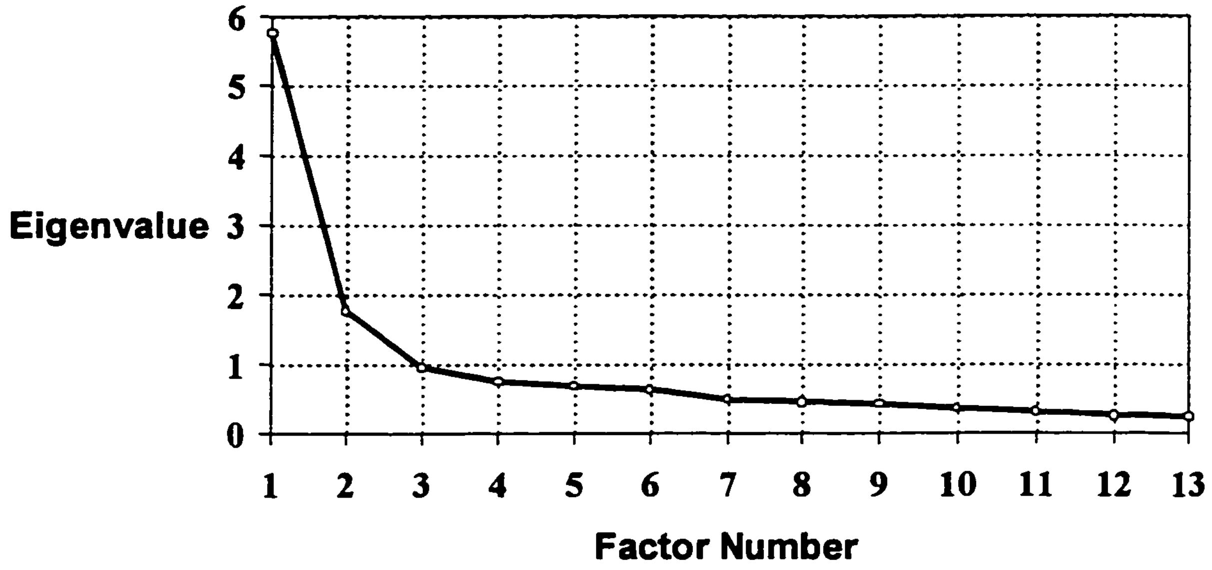


Figure 4.5: Factor Analysis

(a) Factor Scree Plot



(b) Factor Plot in Rotated (Varimax) Factor Space

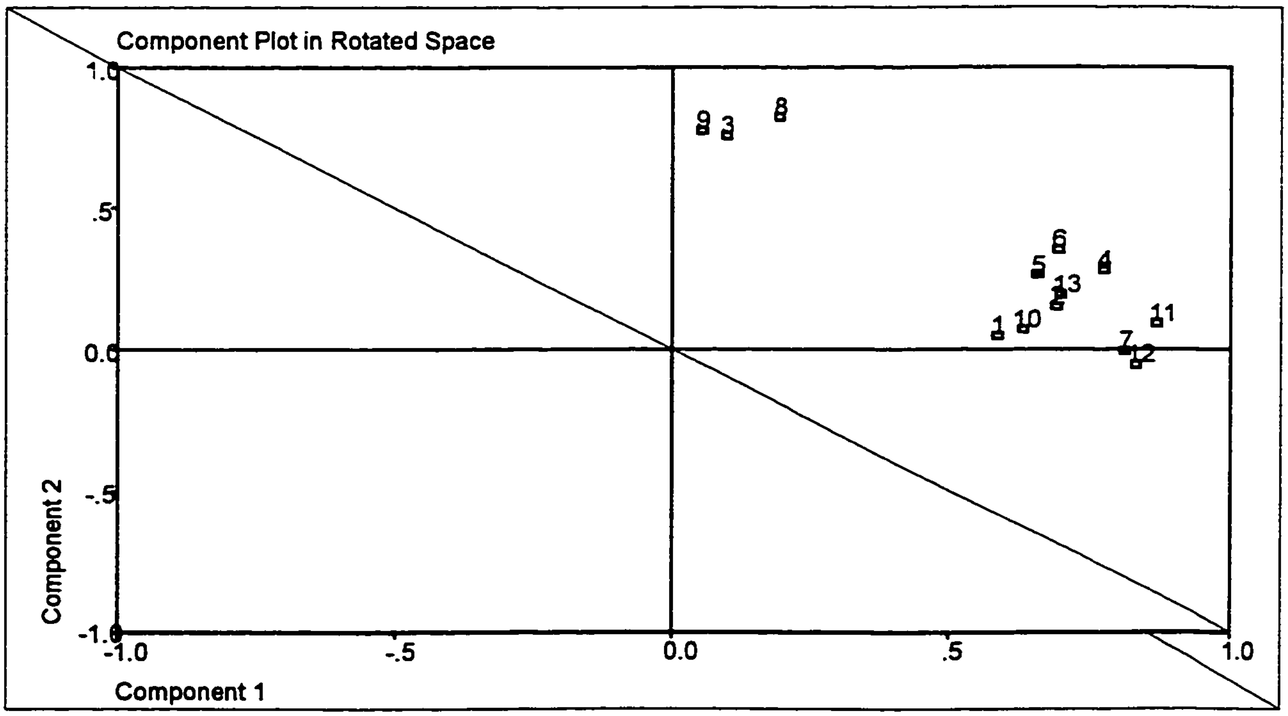


Figure 4.6 Average Daily Costs by Episode Month

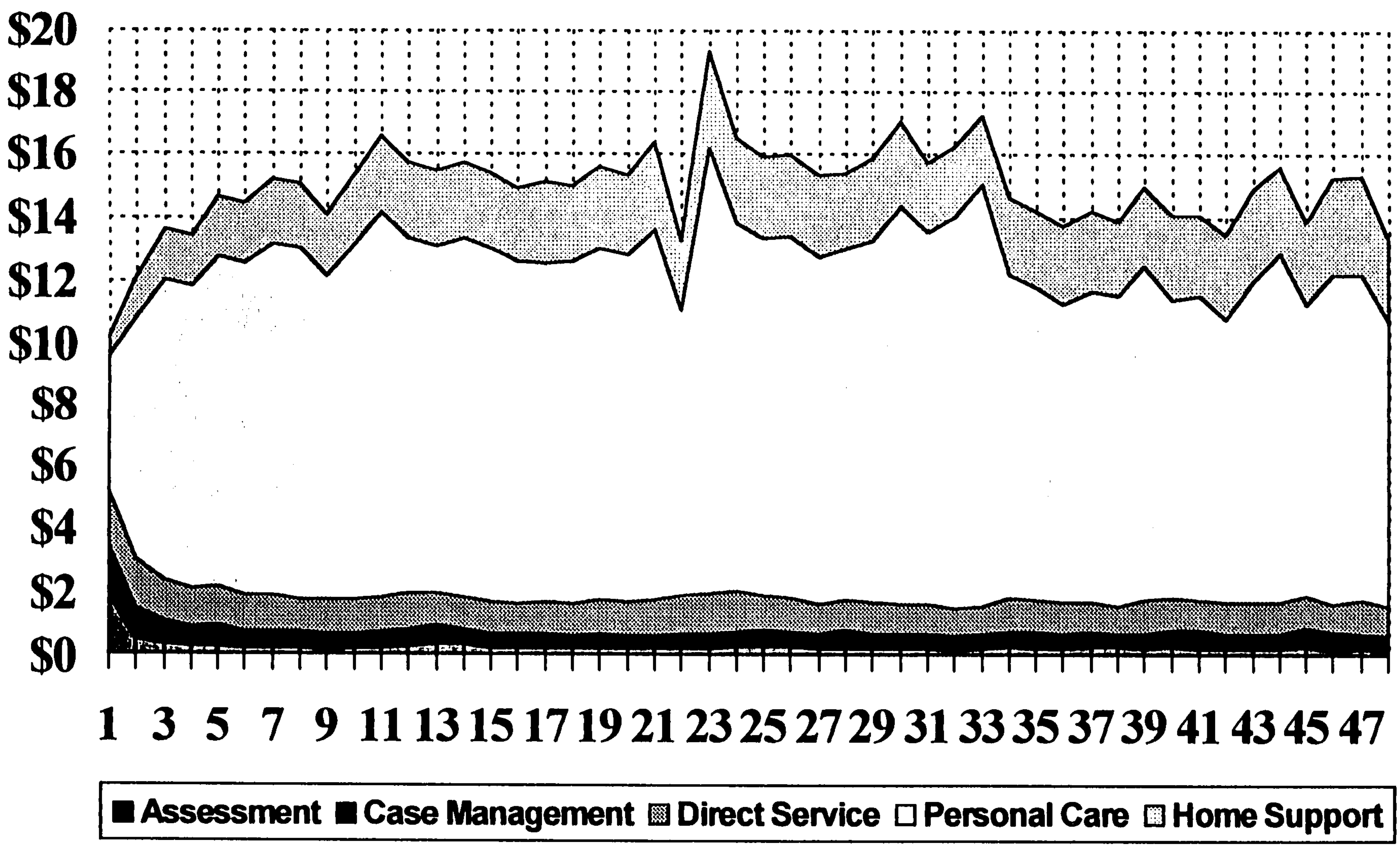


Figure 4.7 Average Daily Costs and Trend by Episode Month

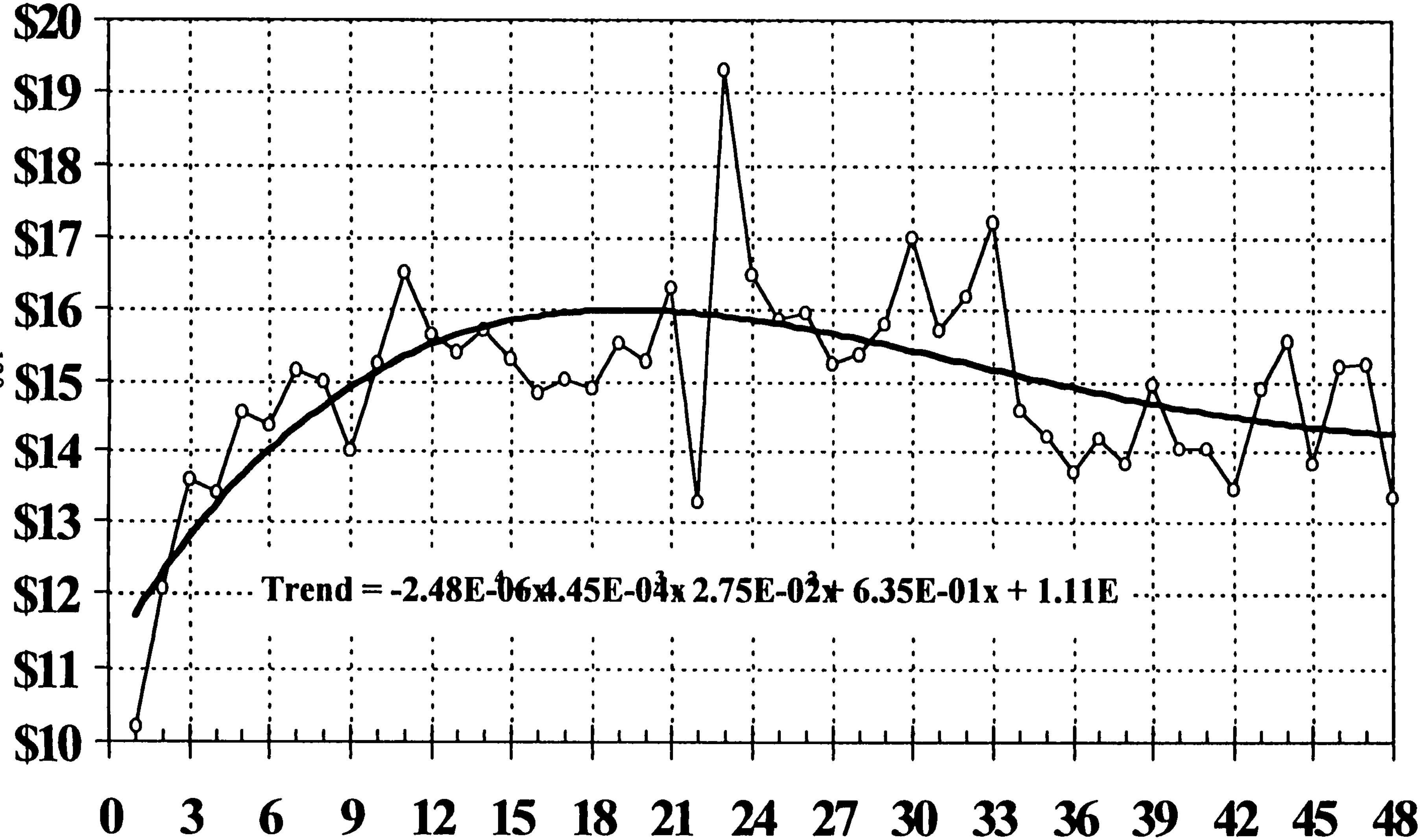
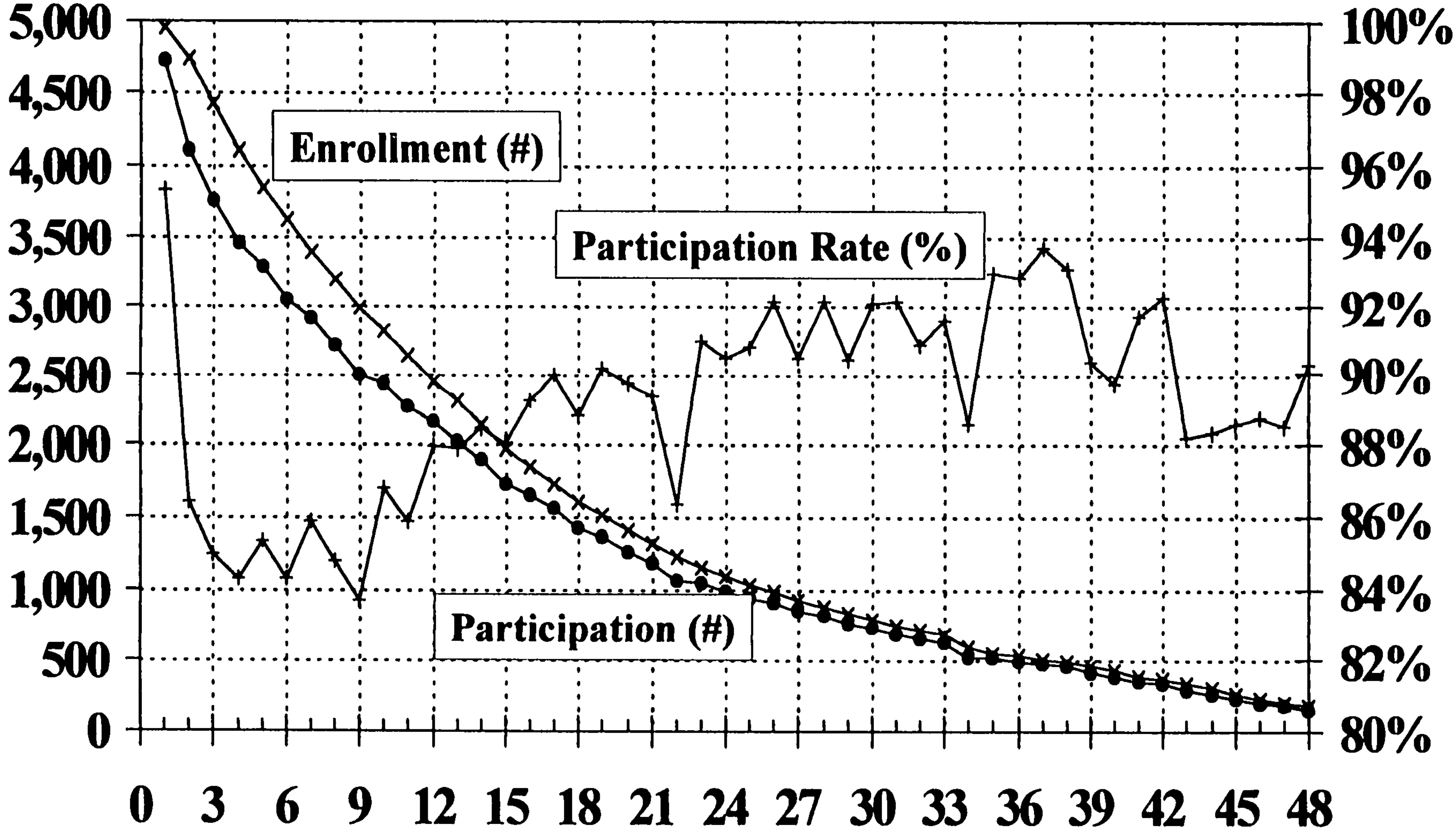


Figure 4.8 Enrollment, Participation and Participation Rate by Episode Month



Bibliography

Andersen, R. and Newman, J.F. (1973): Societal and individual determinants of medical care utilization in the United States. *Milbank Memorial Fund Quarterly*, 51:95-124.

Anderson, A., A. Auquier, W.W. Hauck, D. Oakes, W. Vandaele, and H.I. Weisberg (1980). Statistical Methods for Comparative Studies: Techniques for Bias Reduction. Wiley, New York, pp. 214-230.

Anderson, P.K. and M. Vaeth (1988). Survival Analysis. *Encylopedia of Statistical Sciences*, 9: 119-29.

Branch LG, Goldberg HB, Cheh VA, Williams J. Medicare home health: a description of total episodes of care. *Health Care Financing Review* 1993;14(4):59-74.

Bull K, Spiegelhalter DJ. Survival analysis in observational studies. *Statistics in Medicine* 1997;16(9):1041-74.

Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *Journal of Chronic Diseases* 1987;40(5):373-83.

Cox, D.R. (1972). "Regression Models and Life Tables," *JR Stat Soc (Series B)*, 34: 187-202.

Ellenberg JH. Selection bias in observational and experimental studies. *Statistics in Medicine* 1994;13(5-7):557-67.

HCIS Data Standards Manual, "Standardization of Data Elements for the Home Care Information System," Alberta Health, October 1994.

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

Jacobs P, Bachynsky J, Hall E. 1995. A manual of standard costs for pharmacoeconomic studies. Ottawa: Canadian Coordinating Office for Health Technology Assessment.

Lunn M, McNeil D. Applying Cox regression to competing risks. *Biometrics*. 1995;51(2):524-32.

Ng'andu NH. An empirical comparison of statistical tests for assessing the proportional hazards assumption of Cox's model. *Statistics in Medicine* 1997;16(6):611-26.

Pagan, A. Econometric issues in the analysis of regressions with generated regressors. *International Economic Review* 1984;25:221-247.

Parmar MKB and Machin D. (1995) Survival Analysis: A Practical Approach. John Wiley and Sons: New York.

Roeder, B.J. and Williams, D.N. (1985): Diagnosis-specific home care. The Park Nicollet model. *Postgraduate Medicine*, 77:79-81.

Romano PS, Roos LL, Jollis JG. Adapting a clinical comorbidity index for use with ICD-9-CM administrative data: differing perspectives. *Journal of Clinical Epidemiology* 1993;46(10):1075-9 discussion 1081-90.

Roos LL, Sharp SM, Wajda A. Assessing data quality: a computerized approach. *Social Science & Medicine* 1989;28(2):175-82.

Roos LL, Walld R, Wajda A, Bond R, Hartford K. Record linkage strategies, outpatient procedures, and administrative data. *Medical Care* 1996;34(6):570-82.

Roos NP, Roos LL, Mossey J, Havens B. Using administrative data to predict important health outcomes. Entry to hospital, nursing home, and death. *Medical Care* 1988;26(3):221-39.

Valsecchi MG, Silvestri D, Sasieni P. Evaluation of long-term survival: use of diagnostics and robust estimators with Cox's proportional hazards model. *Statistics in Medicine*. 1996;15(24):2763-80.

Valsecchi MG. Silvestri D. Sasieni P. Evaluation of long-term survival: use of diagnostics and robust estimators with Cox's proportional hazards model. *Statistics in Medicine* 1996;15(24):2763-80.

Zellner, A. Estimation of regression relationships containing unobservable independent variables. *International Economic Review* 1970; 11:441-454.

Zhao Y. Lee AH. Hui YV. Influence diagnostics for generalized linear measurement error models. *Biometrics* 1994;50(4):1117-28.

Zohoori N, Savitz DA. Econometric approaches to epidemiologic data: relating endogeneity and unobserved heterogeneity to confounding. *Annals of Epidemiology* 1997:7(4):251-7.

V. Description, Analysis and Interpretation of Data

In this chapter results of the statistical procedures outlined in this dissertation are presented. To begin with, descriptive statistics, charts and graphs for each of the dependent and independent variables are described. When appropriate, these statistics are supplemented by histograms illustrating frequencies and distribution of the variables.

All descriptive statistics are provided for sample II, while the Cox model results are provided for sample I and II. Sample I is the sample of 2,914 that were assessed within one month of admission. By default, this sample only includes individuals admitted to the program on or after April 1, 1994. Results from this sample may be deemed more valid due to the potential for greater homogeneity of unobservable patient characteristics. Sample II consists of the 4,962 clients that meet the basic data requirements. Results from these individuals are expected to be more generalizable to all homecare clients in the region.

In section B, the results for both the univariate and multivariate Kaplan-Meier are presented. Mean survival times and associated statistics will be provided for each stratum. The univariate analysis generates two separate survival curves for each independent variable. Each independent variable has been categorized into two categories for this purpose. Similarly, the stratified K-M analysis will consist of two sets of survival curves for each independent variable other than service intensity.

Results for the main and two service intensity submodels are then presented. Firstly, the results and accompanying econometric statistics are reported for two ordinary least squares

(OLS) models. Next the Cox model results are summarized for the four model specifications and the two samples. This is followed by a review of the model selection criteria. Based on the model selection criteria, one model is proposed as explaining variations in utilization better than the others. Test results for proportionality are then presented for this model. Section E then concludes this chapter with a review and interpretation of the data and results.

A. *Descriptive Statistics and Tabulation*

In this section I provide selected descriptive statistics for total resource utilization and for the dependent plus independent variables in the model. Maximum, minimum, mean and standard deviation are recorded for all of the variables. As well, pie charts, histograms, and time trend data for selected variables are also provided. Grouping of the variables, for convenience, is in accordance with Andersen's model: predisposing, enabling, need, external environment, and the health care system.

1. Total Resource Utilization

Explanation of variation in total resource utilization is the primary objective of the dissertation. Recall from the last chapter that total resource utilization was a function of both length of service and average intensity of resource utilization. Length of service and average intensity of resource utilization is discussed in the next two subsections. Although total resource utilization is not explicitly modelled, it is described here. For each month between May 1991 and Dec 1995, the discharge events are tabulated and the proportion discharged as of study completion. It is apparent that monthly admission have steadily increased from about 20/month in mid 1991 to about 200/month by early 1995. This increase in the admission rate reflects an overall increase in homecare program utilization. This increase reflects both an increase in demand (eg. population growth, decreased access to alternatives and increasing awareness of the service) and supply (eg. primarily increase in funding) factors.

Average home care costs per client per day totalled approximately $13. This number however cannot be used for planning purposes. Although variation in average costs by month is observed, the systematic variation in total utilization by episode month is more dramatic. As was seen in Figure 4.6, the first month's average daily cost approximated $10 while this cost escalated to $16 by the 18-month mark. Therefore, average total resource costs may be misleading and this section is only meant to provide a brief review of program expenditures.

Homecare services are described along two dimensions: provider and service. Services can be provided by a combination of nurses, therapists, home care aides, and/or social workers. Nurses are further differentiated into nurses and licensed practical nurses; while therapists are divided into occupational and respiratory therapy. Services include assessment, case management, direct services, personal care and home support. Personal care refers to help with functional limitations while home support refers to things like housecleaning. Figures 5.1 and 5.2 are pie charts that illustrate the average resource consumption by each category. For example, home care aides provide the bulk of services while personal care provides the majority of services received.

Figures 5.3 and 5.4 illustrate the monthly variations in average service intensity. It becomes apparent that there are both periodic and random fluctuations in these series. For example, LPN service intensity peaks three times in the summer months and twice around the end of the year. These may be explained by nursing strikes or by scheduling for holidays. Likewise, Figure 5.3 illustrate that nursing resource intensity is increasing at a steady pace. This contradicts the general trend in the hospital sector where LPN's appear to be replacing nurses. As shown in Figure 5.4, other than an apparent disruption in the intensity of personal care provision in mid 1993, the mix of services provided has remained constant. This suggests that although staffing patterns may change for institutional reasons, patterns of practice remain constant.

2. Dependent Variables

The dependent variables consist of the two components of total resource utilization: length of service and formal plus informal resource intensity. The dichotomization of service intensity reflects the fact total costs are comprised of direct (i.e. formal) and indirect (i.e. informal) costs. Direct costs refer to the expenditures for professional care of the client. Indirect costs for the home care population are mainly comprised of caregiver time and are calculated and not observed variables.

Descriptive statistics including the minimum, maximum, mean and standard deviation are included in Table 5.1 for the service intensity variables (dependent only):

Table 5.1 Dependent Variable Descriptive Statistics

Variable	Minimum	Maximum	Mean	Standard Deviation
Formal Service Intensity				
$/Day	0.00	315.02	13.06	19.86
Hour/Day	0.00	23.16	0.96	1.68
Visits/Day	0.00	13.53	0.41	0.57
Informal Service Intensity				
HCIS Score	1	5	3.12	1.22
$/Day	0.00	117.96	4.55	9.22

Table 5.1 illustrates the diversity of the homecare population. The minimum of zero means that some individuals have no recorded resource usage during their (usually short) tenures. In special circumstances, an individual may receive more than thirteen visits per day (calculated) and up to 23 hours per day (observed). The maximum costs observed for an individual was $315/day.

a) Length of Service

Length of service (LOS) is defined as the number of days from admission to the homecare program until the time of discharge to a nursing home (institutionalization). Data on hospitalization is not part of HCIS and therefore not considered here. Because length of service is censored there exists many problems in the graphical representation of survival times (Goldman 1992). Despite these problems a plot of an event history diagram is illustrated in Figure 5.5. Clients discharged for any reason have been distinguished from those individuals not discharged. The line created by those individuals not discharged is termed the "now-line." There do not appear to be any trends among those discharged.

Figure 5.6 further illustrates this last point. The steady decrease after mid 1994 is expected due to the increasingly shortened experience in the homecare program. Prior to this point there appear to be no systematic trends which would otherwise form a basis for concern.

b) Formal Service Intensity

Formal service intensity (FS) refers to the average observed or predicted level of homecare service provision over the time period in which services were actually provided (length of service). It may be specified in units of hours, dollars (1995 Canadian), or number of visits per day. A histogram illustrating the mean $13/day cost is shown in Figure 5.7. As expected, the distribution is highly skewed to the left with a mode of $5.

Figure 5.8 illustrates 6 average formal service intensity indicators over the 56 month period. As mentioned before $/hr is increasing due to the increased skill mix. Dollars per day has increased very slightly with an obvious anomaly in April 1993. Commensurate with the increase in $/hr is the decrease in $/visit. Conclusions based on visits are not made due to the fact that visits are a calculated variable.

The only concern to this point is the fact that increased skill mix may have an important role in client outcomes; particularly the timing of discharge. However, since the increase in nursing intensity has really only taken place in 1995 the impact on the overall results are mitigated. In other words, the model is based on 56 months of observations and only 12 of those exhibited higher than average skill mix. Even though patient admissions are higher during this period, skill mix is not factored directly into the model.

c) Informal Service Intensity

Informal service intensity (IS) is defined by two variables. This first is an observed, categorical variable specifying the category of informal service intensity. The second is calculated (later this chapter, section B.5) and expressed in the same units as formal service intensity. Informal services are defined as the level of indirect costs or resource consumption for which no payment is observed.

From Table 5.1 we note that informal resource intensity is on average one third that of formal resource intensity with much less variation. The decrease in variation may be explained due to the fact that informal resource intensity is modelled and not observed. In categorical terms, "most (>½) informal support is available" (Table 4.1, page 84).

3. Independent Variables

The independent variables are grouped according to the Andersen framework nomenclature. Descriptive statistics including the minimum, maximum, mean and standard deviation are included in Table 5.2 for all of the independent variables:

Table 5.2 Independent Variable Descriptive Statistics

Variable	Minimum	Maximum	Mean	Standard Deviation
Predisposing Characteristics				
Age	16.33	101.51	73.32	16.31
Gender	0	1	0.63	0.48
Enabling Resources				
Predicted FS	-2.89	134.99	13.06	13.07
Predicted IS	-7.01	85.02	4.55	5.35
Client Need Variables				
HCS	0	9	5.45	2.83
HCFN	1	5	3.05	1.38
PFn	-3.42	1.25	-.003	.9912
MFn	-3.47	2.57	-.06	1.0177
WCI	0	10	5.15	1.52
External Environment				
English	0	1	0.93	0.26
Marital Status	0	1	0.35	0.48
Living	0	1	0.54	0.50
Family Size	1	7	1.40	0.63
Income	7,860	81,645	15,988	13,275
Health Care System				
RUT	0.26	0.57	0.39	0.04
IWL	0.85	1.18	1.04	0.04
CHA	0	1	0.62	0.31

a) Predisposing Characteristics

(1) Age

Age is a continuous variable calculated as the difference in days between the date of birth and the date of first assessment divided by 365.25. The mean age for this client sample is 73.3 years. A histogram of client age is presented in Figure 5.9. The ages are strongly skewed to the right with a mode of 80 years of age. Figure 5.10 illustrates that the trend in average admission age has not varied but for the anomaly in April 1993.

(2) Gender

Gender is a dichotomous variable with females coded as '1' with males coded as '0'. Approximately 63% of the clientele are female, however this trend may be decreasing. In Figure 5.11 there appears to be a decrease in the proportion of females. Over the five year period the proportion has decreased about 10%. It is not known whether this phenomenon is attributable to the changing population demographics or a bias in the admission process.

b) Enabling Resources

There are two independent variables that summarize enabling resources in the model: predicted formal support ($\overline{FS}$) and predicted informal support ($\overline{IS}$). Observed formal support and observed informal support have been discussed earlier and are dependent variables in the two service intensity submodels.

(1) Predicted Formal Service Intensity

Predicted mean formal costs per day are also \$13 but with a smaller standard deviation (\$13.07 as compared to \$19.86). This is a consequence of the submodel. Since the number of unique groups identified by the independent variables of the formal service intensity submodel is less than the number of individuals, the variance must be lower. In addition, the predicted mean cost/day also remains the same by virtue of the submodel.

(2) Predicted Informal Service Intensity

Predicted informal support displays the same characteristics in comparison to observed informal support intensity as do the formal support counterparts. Mean independent costs are \$4.55/day with a lowering of the standard deviation from \$9.22 to \$5.35. The proportionately larger reduction in variance is indirectly because of the large number of missing informal service intensity submodel covariates. Since missing demographic covariates have been replaced by their mean, the number of unique categories to which individuals belong has been even further reduced. Lastly, the minimum value of -$7.01 is calculated and, as such, is no cause for alarm.

c) Need Variables

Need is defined by the client's functional status (HCS, HCFN, PFn, MFn) and disease burden or the weighted comorbidity index (WCI). The functional status measures represent different levels of aggregation of the need concept. Figure 5.12 illustrates the trend in need indices over the 56-month study period. Only the combined home care classification score and the weighted comorbidity index demonstrate an upward trend. A *bona fide* trend toward increasing disease burden cannot be explained by a possible increase in the recording of medical diagnoses. Similarly, the increase in the HCS index may be a result of coder bias. It should also be noted that once component of the HCS index is HCIS, an enabling resource. At any rate the trend does not appear steep and there is no visible movement in the physical or mental functioning levels.

(1) The Homecare Client Classification Scale

The combined homecare classification score (HCS) is a casemix variable which controls for diagnosis, severity of illness, and other predictors of resource utilization. It is an ordered ordinal variable that goes from '0', the level associated with the lowest predicted resource consumption, to the highest, '9'. It comprises both enabling resources and level of need in that it is the sum of both HCIS and HCFN. The average score is 5.45.

(2) The Homecare Function Need Index

The homecare functional need score (HCFN) is an ordered ordinal variable summarizing the levels of need as captured by thirteen indicators of functional status. The mean index score is 3.05 with a standard deviation of 1.38.

(3) Physical Function

Average physical function levels are not exactly zeroed due to the fact that the principal component analysis was performed on the larger sample of 6,084. By construction the index was defined so that the average would be close to zero and the standard deviation defined as one. Physical functioning levels vary from -3.42 standard deviations to 1.25.

(4) Mental Function

Average mental function levels are also not exactly zero due to the fact that the principal components analysis was performed on the larger sample of 6,084. A +1 indicates one standard deviation above average functional status and -1, one standard deviation below average functional status. Mental functioning varied from -3.47 to 2.57 standard deviations.

(5) Disease Burden or Comorbidity

Disease burden is measured by a (modified) weighted Charlson index of disease comorbidity (WCI). The index is a ratio scale that is observed to vary from 0 to 10. The mean index value was 5.15 with a standard deviation of 1.52.

d) External Environment

The external environment, more commonly referred to as socio-economic variables, is characterized by sociological variables over which the individual or the health care system has little control. The five external environmental variables are English speaking (ENG), marital status (MCL), living arrangement (LSO), families size (FSZ), and income (INC).

(1) English Speaking

Upon admission to the homecare program, clients' preferred primary and secondary languages are recorded. English speaking (ENG) is coded as a binary variable where 1 signifies the affirmative and 0 refers to all other languages. Approximately 93% of this sample are recorded as English speaking.

(2) Marital Status

Marital status has been aggregated to reflect that only 35% of this population is in a married or common law relationship. Because of the relatively high proportion of missing information, it is unknown whether this value reflects the true population values due to the possible presence of bias. For example, hospital coders may mistakenly assume that some individuals are single when in fact they are not.

(3) Living Arrangement

Likewise, the living arrangement (LSO) has been reclassified using judgement. Approximately 54% of clients live with a spouse and or others. Although the missing data may not correspond to the marital status, it could safely be assumed that large proportions of these individuals live with someone other than their partner.

(4) Family Size

Family size, FSZ, is treated as a continuous variable where the number reflects the combined number of adults and children in the family unit. Note that they need not reside at the same location. Family size varies from 1 to 7 individuals with a mean of 1.40. As with marital status, this variable reflects the lack of social support that prevents the clients from remaining independent.

(5) Income

Income (INC) has been recalculated as a continuous variable from discrete income categories using median values. Income varies from \$7,860 to \$81,645 per year for those individuals whose values are not missing. The mean level of income is \$15,988.

e) Health Care System

The health care system variables are those which indirectly influence length of service through formal service intensity. They act as constraints in the homecare managers' decision as to the appropriate level of service intensity provision. As mentioned before, the following health care system variables have been constructed which translate the effects of the health care system into individually observable influences.

(1) Ratio of Urgent to Total Referrals

The ratios of urgent to total referrals vary from 26% to 57% which is mirrors the trend in variation over the last 21 months (26% to 58%). The standard deviation is very small relative to the magnitude.

(2) Institutional or Nursing Home Waiting Lists

Average institutional or nursing home waitlist are 3.5% above the average monthly system average. This reflects the fact that those months at the end of an episode are weighted more heavily that those at the beginning. While the waiting list has varied from 82% to 119% over the last 21 months, the range of individual indices does the sae (85% to 118%).

(3) Management of the Homecare Program

The CHA variable is an arithmetic average of all episode months where CHA=1 and EBH=0. Therefore a value of 1 signifies an episode of home care that falls entirely within the Capital Health Authority's (CHA) management period. A value of 0 is assumed when the episode falls entirely under the Edmonton Board of Health's (EBH) management. As expected some episodes do encompass a time period characterised by one regime as opposed to the other. Therneau index is 0.62, which means that 62% of all client experience in this cohort falls under CHA's jurisdiction or management.

B. Kaplan-Meier Survival Analysis

In this section, the Kaplan-Meier probability of still receiving home care services at each point in time will be presented. Mean survival time, standard error and the 95% confidence intervals of cumulative survival are reported for each category of each independent variable. As mentioned earlier, the mean survival time is simply not the average of the observed survival times, but is calculated using special techniques in SPSS. Also, median times could not be calculated due to the fact that cumulative survival had not fallen below 50% during the 56-month (1703 days) period of observation.

The p-value associated with the log rank test statistic for equality of survival distributions is also tabulated. One p-value is provided for each independent variable in the univariate setting while two are presented in the stratified setting (one test statistic for each pair of survival curves at the two formal service intensity levels).

1. Kaplan-Meier Univariate Results

Univariate Kaplan-Meier analysis is used to elucidate the individual influences of explanatory variables on the cumulative survival function, in this case - length of home care service. Separate survival curves are calculated for individual values of discrete variables. Table 5.3 lists the variable name, category, standard error, 95% confidence intervals, and the log-rank p-value for each independent variable in the study.

The mean survival time for all covariates conform to expectations except those of formal service intensity (observed and predicted). For example, individuals over 75 years of age remain independent for an additional 207 days, on average. Similarly females remain in the home care program longer as do those with greater informal support. For all need variables, higher level of need corresponds to a lower mean survival time. Again with the exception of observed formal service intensity, the p-values are highly significant reflecting the fact that individual survival curves do differ from one another (between categories).

Table 5.3 Kaplan-Meier Univariate Results

Variable	Category	Mean Survival (Days)	Standard Error	Confidence Intervals Lower (5%)	Upper (95%)	Log-Rank Test p-value
Age	<75	1486	16	1455	1517	<0.0000
	>75	1279	18	1244	1314	-
Gender	Male	1304	23	1259	1350	<0.0000
	Female	1401	14	1373	1430	-
FS	<$7.50	1389	15	1359	1418	0.0445
	>$7.50	1340	21	1299	1381	-
Predicted FS	<$7.50	1460	18	1425	1494	<0.0000
	>$7.50	1309	17	1277	1342	-
HCIS	1	1573	26	1523	1623	<0.0000
	2	1518	19	1481	1555	-
	3	1371	25	1323	1419	-
	4	1231	29	1173	1288	-
	5	1179	34	1113	1246	-
IS	<$2.50	1331	17	1298	1364	<0.0000
	>$2.50	1424	18	1389	1460	
Predicted IS	<$2.50	1202	23	1156	1248	<0.0000
	>$2.50	1441	14	1414	1468	-
HCS	0	1618	28	1564	1673	<0.0000
	1	1489	33	1426	1553	-
	2	1512	32	1450	1574	-
	3	1540	27	1487	1593	-
	4	1446	34	1379	1512	-
	5	1378	37	1306	1449	-
	6	1345	51	1246	1445	-
	7	1208	39	1131	1286	-
	8	1204	42	1122	1286	-
	9	1210	29	1153	1267	-
HCFN	1	1508	24	1462	1555	<0.0000
	2	1502	20	1462	1542	-
	3	1371	29	1315	1428	-
	4	1199	29	1142	1255	-
	5	1223	29	1165	1281	-
PFn	<0	1248	22	1205	1291	<0.0000
	>0	1441	14	1413	1469	-
MFn	<0	1142	23	1097	1186	<0.0000
	>0	1502	13	1476	1527	-
WCI	<5	1438	16	1407	1468	<0.0000
	≥5	1294	19	1256	1332	-

Figures 5.13a-l illustrate cumulative survival curves for each category of each variable. The log-rank test statistic indicates on average that the curves do not cross. However, the graphs provide additional information. Although panels 5.13g-i demonstrate an overall spread of approximately parallel line, two adjacent lines may cross. For example, in Table 5.3 HCS category 1 and 2 are reversed and so forth.

The observation that formal service intensity does not conform to expectations is the basis of the self-selection problem. Persistence of the bias remains even after regressing on predisposing, enabling, need and system variables (i.e. formal service intensity submodel). In fact, comparison of panels 5.13c and 5.13d reveal that the bias has worsened. On the basis of this evidence alone, it would be incorrect to conclude that provision of these services is ineffective. The observation that higher intensity levels are associated with reduced mean survival times may signal the fact that these individuals' needs are greater.

There are basically three strategies to deal with this problem. Firstly, additional information (especially need variables) may correct the problem. This is done with the stratified K-M and Cox proportional hazards models. The second is to employ econometric techniques to remove the bias in the variable. This is accomplished by construction of fitted (predicted) service intensity and inclusion in the Cox model. Finally, the undertaking of a randomized trial (with predetermined service levels) is generally considered the most valid but least practical.

2. Stratified Kaplan-Meier Results

Stratified Kaplan-Meier analysis is used to elucidate the influences of formal resource intensity (cost per day), on survival stratified by each of the other independent variables. The primary strength of proceeding from univariate to stratified analysis is to determine whether the self-selection bias inherent in the provision of formal services can be corrected with any <u>one</u> covariate. The Cox model then will adjust for all covariates simultaneously.

Mean survival time, the standard error, the 95% confidence intervals will be reported for each individual stratum and the p-value associated with the log-rank test statistic are provided in the Table 5.4.

Table 5.4 Stratified Kaplan-Meier Results

Variable	Strata	Factor	Mean Survival (Days)	Standard Error	Confidence Intervals		Log-Rank Test p-value
					Lower (5%)	Upper (95%)	
Age	<75	<$7.50	1521	27	1469	1574	0.0056
		>$7.50	1467	18	1431	1502	-
	>75	<$7.50	1430	21	1389	1471	<0.0000
		>$7.50	1076	27	1023	1130	-
Gender	Male	<$7.50	1363	38	1289	1436	0.0057
		>$7.50	1271	28	1215	1327	-
	Female	<$7.50	1487	19	1449	1525	<0.0000
		>$7.50	1334	21	1293	1374	-
HCIS	1	<$7.50	1590	28	1535	1644	0.2477
		>$7.50	1497	57	1386	1608	-
	2	<$7.50	1545	22	1502	1588	0.0125
		>$7.50	1448	34	1380	1515	-
	3	<$7.50	1306	38	1232	1380	0.2251
		>$7.50	1401	31	1340	1462	-
	4	<$7.50	1138	65	1011	1266	0.1383
		>$7.50	1240	32	1178	1303	-
	5	<$7.50	904	123	662	1145	0.9010

		>$7.50	1179	34	1112	1246	-
IS	<$2.50	<$7.50	1415	20	1375	1455	<0.0000
		>$7.50	1220	26	1169	1270	-
	>$2.50	<$7.50	1535	31	1475	1595	0.0003
		>$7.50	1391	21	1349	1433	-
Predicted IS	<$2.50	<$7.50	1359	31	1298	1419	<0.0000
		>$7.50	1081	32	1018	1143	-
	>$2.50	<$7.50	1494	20	1454	1534	0.0003
		>$7.50	1408	18	1372	1443	-
HCS	0	<$7.50	1639	25	1591	1688	0.2076
		>$7.50	1044	47	952	1135	-
	1	<$7.50	1489	36	1418	1560	0.8158
		>$7.50	1472	64	1346	1598	-
	2	<$7.50	1491	34	1424	1559	0.7840
		>$7.50	1542	58	1429	1656	-
	3	<$7.50	1552	31	1492	1613	0.2566
		>$7.50	1445	50	1346	1543	-
	4	<$7.50	1440	43	1355	1525	0.8963
		>$7.50	1452	55	1344	1560	-
	5	<$7.50	1351	53	1246	1456	0.5024
		>$7.50	1375	46	1286	1465	-
	6	<$7.50	1022	70	885	1160	0.0662
		>$7.50	1417	56	1308	1527	-
	7	<$7.50	1064	82	904	1224	0.0329
		>$7.50	1243	44	1156	1329	-
	8	<$7.50	1157	96	968	1345	0.2397
		>$7.50	1205	43	1120	1290	-
	9	<$7.50	602	203	203	1000	0.1447
		>$7.50	1213	29	1156	1271	-
HCFN	1	<$7.50	1516	26	1466	1567	0.3944
		>$7.50	1479	55	1371	1586	-
	2	<$7.50	1499	25	1450	1548	0.9192
		>$7.50	1500	36	1428	1571	-
	3	<$7.50	1365	43	1280	1450	0.9967
		>$7.50	1373	38	1298	1449	-
	4	<$7.50	944	79	789	1099	0.0025
		>$7.50	1222	30	1163	1281	-
	5	<$7.50	-	-	-	-	>0.9999
		>$7.50	1223	29	1165	1281	-
PFn	<0	<$7.50	976	103	773	1178	0.0445
		>$7.50	1254	22	1210	1298	-
	>0	<$7.50	1473	18	1438	1507	0.0011
		>$7.50	1389	24	1342	1437	-
MFn	<0	<$7.50	1213	41	1133	1293	0.0040
		>$7.50	1108	27	1055	1162	-
	>0	<$7.50	1545	18	1510	1579	0.0002
		>$7.50	1463	19	1426	1501	-
WCI	<5	<$7.50	1529	20	1490	1568	<0.0000
		>$7.50	1361	23	1317	1406	-
	≥5	<$7.50	1353	32	1290	1415	0.0002
		>$7.50	1257	24	1210	1304	-

The primary observation to make is that the evidence suggests that any one covariate does not correct for the self-selection bias. In fact, as observed in Figures 5.14a-j, the cumulative survival curves remain 'reversed' for the most part. The only exceptions are HCS(6,7), HCIS(3,4), HCFN(4), and PFn<0. Of these, only HCS(7), HCFN(4) and PFn<0 are significant. Optimism is not warranted at this point as within these categories, there are only two dozen or so observations less than \$7.50/day (predicted formal support). The p-values for the remaining 3 reversals are insignificant.

Observation of the p-values for the log-rank statistics supply evidence for inclusion of age, gender, informal support (observed and predicted), physical plus mental function, and the weighted comorbidity index. Evidence would suggest poor performance of the combined home care classification index and its two components: functional need and informal support scores. One explanation for this observation is that multiple categories reduce the number of events for each strata-factor combination thereby increasing the potential for wide confidence intervals. For example, as seen in Table 5.4, the combination of HCFN(5) and predicted FS<\$7.50 contains no observations and therefore statistics could not be calculated.

C. *Service Intensity Submodels*

Before the Cox proportional hazards model results are reviewed, the service intensity submodel results are tabulated. Since these models are ordinary least squares, the standard econometric statistics are tabulated: R^2, standard error of the estimate, and the F-test of model significance (that all covariates = 0). Covariate statistics will consist solely of the coefficients and the p-values associated with the Student's t-test of individual significance.

For each submodel, three variants of service intensity are tested. In each case the numerator alternates between cost, hour and visit while the denominator equals day. Although formal and informal costs per day are the dependent variables of interest, the two other service intensity model results are provided for insight.

1. Formal Service Intensity

For each of the three variants of formal service intensity the following statistics are provided in Tables 5.5a-b: R^2, standard error of the estimate, the F-test of model significance, individual coefficients, and the Student's t-test of individual significance.

Table 5.5.a Formal Service Intensity Model Results

Dependent Variable	R^2	Standard Error	F-test Statistic	F-test p-value
Dollars/Day	0.433	14.96	420.7	<0.000
Hours/Day	0.475	1.22	497.6	<0.000
Visits/Day	0.309	0.48	246.3	<0.000

To begin with, approximately half of the covariates are insignificant for each of the three submodels. As evidenced both by the number of insignificant covariates and the R^2 statistics, the visit per day model does not perform as well. When the dependent variable is hours per day, the model fit is slightly better. It should be noted that R^2 should not be solely relied upon for model fit. Finally, the evidence suggests that none of the models can be rejected at this time.

Table 5.5.b Formal Service Intensity Model Results

Dependent Variable	**Covariate**	**Beta Coefficient**	**Students t-test**	
			Statistic	**p-value**
Dollars/Day	Constant	41.3	3.3	0.001
	Age	-0.2	-14.5	<0.000
	Gender	-0.3	-0.8	0.443
	Informal $/Day	1.1	46.5	<0.000
	PFn	-6.5	-29.6	<0.000
	MFn	-1.3	-6.0	<0.000
	WCI	0.07	0.5	0.639
	RUT	-10.0	-0.7	0.455
	IWL	-15.1	-1.6	0.104
	CHA	0.8	0.5	0.631
Hours/Day	Constant	3.5	3.4	0.001
	Age	-0.02	-17.6	<0.000
	Gender	-0.01	-0.3	0.754
	Informal hr/Day	1.1	49.3	<0.000
	PFn	-0.5	-30.9	<0.000
	MFn	-0.09	-5.2	<0.000
	WCI	-0.008	-0.6	0.519
	RUT	-1.1	-1.0	0.315
	IWL	-0.9	-1.2	0.225
	CHA	-0.1	-0.7	0.455
Visits/Day	Constant	0.8	2.1	0.039
	Age	-0.005	-12.3	<0.000
	Gender	-0.02	-1.3	0.183
	Informal vst/Day	1.1	32.7	<0.000
	PFn	-0.2	-25.0	<0.000
	MFn	-0.01	-1.4	0.155
	WCI	0.008	1.8	0.070
	RUT	-0.1	-0.3	0.773
	IWL	-0.1	-0.4	0.653
	CHA	0.04	0.9	0.378

Of the significant variables, only age and informal service intensity consistently violate hypothesized coefficient signs. One should expect that as age increases that, *ceteris paribus*, formal service intensity should increase. This is not observed and no simple explanation can be provided. (Unless one is willing to say that as a client ages, resource intensity is decreased but intended services are provided over a longer period of time - a stretch.) Likewise, a substitution effect should be observed for the informal service intensity covariate. As increasing levels of informal services are available, then one should expect a decrease in the provision of formal services.

The next observation is that physical and mental functioning levels are consistent determinants of formal service provision. Only the mental functioning level in the visits per day model is insignificant (p=0.155), but the correct sign. In terms of the magnitude of the coefficients as well, it should be noted that these two coefficients are strong drivers of formal resource intensity.

The last observation is that the system variables are consistently insignificant. That is to say that the evidence suggests that formal service intensity is independent of institutional arrangements arising from program management and nursing home waiting lists. Although this result was not expected, a working hypothesis suggested earlier is that the source of selection bias is based on unobservable need variables. This would suggest mis-specification of the model through missing explanatory variables. Interpretation of the coefficients therefore must be made with caution.

2. Informal Service Intensity

For each of the three variants of informal service intensity the following statistics are provided in Tables 5.6.a-b: R^2, standard error of the estimate, the F-test of model significance, individual coefficients, and the Student's t-test of individual significance.

Table 5.6.a Informal Service Intensity Model Results

Dependent Variable	R^2	Standard Error	F-test Statistic	p-value
Dollars/Day	0.580	7.5	228.4	<0.000
Hours/Day	0.368	0.6	261.7	<0.000
Visits/Day	0.222	0.2	128.2	<0.000

To begin with, more than half of the covariates is insignificant for each of the three submodels as evidenced by the number of insignificant covariates. The R^2 statistics provide evidence that when service intensity is measured with dollars per day, that a significantly greater proportion of the model variance is explained. Again it should be noted that R^2 should not be solely relied upon for model fit. As before, the F-statistics suggests that none of the models can be rejected at this time.

Of the significant variables, only formal service intensity consistently violates the hypothesized coefficient signs. Although one should expect, *ceteris paribus*, a substitution effect between formal and informal services, the lack of evidence here is not as of great concern as it was in the formal service intensity submodel. That is to say that as formal service intensity increases, the provision of informal services need not be observed to decrease. The only problem with this argument is that only a single (usually) observation is made regarding the capacity for informal service provision. Longitudinal informal service provision is not recorded and this is a serious limitation of the homecare database.

Table 5.6.b Informal Service Intensity Model Results

Dependent Variable	Covariate	Beta Coefficient	Students t-test Statistic	p-value
Dollars/Day	Constant	1.6	1.8	0.069
	Age	-0.02	-2.1	0.034
	Gender	0.4	1.9	0.062
	Formal $/Day	0.3	46.9	<0.000
	PFn	2.2	18.3	<0.000
	MFn	0.03	0.7	0.500
	WCI	-0.04	-0.5	0.622
	English	-0.6	-1.4	0.149
	Marital Status	-0.06	-0.2	0.847
	Living	1.3	4.8	<0.000
	Family Size	0.2	0.6	0.523
	Income	-0.000002	-0.2	0.833
Hours/Day	Constant	0.06	0.8	0.453
	Age	-0.0008	-1.2	0.222
	Gender	0.04	1.9	0.056
	Formal hr/Day	0.3	49.7	<0.000
	PFn	0.2	18.8	<0.000
	MFn	0.006	0.6	0.542
	WCI	0.002	0.3	0.800
	English	-0.05	-1.3	0.198
	Marital Status	0.006	0.2	0.812
	Living	0.1	4.4	<0.000
	Family Size	0.02	1.1	0.283
	Income	-0.0000002	-0.2	0.823
Visits/Day	Constant	0.1	5.9	<0.000
	Age	-0.001	-7.2	<0.000
	Gender	0.01	1.7	0.081
	Formal vst/Day	0.2	32.9	<0.000
	PFn	0.03	12.0	<0.000
	MFn	-0.005	-1.9	0.060
	WCI	-0.001	-0.7	0.465
	English	-0.02	-1.7	0.087
	Marital Status	-0.01	-1.7	0.096
	Living	0.04	5.4	<0.000
	Family Size	-0.002	-0.4	0.679
	Income	-0.0000002	-0.6	0.544

The last observation is that there is evidence that only three variables are important correlates of informal resource intensity. They are age at admission, the level of physical function, and whether the client lives with a spouse and/or others. The fact that informal support decreases, as age increases is unavoidable. This simply reflects the lack of the informal support network as one gets older. The evidence is not uniform in that age is not significant in the hours/day model. Similarly, the increase of informal support if a homecare client lives with a spouse and/or others reflects the fact that those who are close at hand naturally provide informal support. That is to say that withholding support is difficult when you live with someone. Out of five environmental variables, only this one was significant. Lastly, the fact that informal support increases strongly with an increase in physical functioning is not readily explainable.

D. *Cox Proportional Hazards Model*

In this section, the results and tests underlying the assumptions behind the Cox proportional hazards model are provided (Cox and Oakes 1984; Kalbfleisch and Prentice 1980). In the first section, relax assumptions regarding the covariates, selection criteria, and definition of an event (due to competing hazards). The results for 32 separate regressions are presented and discussed.

After the relative merits of each model are argued, non-traditional model selection criteria are proposed in section D.2. On the basis of test statistics derived from the log-likelihood function, the test statistics across the 32 models were aggregated. To make some specifications about the model, sample and hazard definition which is most consistent with theory and *a prior* assumptions and best describe variations in individual utilization of home care resources. Finally, in section D.3, assumptions regarding the underlying proportionality of the Cox model are tested graphically for this model.

1. Model Results

Cox model results are provided for four models (Figure 4.3), two samples (I and II), and four different event definitions (death, institutionalization, other, all discharges). The four models are not nested, as they do not reflect merely the consideration of additional variables. As one progresses from Model A to Model C, additional modelling assumptions have been made. Model D incorporates the two auxiliary models and attempts to fully compensate for the self-selection bias.

Up to this point, only sample statistics for the less restrictive model have been provided. The evidence suggests that mis-specification errors and the presence of bias require the full Cox model. However, the exact nature of the unobservable influences remains unknown but is hypothesized to be related to client need. As discussed in Chapter IV, the timing of assessment may be related to client need. Therefore, it is imperative to include the more restrictive sample in the analysis.

The notion of competing hazards likewise plays an important role in the analysis. Of the 4,962 individuals, 62.9 have not been discharged, 9.0% are deceased, 13.1% are institutionalized, 9.8% no longer require services, services are not provided to 2.2%, and reasons for discharges are not documented for 3.1% of the clients. The latter three categories have been aggregated to other discharges. Up to this point it would have been too much to analyze the statistics for both samples and the four hazard definitions. It is also desirable to be able to generalize the statistical findings to the typical homecare client. However, since the Cox model incorporates all assumptions made to this point in time, validity of results becomes an issue. It is imperative that as few unknowns as possible confound negative results.

Table 5.7 summarizes the findings of the 32 Cox models. The coefficient is reported as a hazard rate where a positive coefficient corresponds to a hazard greater than 1. Likewise, a negative coefficient corresponds to a hazard less than 1. The p-values associated with significance of the covariates are also provided.

Table 5.7 Informal Service Intensity Model Results

Model A *FS=$/Day Sample I*

	Death		Institutionalization		Other Discharges		All Discharges	
	Exp(B)	Sig	Exp(B)	Sig	Exp(B)	Sig	Exp(B)	Sig
Age	1.035	0.000	1.031	0.000	0.994	0.030	1.014	0.000
Sex	0.671	0.001	0.800	0.031	0.906	0.271	0.808	0.000
FS	1.002	0.431	1.003	0.148	0.972	0.000	0.995	0.003
HCS	1.100	0.000	1.228	0.000	0.954	0.005	1.055	0.000
WCI	1.294	0.000	1.135	0.000	1.007	0.822	1.110	0.000

Model B *FS=$/Day Sample I*

	Death		Institutionalization		Other Discharges		All Discharges	
	Exp(B)	Sig	Exp(B)	Sig	Exp(B)	Sig	Exp(B)	Sig
Age	1.035	0.000	1.031	0.000	0.994	0.029	1.013	0.000
Sex	0.678	0.002	0.782	0.018	0.907	0.277	0.808	0.000
FS	1.002	0.428	1.002	0.465	0.972	0.000	0.994	0.001
HCIS	0.992	0.919	1.418	0.000	0.990	0.869	1.096	0.019
HCFN	1.238	0.003	1.165	0.013	0.919	0.087	1.060	0.078
WCI	1.289	0.000	1.135	0.000	1.008	0.792	1.109	0.000

Model C FS=$/Day IS=$/Day Sample I

	Death		Institutionalization		Other Discharges		All Discharges	
	Exp(B)	Sig	Exp(B)	Sig	Exp(B)	Sig	Exp(B)	Sig
Age	1.034	0.000	1.032	0.000	0.994	0.023	1.014	0.000
Sex	0.673	0.001	0.793	0.024	0.916	0.331	0.819	0.001
FS	1.002	0.495	1.003	0.304	0.982	0.000	0.997	0.086
IS	0.998	0.758	1.000	0.964	0.959	0.000	0.986	0.002
PFn	0.694	0.000	0.617	0.000	1.126	0.038	0.800	0.000
MFn	1.246	0.000	0.708	0.000	1.001	0.987	0.937	0.016
WCI	1.302	0.000	1.087	0.025	1.007	0.822	1.100	0.000

Model D FS=Predicted $/Day IS=Predicted $/Day Sample I

	Death		Institutionalization		Other Discharges		All Discharges	
	Exp(B)	Sig	Exp(B)	Sig	Exp(B)	Sig	Exp(B)	Sig
Age	1.034	0.000	1.032	0.000	0.985	0.000	1.011	0.000
Sex	0.672	0.001	0.795	0.027	0.911	0.306	0.819	0.001
FS	0.999	0.896	1.002	0.660	0.960	0.000	0.989	0.007
IS	1.007	0.537	1.003	0.772	0.945	0.000	0.985	0.036
PFn	0.679	0.000	0.614	0.000	0.978	0.790	0.762	0.000
MFn	1.245	0.000	0.710	0.000	0.952	0.297	0.923	0.004
WCI	1.303	0.000	1.087	0.025	1.009	0.778	1.100	0.000

Model A FS=$/Day Sample II

	Death		Institutionalization		Other Discharges		All Discharges	
	Exp(B)	Sig	Exp(B)	Sig	Exp(B)	Sig	Exp(B)	Sig
Age	1.034	0.000	1.038	0.000	0.996	0.086	1.018	0.000
Sex	0.602	0.000	0.748	0.000	0.811	0.007	0.732	0.000
FS	1.003	0.283	1.000	0.938	0.974	0.000	0.994	0.001
HCS	1.091	0.000	1.254	0.000	0.929	0.000	1.058	0.000
WCI	1.272	0.000	1.177	0.000	1.068	0.006	1.155	0.000

Model B FS=$/Day Sample II

	Death		Institutionalization		Other Discharges		All Discharges	
	Exp(B)	Sig	Exp(B)	Sig	Exp(B)	Sig	Exp(B)	Sig
Age	1.034	0.000	1.038	0.000	0.996	0.069	1.018	0.000
Sex	0.607	0.000	0.736	0.000	0.808	0.006	0.731	0.000
FS	1.002	0.366	0.998	0.391	0.973	0.000	0.993	0.000
HCIS	1.026	0.699	1.525	0.000	1.011	0.840	1.159	0.000
HCFN	1.190	0.003	1.124	0.018	0.851	0.000	1.011	0.688
WCI	1.269	0.000	1.176	0.000	1.070	0.005	1.155	0.000

Model C FS=$/Day IS=$/Day Sample II

	Death		Institutionalization		Other Discharges		All Discharges	
	Exp(B)	Sig	Exp(B)	Sig	Exp(B)	Sig	Exp(B)	Sig
Age	1.036	0.000	1.034	0.000	0.994	0.018	1.016	0.000
Sex	0.611	0.000	0.753	0.001	0.826	0.014	0.749	0.000
FS	1.002	0.582	0.998	0.463	0.984	0.000	0.995	0.012
IS	0.997	0.648	1.005	0.370	0.954	0.000	0.989	0.006
PFn	0.730	0.000	0.678	0.000	1.277	0.000	0.849	0.000
MFn	1.102	0.039	0.630	0.000	0.894	0.005	0.816	0.000
WCI	1.274	0.000	1.122	0.000	1.059	0.021	1.132	0.000

Model D *FS=Predicted $/Day* *IS=Predicted $/Day* *Sample II*

	Death		Institutionalization		Other Discharges		All Discharges	
	Exp(B)	Sig	Exp(B)	Sig	Exp(B)	Sig	Exp(B)	Sig
Age	1.036	0.000	1.035	0.000	0.985	0.000	1.014	0.000
Sex	0.610	0.000	0.758	0.001	0.820	0.011	0.749	0.000
FS	0.998	0.786	1.008	0.120	0.956	0.000	0.992	0.022
IS	1.006	0.580	0.985	0.128	0.954	0.001	0.981	0.005
PFn	0.714	0.000	0.724	0.000	1.060	0.437	0.832	0.000
MFn	1.100	0.046	0.635	0.000	0.845	0.000	0.808	0.000
WCI	1.274	0.000	1.122	0.000	1.061	0.017	1.132	0.000

The major finding is that there is no evidence in favour of formal and informal service intensity as an alternative to death and institutionalization. This is evidenced by the consistent insignificance of these two variables. Rather than view this as negative evidence, it should be noted that the formal resource intensity covariate has been significant but in the wrong direction. It would therefore appear that when taking all confounding variables into account simultaneously, and correcting for self-selection bias that the direction of change in the coefficient is positive.

The second major finding is that the evidence for explaining resource utilization when discharges are performed for other reasons is less than that of the other models. This is somewhat expected as other discharges encompassing both client and system related reasons. Interestingly, predisposing characteristics are significant and possess the correct sign in this context suggesting that either resource provision is effective or that these individuals are not as severe as presented by the observable covariates.

The excellent fit for the relatively large number of clients in the other discharges category drives all discharges. Results must be interpreted with caution in that the competing hazards are like apples and oranges. Not only are the competing hazards incomparable, it is not clear whether they are independent. For example, if an individual is discharged dead, it is likely that the same mechanism is responsible for institutional discharges. Estimation of all discharges assumes independency and as such the hazards are additive. On the other hand, program planning requires an estimate of overall resource consumption. Furthermore, clients cannot be identified as belonging to one group or another at the time of admission.

2. Model Selection

Although the models are not nested, test statistics based on the likelihood function are used determine which of the four models explains variations in client utilization of home services the best. Four each of the 32 regressions, the number of observations, events, censorship rate, log-likelihood function, likelihood ratio test statistic and the score (also known as the global chi-square) statistic are reported in Table 5.8.a.

Table 5.8.b summarizes the two test statistics by providing the minimum, maximum, median and mean statistics so that comparisons can be made between the models.

Comparisons of the likelihood ratio test statistic and the score statistics are not standard statistical practice. Since the models are not nested, they cannot be directly compared. On the other hand they do provide information in that they provide a measure of how well the data are explained by the various models. The ranges and median are provided in addition to the mean as the distribution of non-nested model test statistics is not known.

Table 5.8.a Model Selection Statistics

SAMPLE	MODEL	HAZARD	N	E	CENSORED	LNLF	χ (LR)	χ(SCORE)
I	A	Death	2902	274	90.6	3918.4	126.3	118.1
I	A	Institution	2896	397	86.3	5632.7	190.9	176.2
I	A	Other	2914	538	81.5	7847.6	100.7	76.3
I	A	All	2914	1209	58.5	17686.0	130.6	125.2
I	B	Death	2902	274	90.6	3915.1	129.6	122.7
I	B	Institution	2896	397	86.3	5620.2	203.4	193.3
I	B	Other	2914	538	81.5	7848.6	99.7	74.1
I	B	All	2914	1209	58.5	17677.8	138.8	133.5
I	C	Death	2902	274	90.6	3879.4	165.4	168.9
I	C	Institution	2986	397	86.7	5578.9	244.6	268.0
I	C	Other	2914	538	81.5	7838.2	110.1	77.6
I	C	All	2914	1209	58.5	17630.6	186.0	181.2
I	D	Death	2902	274	90.6	3879.4	165.3	169.0
I	D	Institution	2896	397	86.3	5579.6	243.9	269.0
I	D	Other	2914	538	81.5	7839.5	108.8	76.8
I	D	All	2914	1209	58.5	17630.7	185.9	182.0
II	A	Death	4950	446	91.0	6764.8	210.0	197.9
II	A	Institution	4944	651	86.8	9701.4	416.7	378.6
II	A	Other	4962	746	85.0	11631.0	156.9	131.9
II	A	All	4962	1843	62.9	28552.9	327.7	313.7
II	B	Death	4950	446	91.0	6760.6	214.1	202.9
II	B	Institution	4944	651	86.8	9685.9	432.1	401.0
II	B	Other	4962	746	85.0	11630.9	157.0	129.6
II	B	All	4962	1843	62.9	28537.2	343.5	328.5
II	C	Death	4950	446	91.0	6737.7	237.0	230.2
II	C	Institution	4944	651	86.8	9622.9	495.1	561.5
II	C	Other	4962	746	85.0	11611.0	176.9	132.5
II	C	All	4962	1843	62.9	28450.1	430.5	430.2
II	D	Death	4950	446	91.0	6737.8	237.0	230.2
II	D	Institution	4944	651	86.8	9620.7	497.3	565.7
II	D	Other	4962	746	85.0	11614.4	173.4	130.4
II	D	All	4962	1843	62.9	28450.9	429.7	430.3

Table 5.8.b Model Selection Statistics, Summary Statistics

	Likelihood Ratio Test Statistic				Score Test Statistic			
	Min	Max	Median	Mean	Min	Max	Median	Mean
A	100.69	416.66	173.89	207.47	76.27	378.60	154.02	189.72
B	99.69	432.07	180.18	214.77	74.11	400.99	163.39	198.20
C	110.15	495.06	211.51	255.70	77.64	561.55	205.72	256.26
D	108.82	497.31	211.44	255.18	76.83	565.71	206.12	256.70
I	99.69	244.64	152.07	158.14	74.11	269.04	151.20	150.75
II	156.89	497.31	282.37	308.43	129.59	565.71	271.95	299.69
Death	126.33	237.03	187.69	185.60	118.13	230.21	183.43	179.99
Institution	190.89	497.31	330.65	340.49	176.18	565.71	323.82	351.66
Other	99.69	176.87	133.52	135.44	74.11	132.46	103.61	103.65
All	130.61	430.51	256.85	271.59	125.18	430.34	247.86	265.58

The test statistics are generated from a test comparing the hypothesis of no covariates to that of all the covariates (in this respect, similar to the F-test). The larger the test statistic, the more likely that the covariates explain variations in individual utilization. Hence, the larger the test statistic, the better the model. The problem with extending this logic further is that a different sample population may result in a different ordering of the model results.

Nonetheless, the following observations are made. It appears that the evidence favours models C and D over A and B. Given Table 5.8.b, it is impossible to distinguish between C and D other than the grounds of theory consistency. That is to say that since we started with the hypothesis of self-selection bias, we should support the model that corrects it even if the model fit does not improve.

The evidence in favour of sample II is surprising since the assumption has been that sample I should be more valid. This supports an observation made earlier in that the sample characteristics from excluded clients does not differ from those of the study clients. Lastly, the data support the institutional hazard definition of an event. The superiority of this hazard over the all and other categories may be expected for the reasons discussed earlier. The improved performance over the death as reasons for discharge indicate that the covariates in the model explain variations in utilization better than they predict death. This is not surprising in that the unobservable client need characteristics are probably correlated directly with the probability of death. For the purposes of testing proportionality, model D, sample II (less restrictive selection criteria), and treatment of death and other discharges as censored observations (i.e. not an event) will be assumed.

3. Proportionality Test Results

The hazard rate illustrates a very important assumption, that of proportionality. By assumption and for any two individuals, the ratio of hazards is a constant (i.e. independent over time). Four graphical tests are employed to test this assumption: stratified survival, log-minus-log (LML) survival, Schoeldfeld's (1982a,b) partial residual, and Therneau's (1990) martingale residual plots. These tests for proportionality are graphical in nature, and therefore are subject to investigator bias (Lin 1993). Yet these tests are necessary due to the importance of the of the proportionality assumption and the fact that violation of this assumption would require a separate statistical modelling strategy.

The tests are presented in Figures 5.15-17. The stratified and LML survival plots in Figure 5.15a-g are analyzed for parallelism. The overwhelming evidence is in favour of proportionality. For some plots (eg. Figure 5.15b,c and 5.15g), the curves are parallel for the first 1200 days but are not parallel for the last 500 days. Schoelnfeld's residuals are illustrated in Figured 5.16a-g and are analysed for an equal distribution around zero that holds true for most of the plots. Figure 5.16 demonstrates that there exist a few outliers without which there would be absolutely not doubt. Analyses of categorical data as in 5.16g are difficult to analyze. Therneau's martingale residuals are presented in Figures 5.17a-h. Due to the particular construction of this visual statistic, it is difficult to conclude an equal distribution around zero. This happens because of the unequal numbers of institutionalized (13.1%) around '1' and the remainder around '0'.

In conclusion, I would suggest that the evidence support the fact that the assumption of proportionality is not violated for the main Cox model.

E. *Summary Interpretation of Data and Results*

In this section, covariate assumptions and evidence are reviewed for the auxiliary and Cox proportional hazard models. The results are then briefly summarized.

Table 5.9 summarizes the results for the two auxiliary models. The dark crosshatched pattern represents evidence against expectations, the lightly shaded cells represent a lack of evidence, and the clear cells represent the only supportive evidence.

Table 5.9: Evidence Favouring Service Intensity (FS and IS) Auxiliary Model Hypotheses

VARIABLE	AUXILIARY MODEL	
	FORMAL SERVICES (FS)	INFORMAL SERVICES (IS)
Age (AGE)	[illegible]	-
Gender (SEX)	0	0
Formal Service Intensity (FS)	n/a	[illegible]
Informal Support Service Intensity (IS)	-	n/a
Physical Function (PFn)	-	[illegible]
Mental Function (MFn)	-	0
Weighted Comorbidity Index (WCI)	[illegible]	+
Capital Health Authority / Edmonton Board of Health (CHA)	0	n/a
Nursing Home Waitlist (WTL)	0	n/a
Ratio of Urgent to Total Referrals (RUT)	0	n/a
English Speaking (ENG)	n/a	-
Married or Common Law (MS)	n/a	+
Live With Spouse or Others (LIV)	n/a	+
Family Size (FAM)	n/a	+
Income (INC)	n/a	+

Table 5.10 summarizes the hypotheses and results for the main Cox proportional hazards model. The dark crosshatched pattern represents evidence against expectations, the lightly shaded cells represent a lack of evidence, and the clear cells represent the only supportive evidence. Details are provided in the following subsections.

Table 5.10: Evidence Favouring the Main Model Hypotheses

VARIABLE	HAZARD			
	INSTITUTIONAL-IZATION	DEATH	OTHER DISCHARGES	ALL DISCHARGES
Age (AGE)	-	-	+	-
Gender (SEX)	+	+	0	+
Formal Service Intensity (FS)	+	+	0	+
Homecare Classification Score (HCS)	-	-	+	-
Homecare Informal Support Score (HCIS)	-	-	0	-
Homecare Functional Need Score (HCFN)	-	-	+	-
Informal Support Service Intensity (IS)	+	+	0	+
Physical Function (PFn)	+	+	-	+
Mental Function (MFn)	+	+		+
Weighted Comorbidity Index (WCI)	-	-	0	-

The progression from the Kaplan-Meier univariate to the stratified K-M to the various Cox models (models A, B and C) and finally to the all-inclusive Cox model (D) have resulted in increasingly favourable evidence for the formal resource intensity covariate. Regressing length of service on both formal and informal service intensity has resulted in less evidence for the importance in informal services. However, as pointed out earlier, informal service data is collected only at the time of assessments (only once in most cases) and by design informal service intensity is a constructed variable.

The evidence for the two auxiliary models is less compelling. The important finding is the confirmation of the hypothesis that the unobservable influences on length of service are related to the severity of the client's functional status or some other need variable. There is no evidence for system variables and slight evidence for environmental influences.

Finally, Figure 5.18 illustrates the cumulative survival curves for Model D and Sample II for each of the competing hazards. For example, the probability of remaining in the homecare program after one year is 72%. The median time for discharge is approximately three years. At one year an individual has a 7% probability of death, 9% probability of discharge to an institution, and 12% discharge for other reasons. As discussed earlier, these Figures assume independence (hence 7+9+12=28%). Multiplication of the predicted length of service by the predicted service intensity then can be performed to obtain predicted total resource requirements for any individual.

Figure 5.1: Formal Service Intensity: Composition by Provider Type

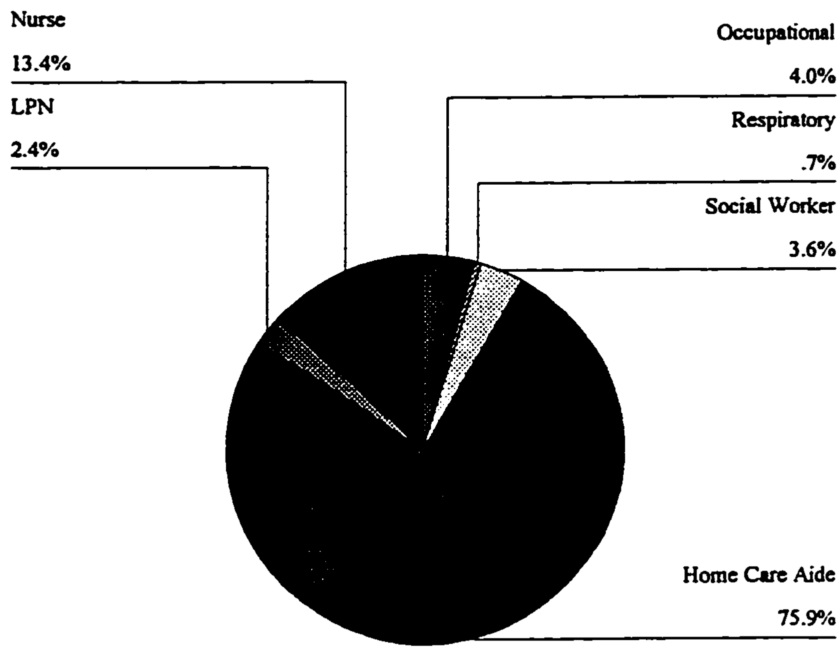


Figure 5.2 Formal Service Intensity: Compostion by Service Type

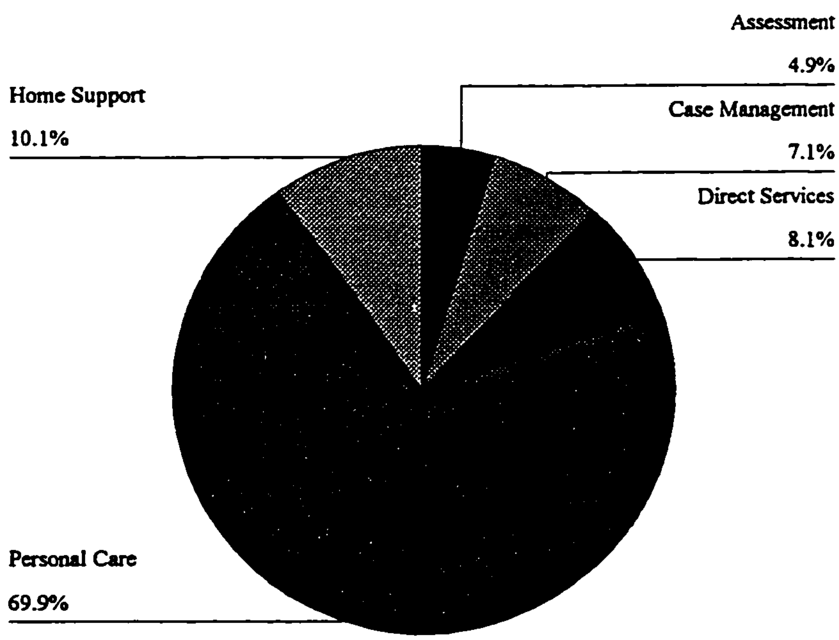


Figure 5.3: Average Service Intensity by Provider Type

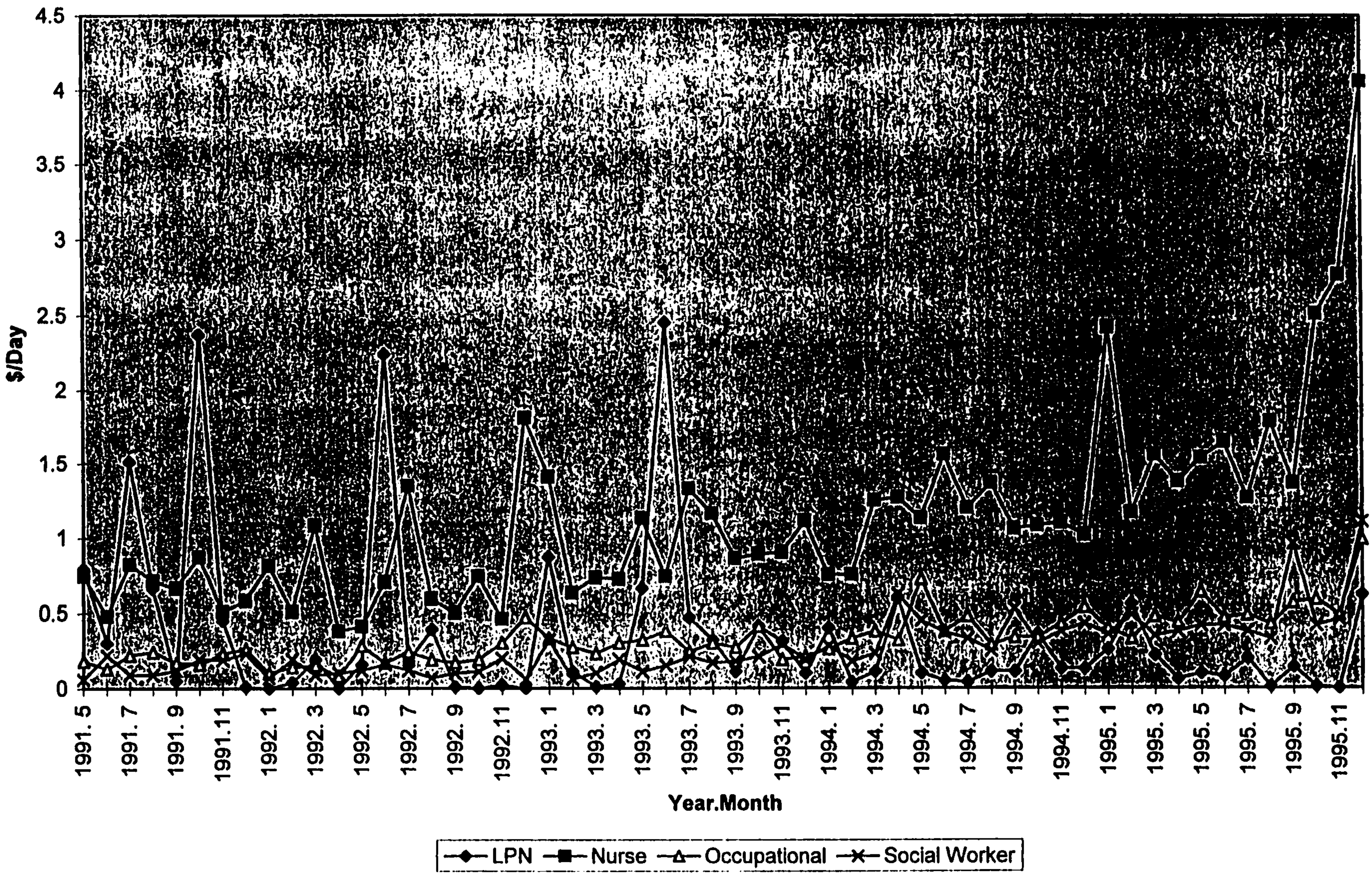


Figure 5.4: Average Monthly Service Intensity by Service Type

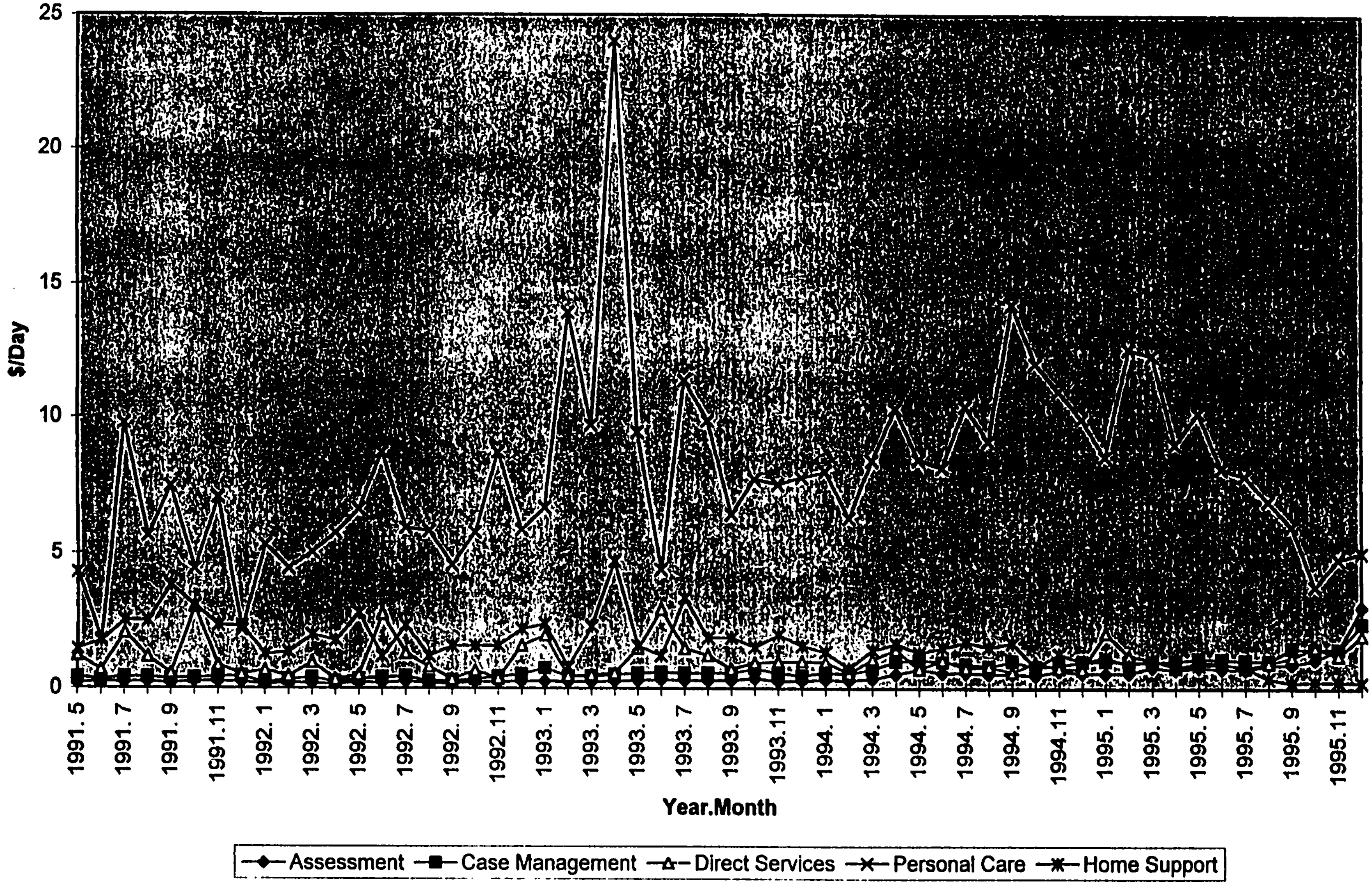


Figure 5.5: Event (Combined) History Diagram (N=4,962)

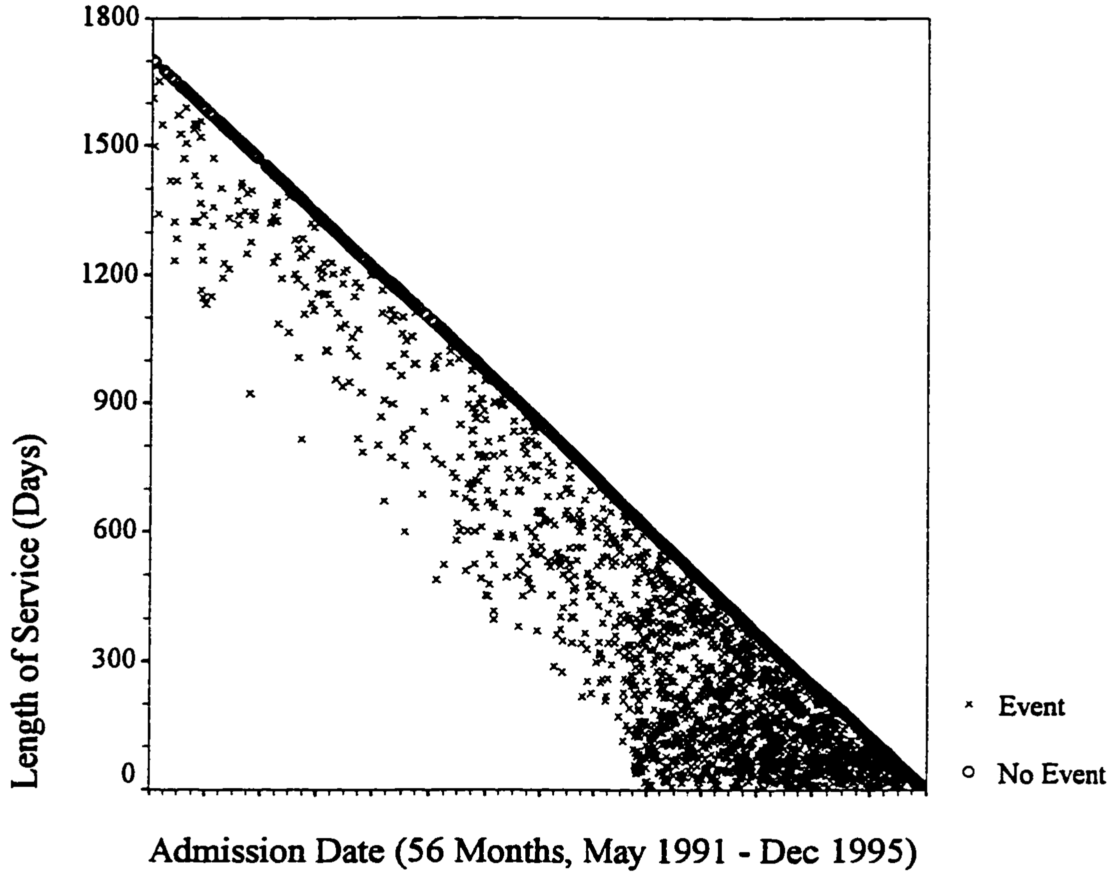


Figure 5.6: Proportion Discharged by Admission Month

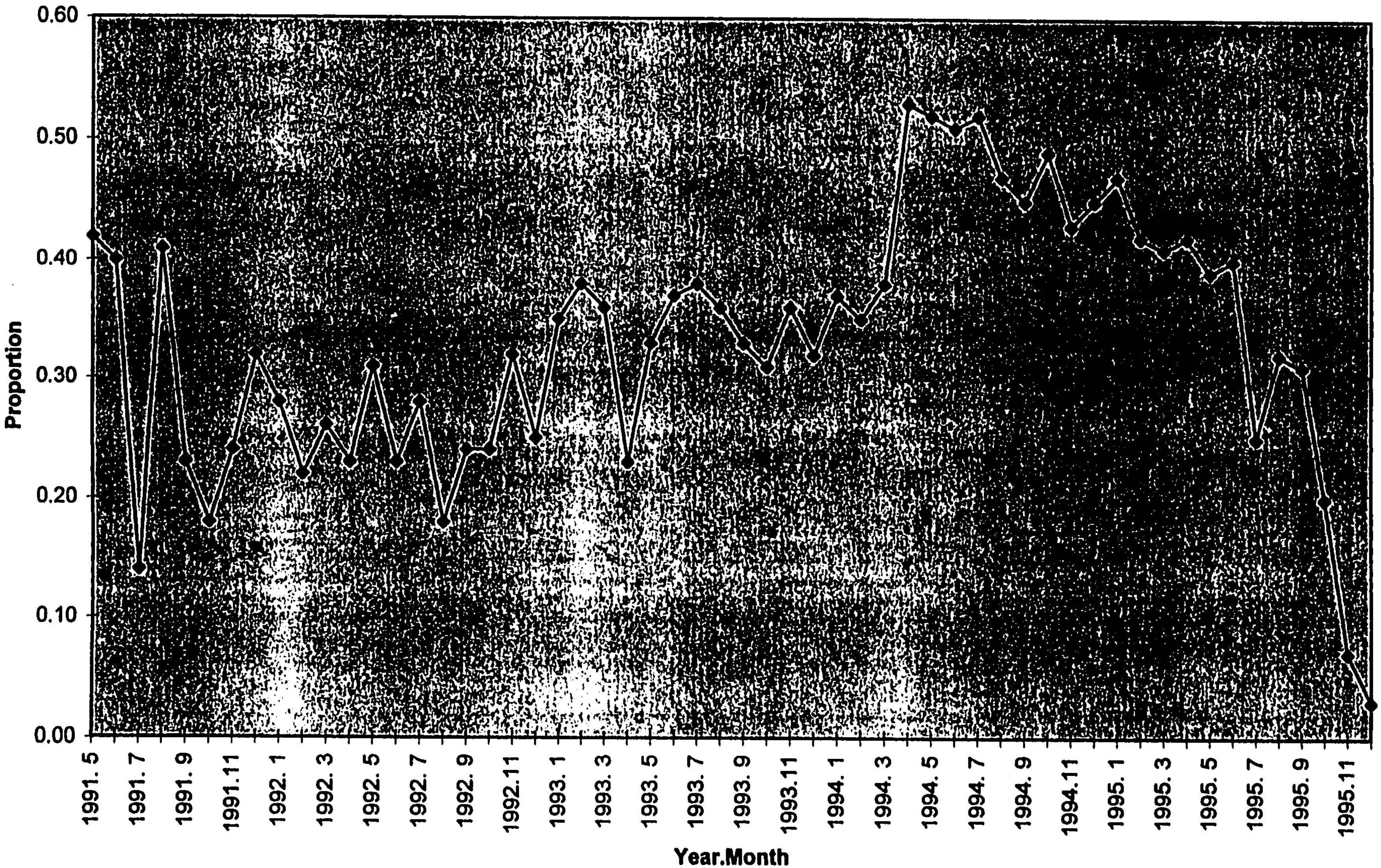


Figure 5.7: Formal Service Intensity Histogram

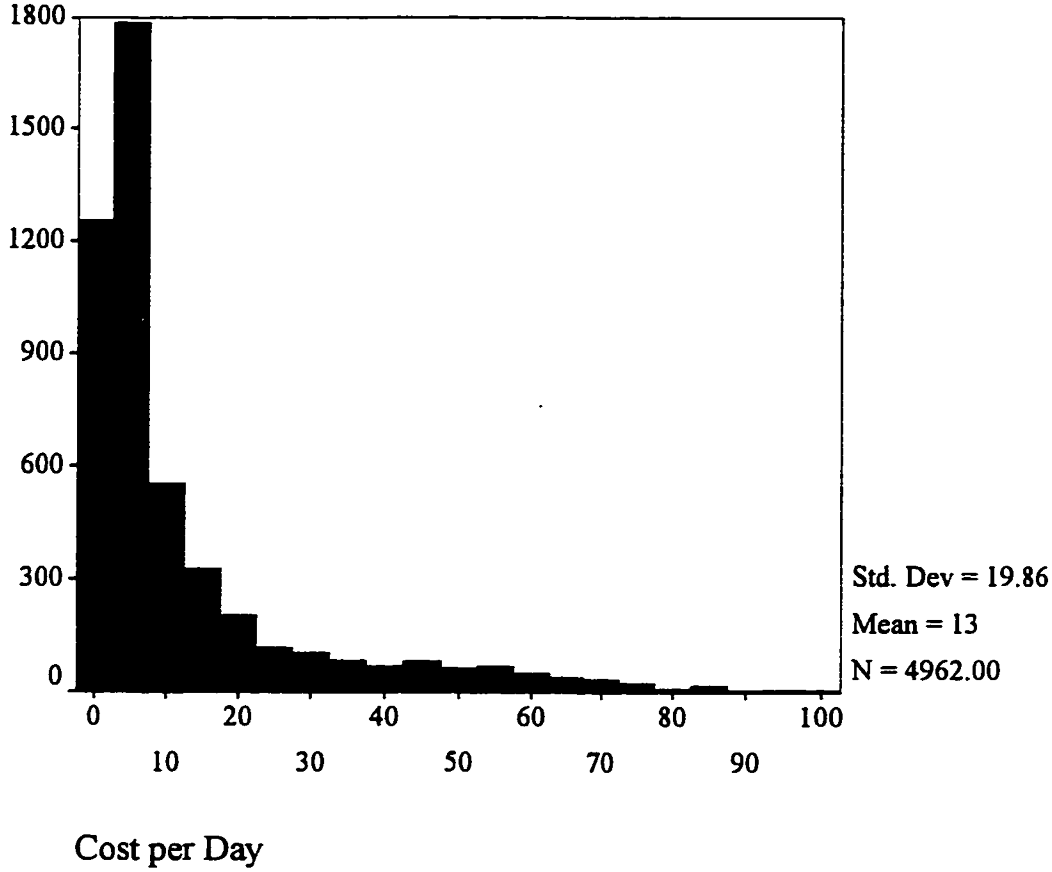


Figure 5.8: Average Service Intensity by Month of Admission

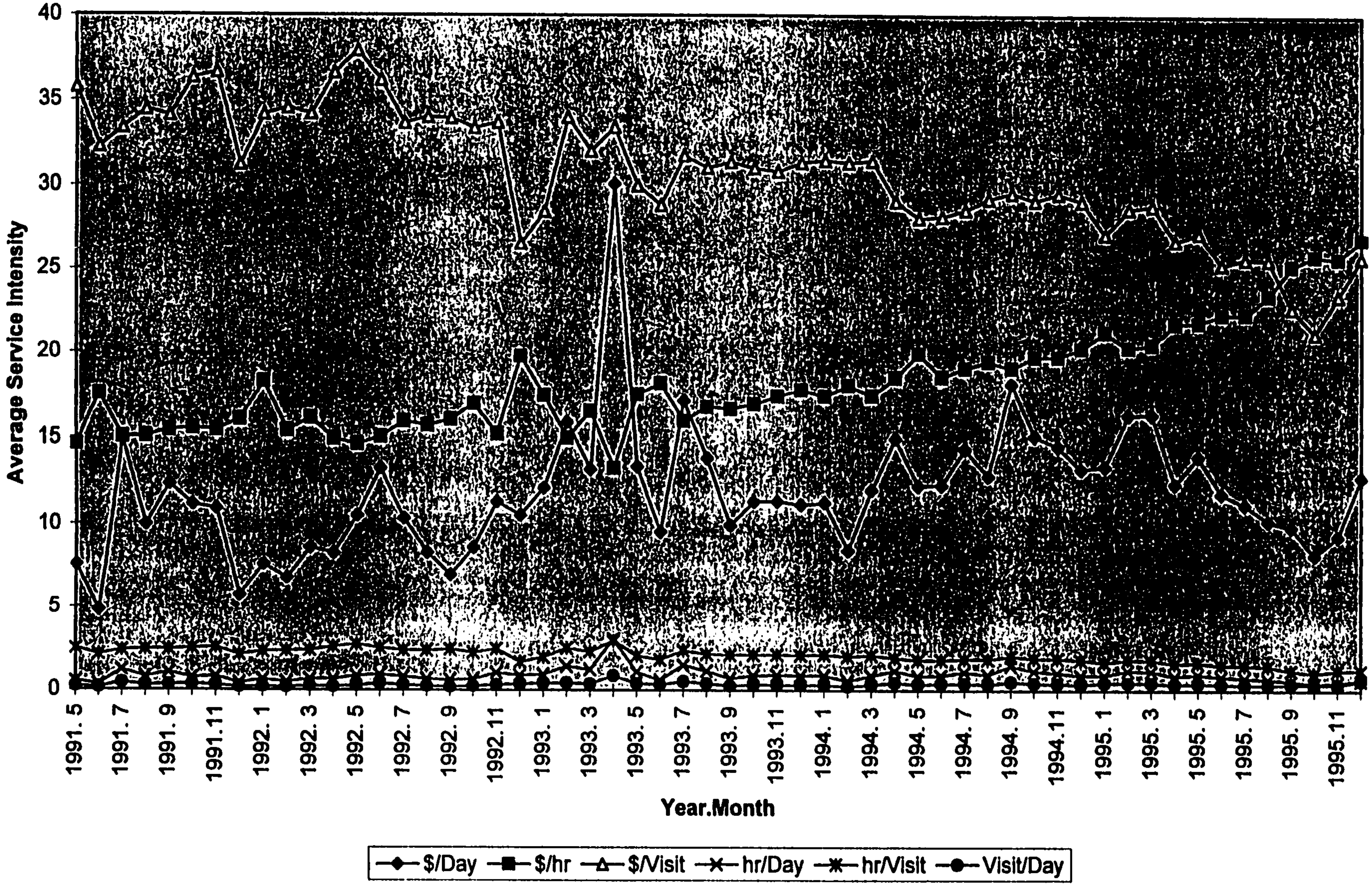


Figure 5.9: Client Age (at Admission) Histogram

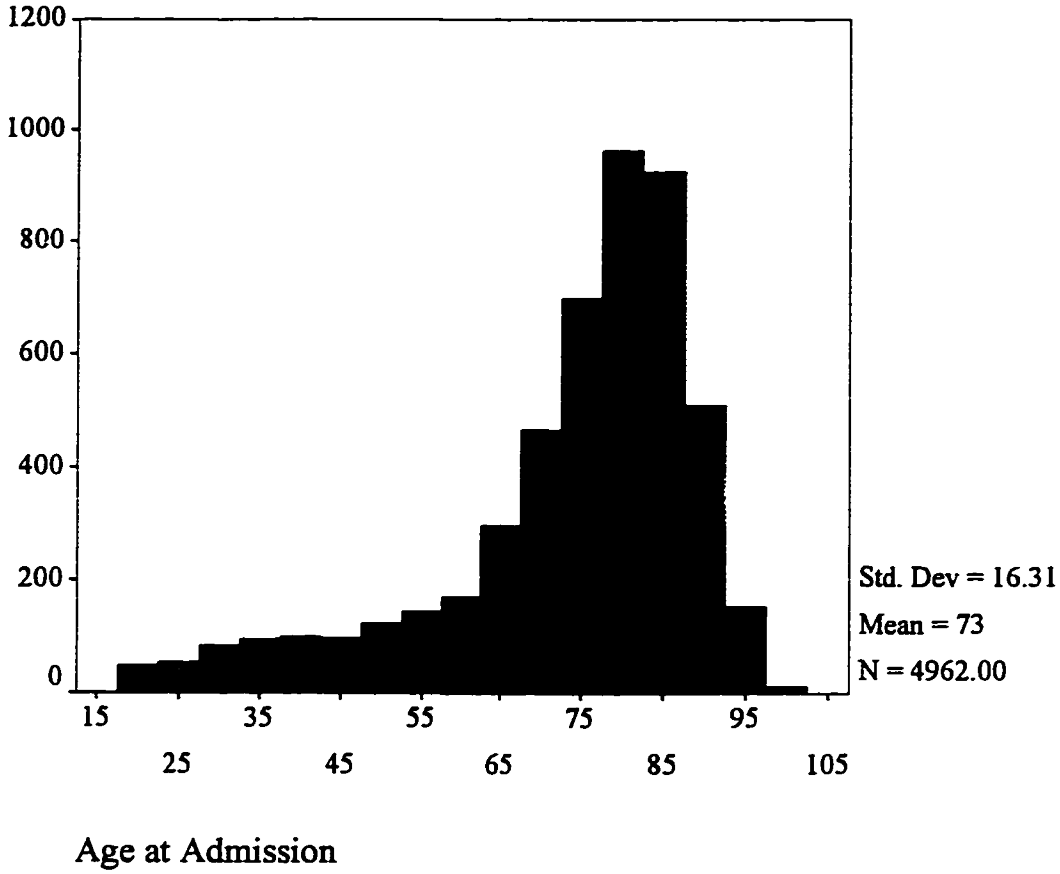


Figure 5.10: Average Client Age at Admission

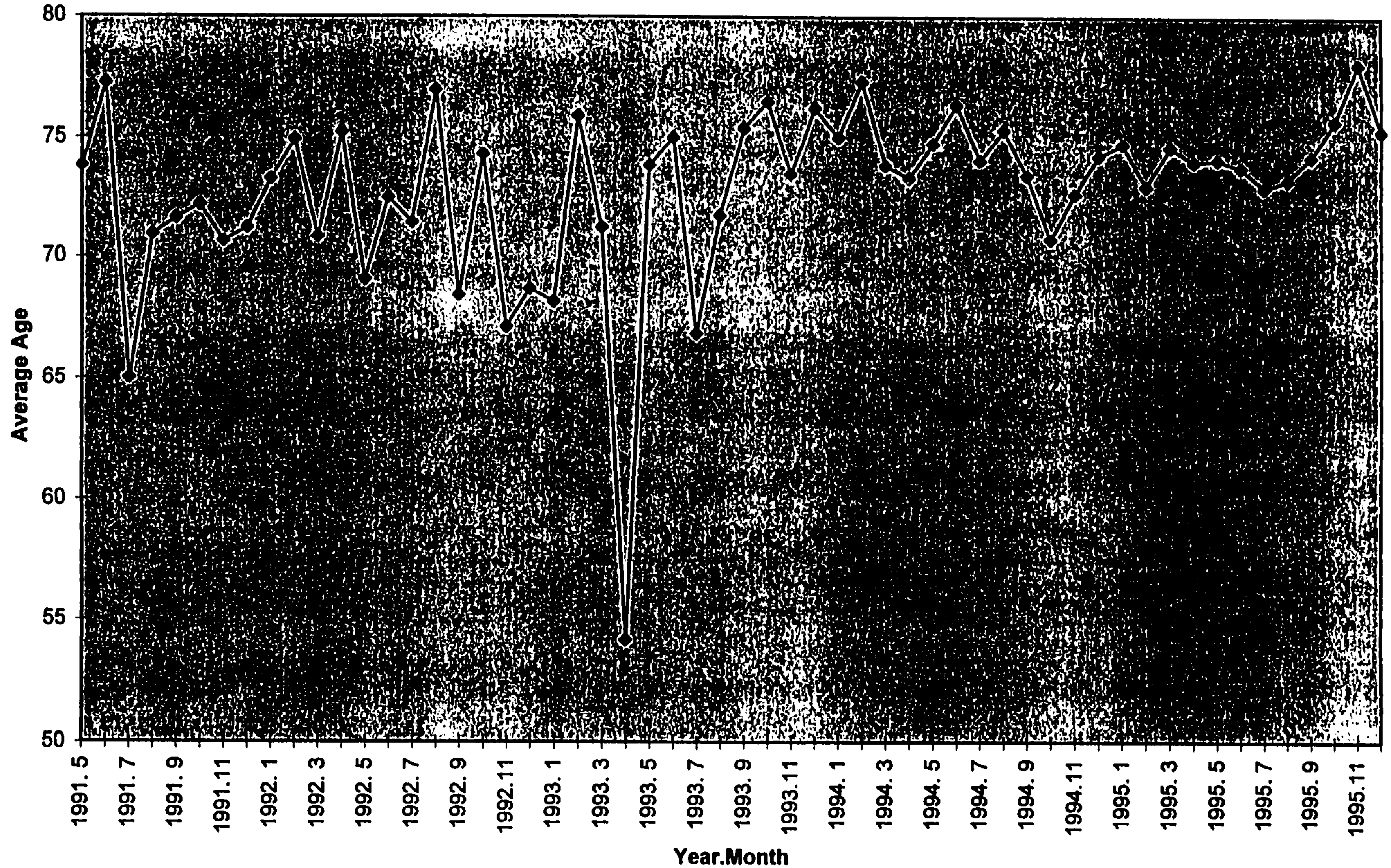


Figure 5.11: Proportion Female Admissions

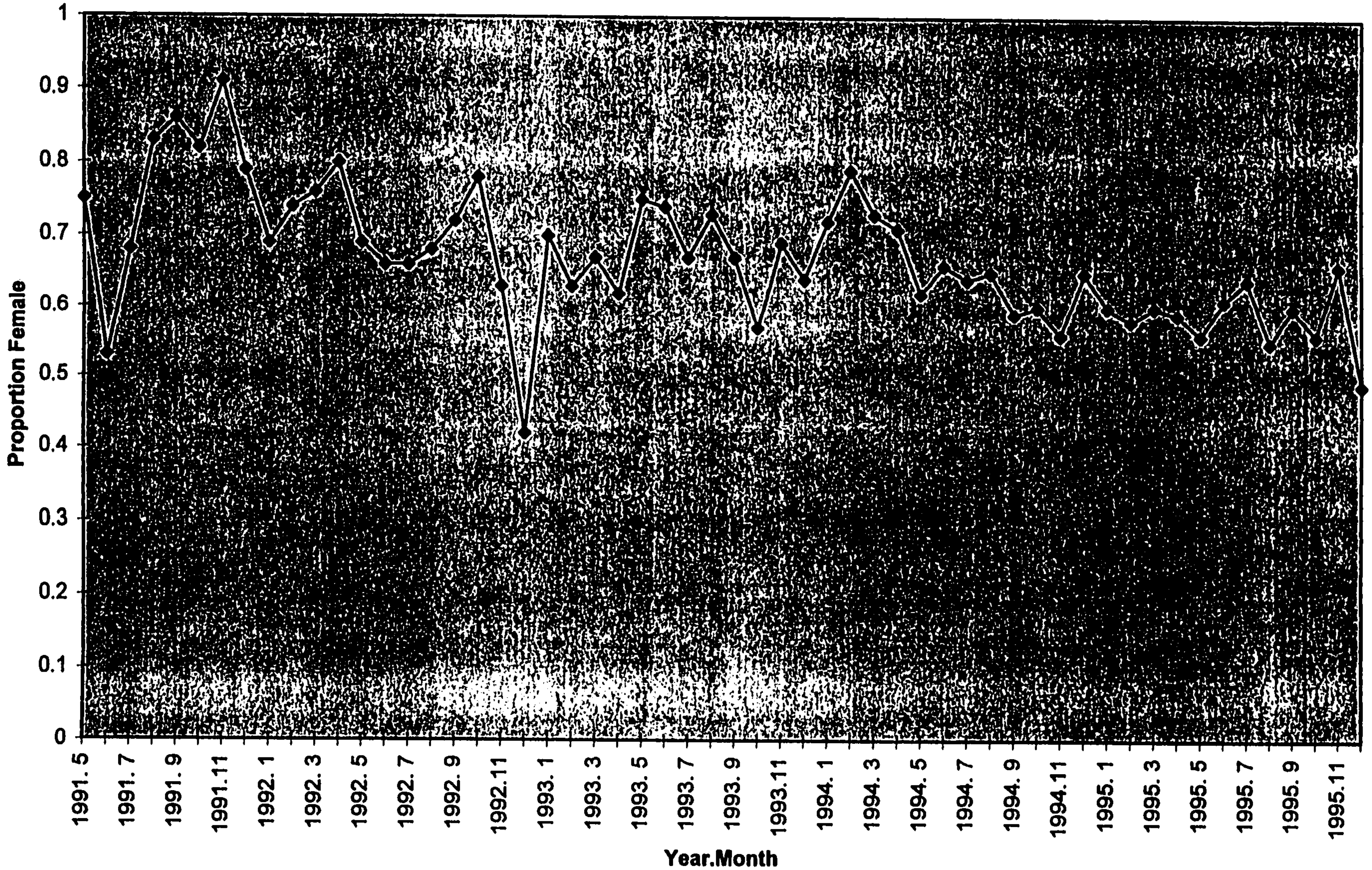


Figure 5.12: Average Client Need by Month

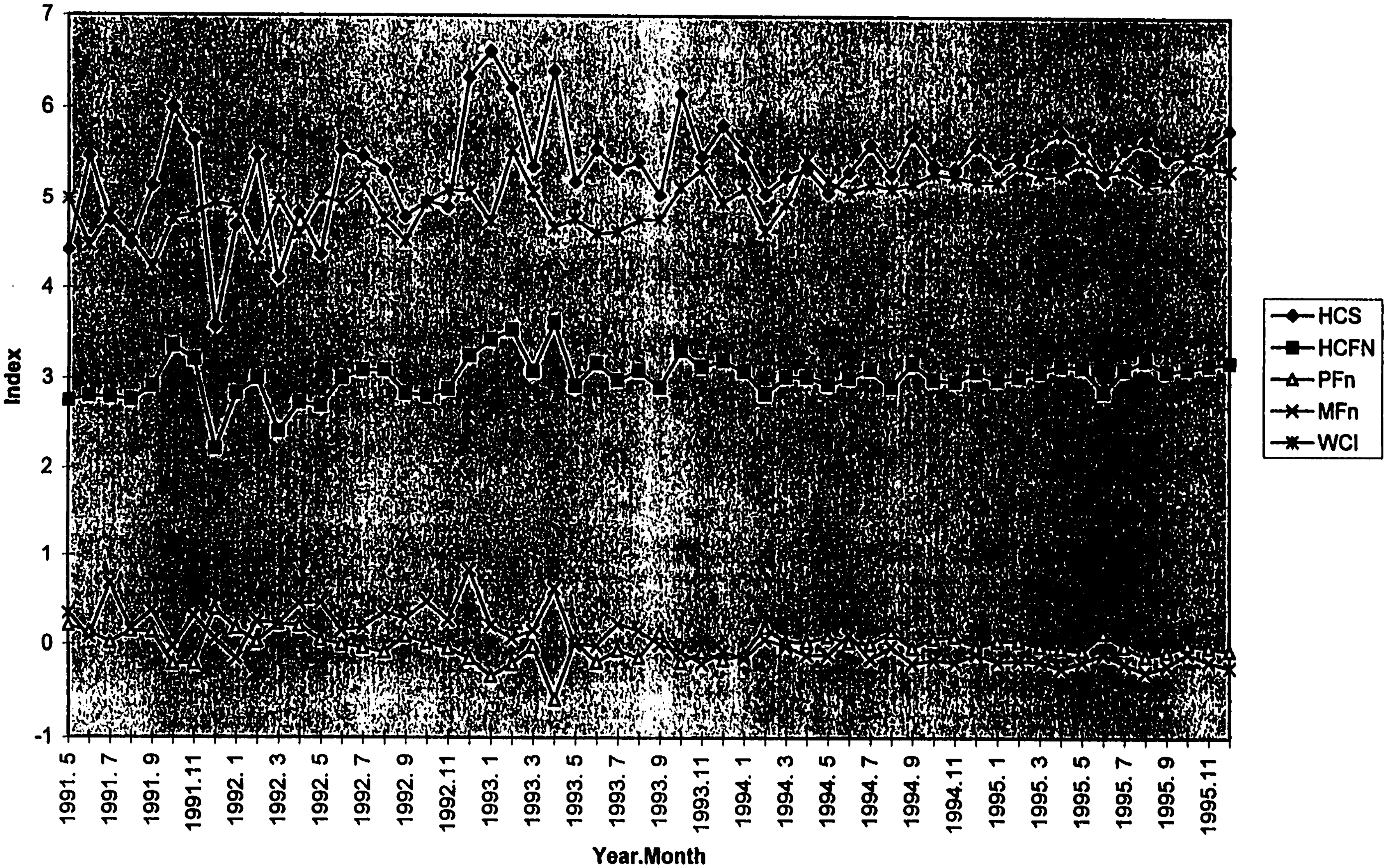
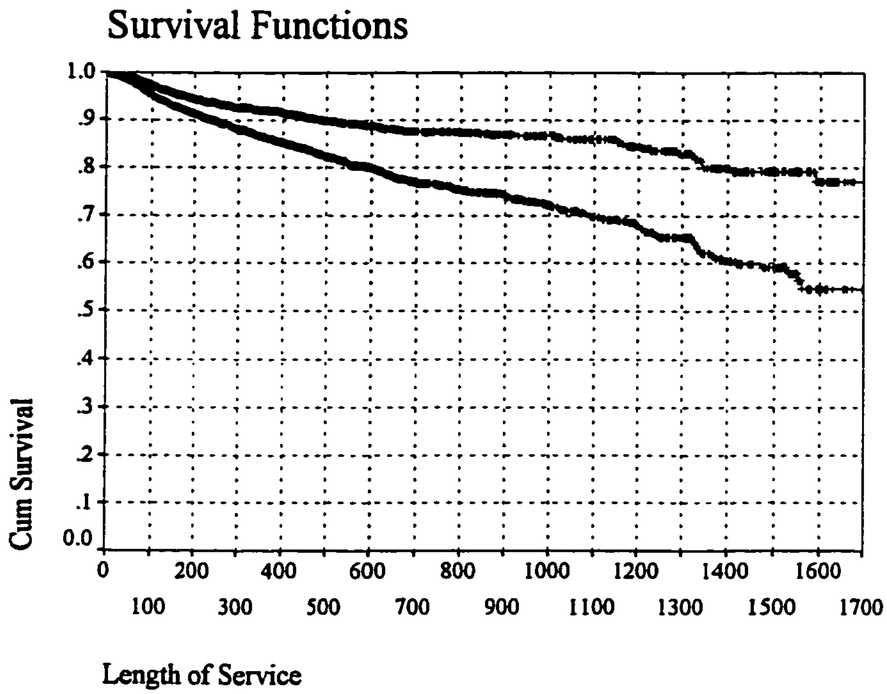
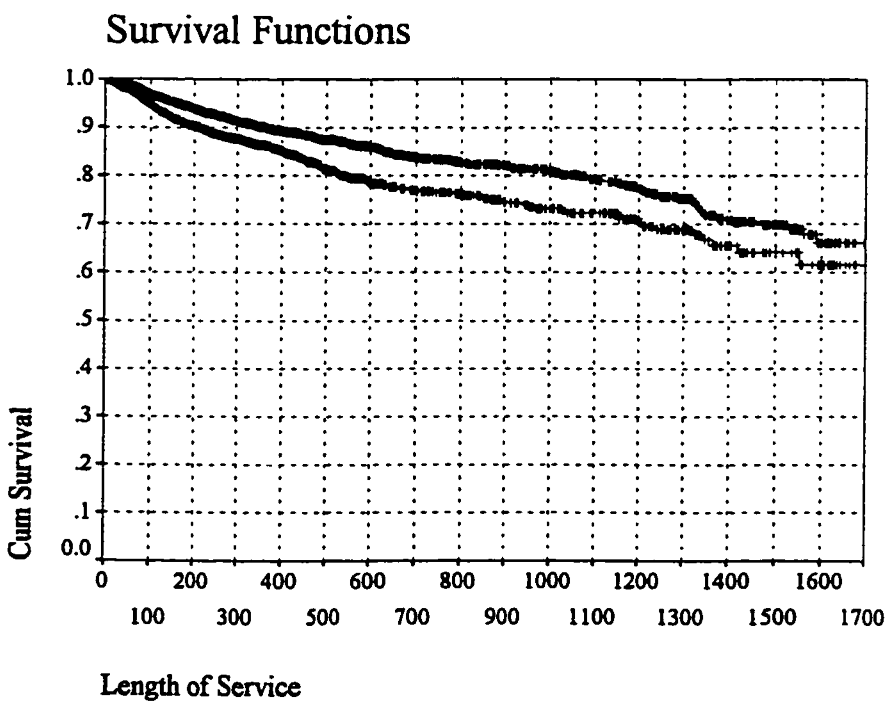


Figure 5.13: Univariate Kaplan-Meier Survival Functions

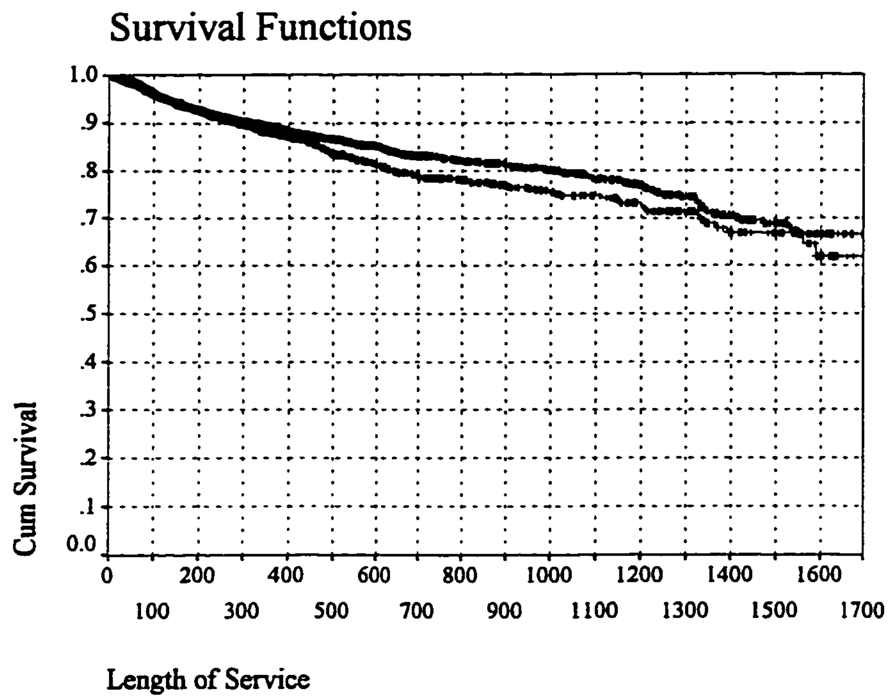
(a) Age



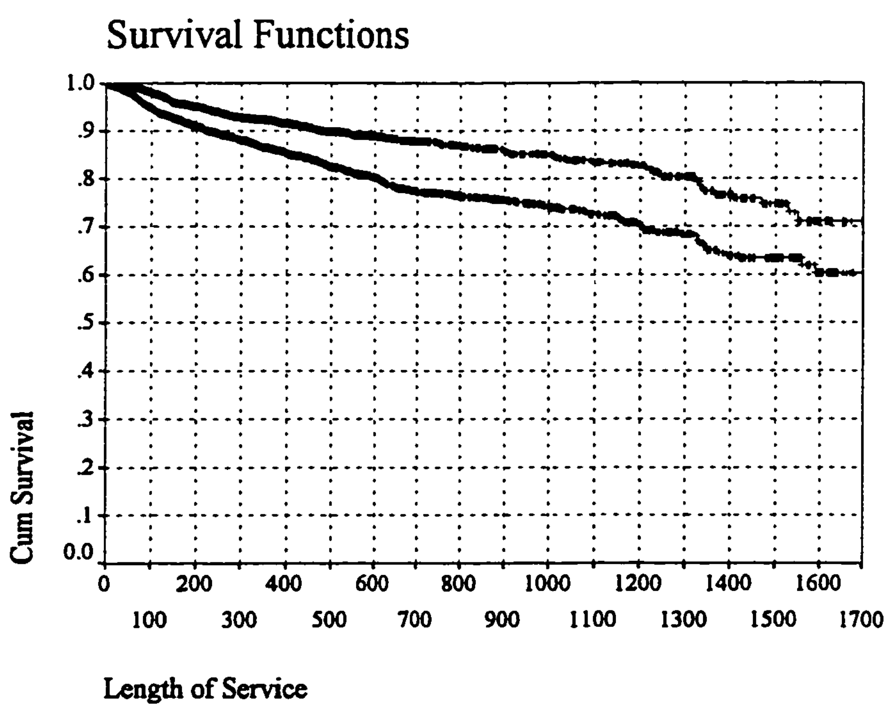
(b) Gender



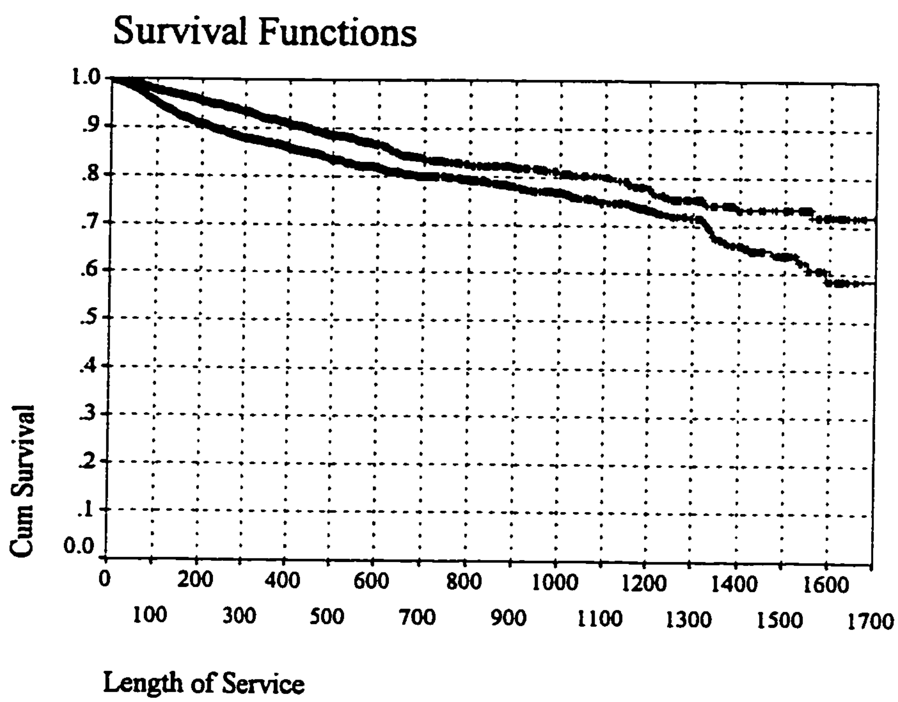
(c) Observed Formal Service Intensity



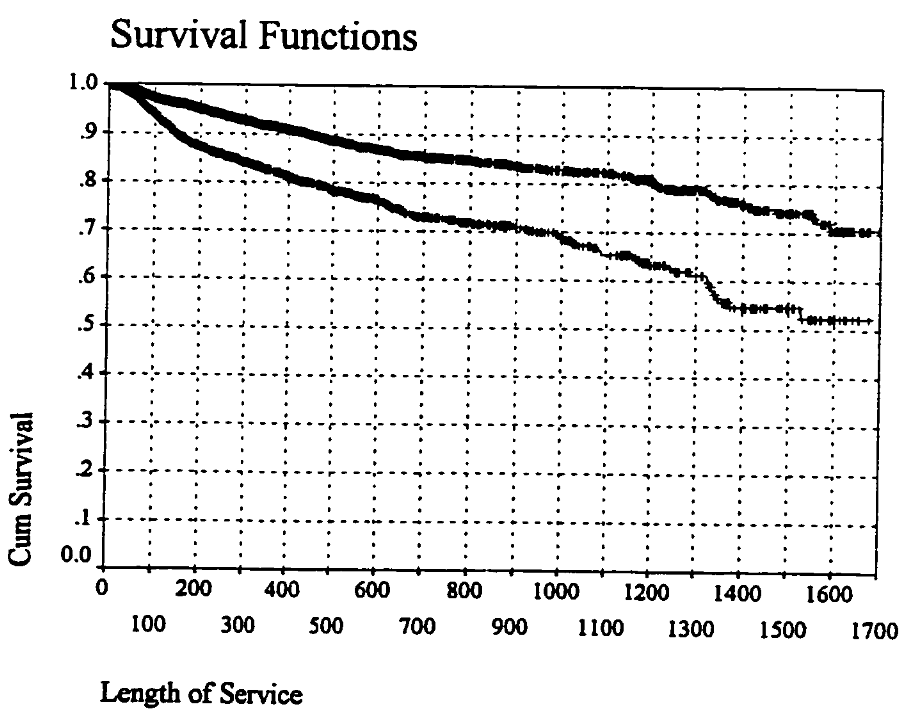
(d) Predicted Formal Service Intensity



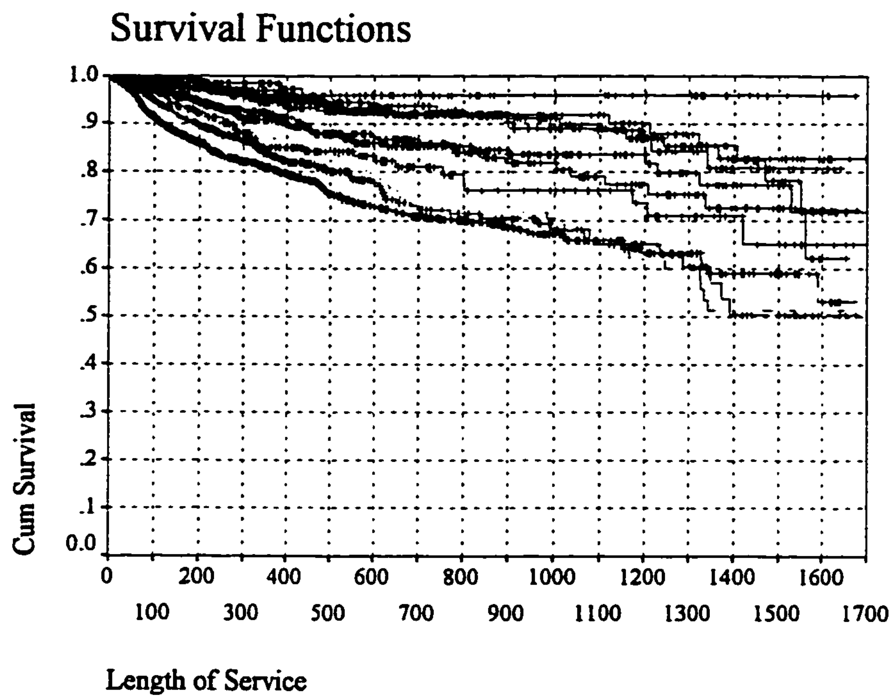
(e) Observed Informal Resource Intensity



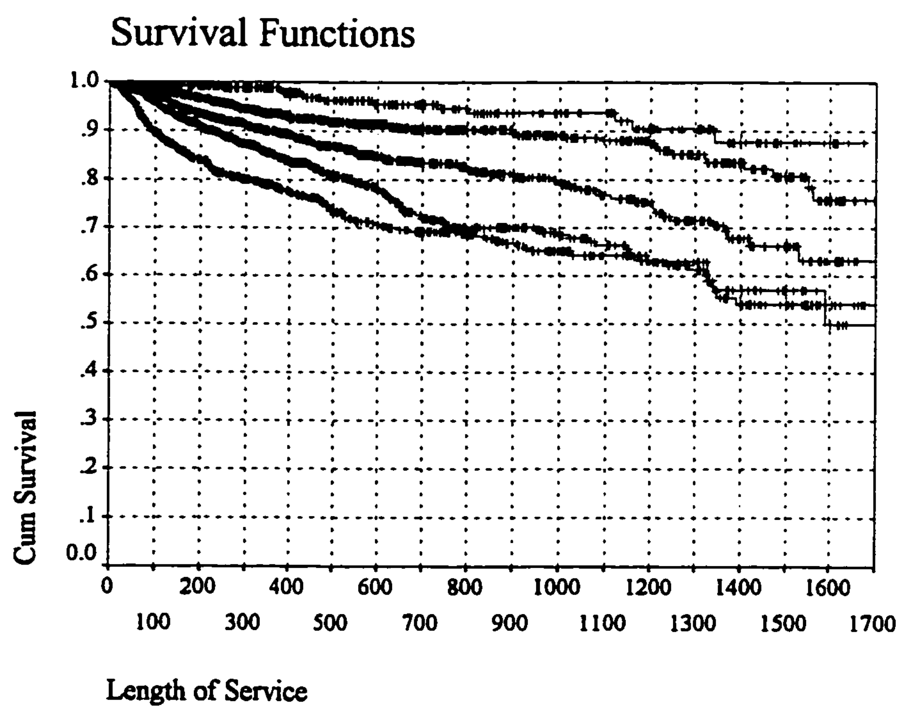
(f) Predicted Informal Resource Intensity



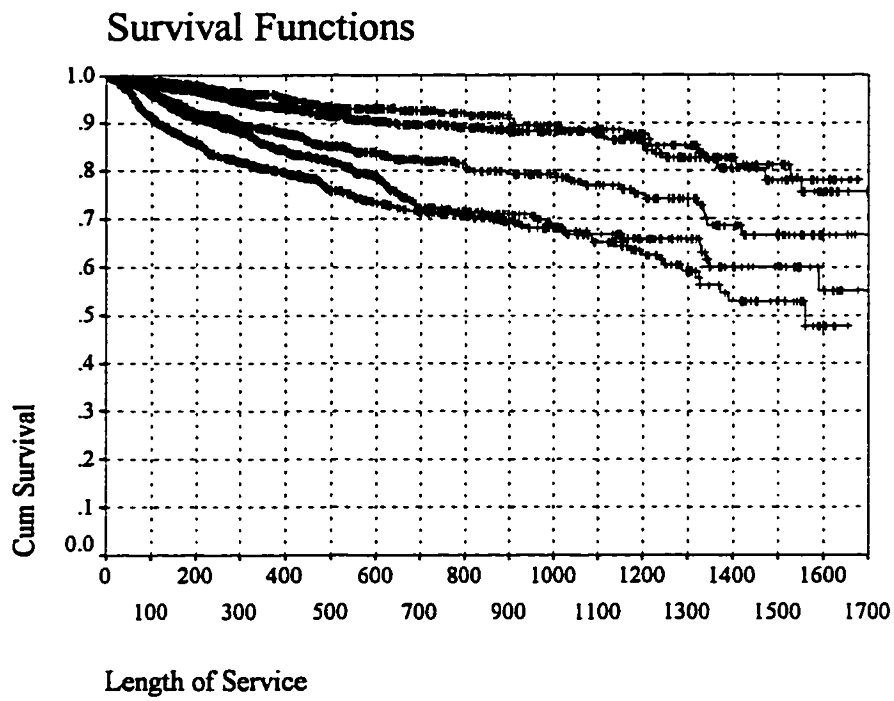
(g) HCS


Survival Functions
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
0.0
Cum Survival
0
200
400
600
800
1000
1200
1400
1600
100
300
500
700
900
1100
1300
1500
1700
Length of Service

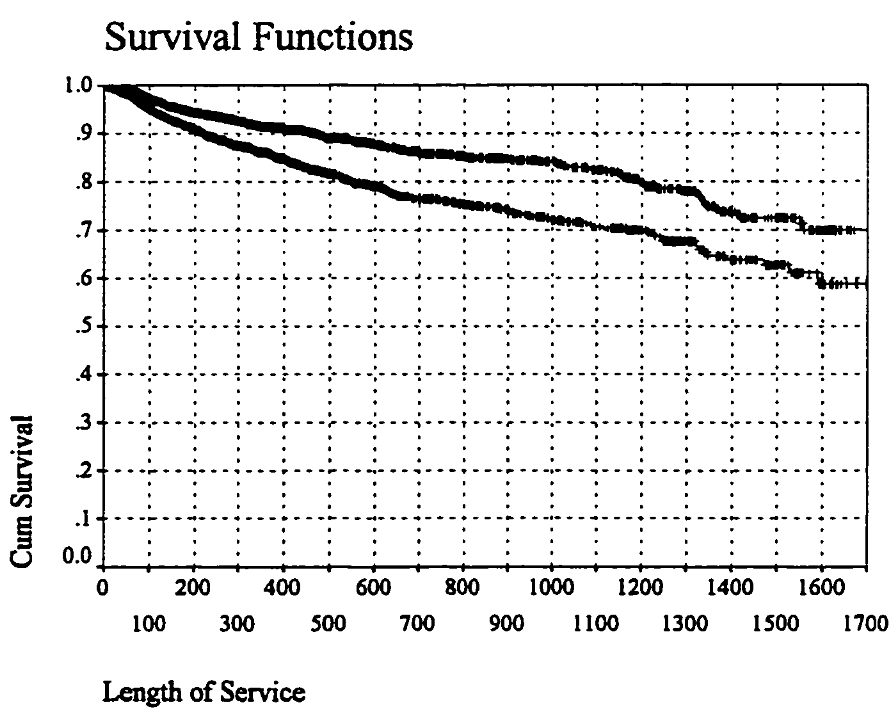
(h) HCIS


Survival Functions
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
0.0
Cum Survival
0
200
400
600
800
1000
1200
1400
1600
100
300
500
700
900
1100
1300
1500
1700
Length of Service

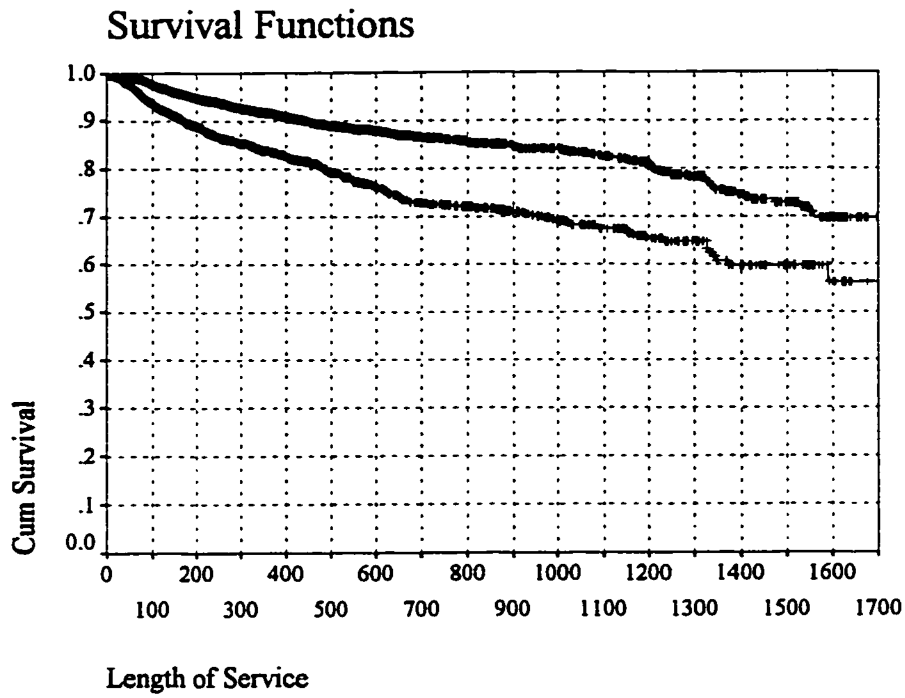
(i) HCFN



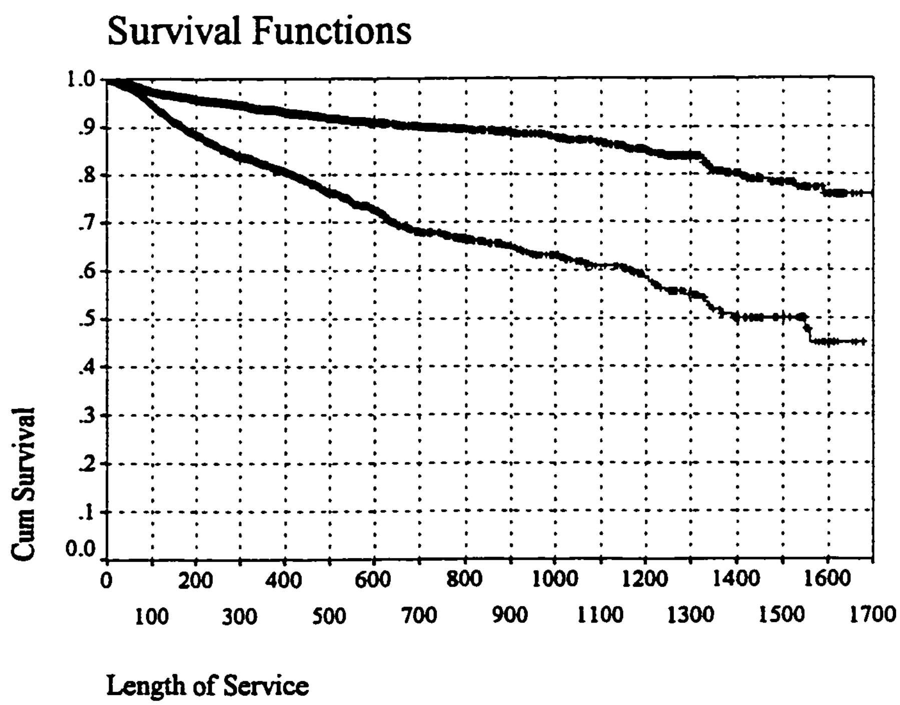
(j) Weighted Comorbidity Index



(k) Physical Function

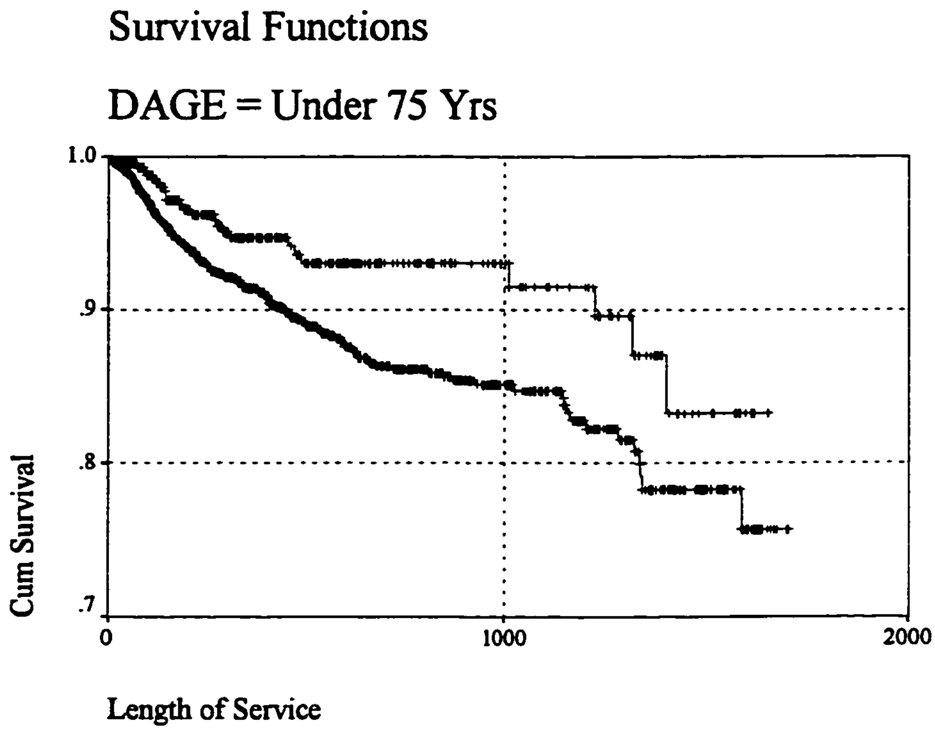


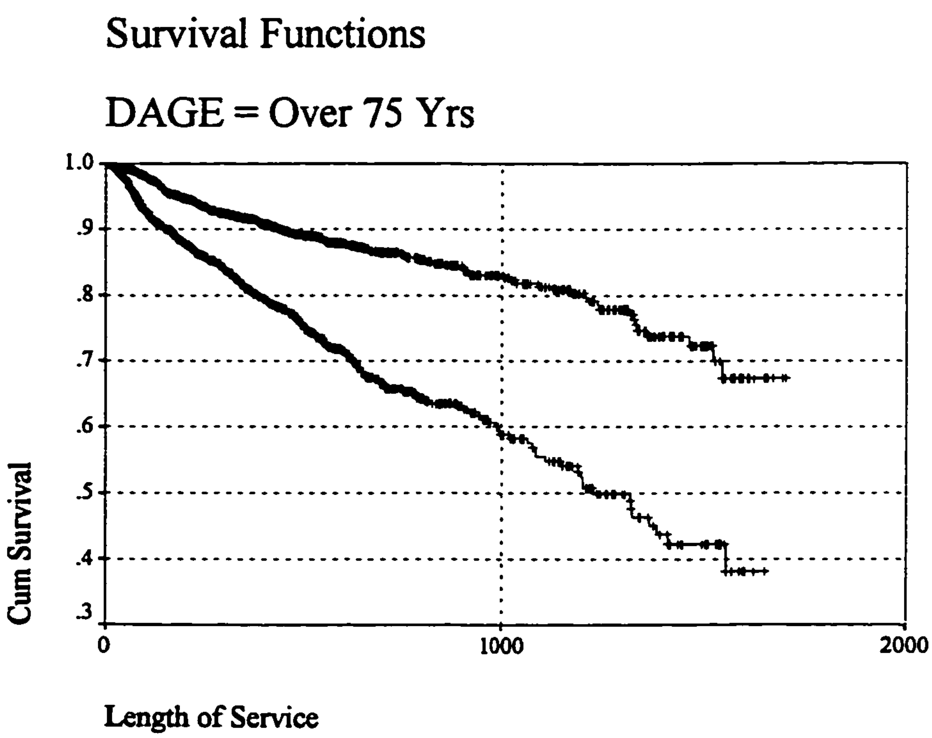
(l) Mental Function



5.14: Stratified Kaplan-Meier Results

(a) Age

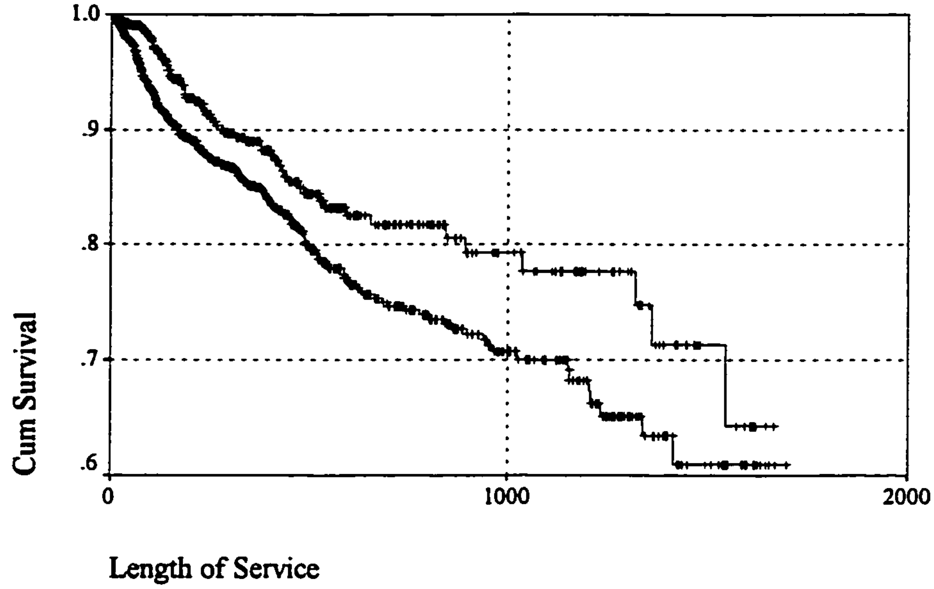




(b) Gender

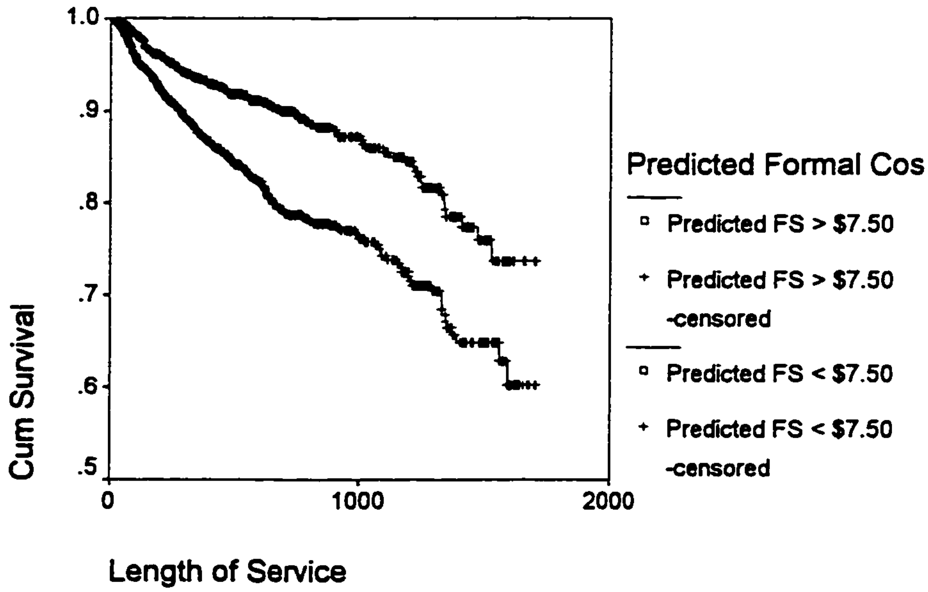
Survival Functions

SEX = Male



Survival Functions

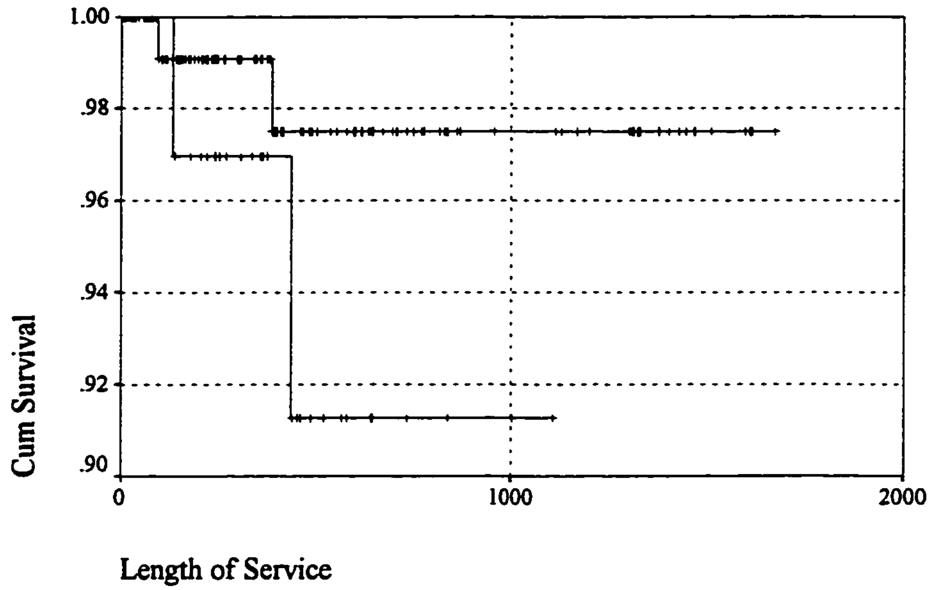
SEX = Female



(c) Combined Home Care Classification Score

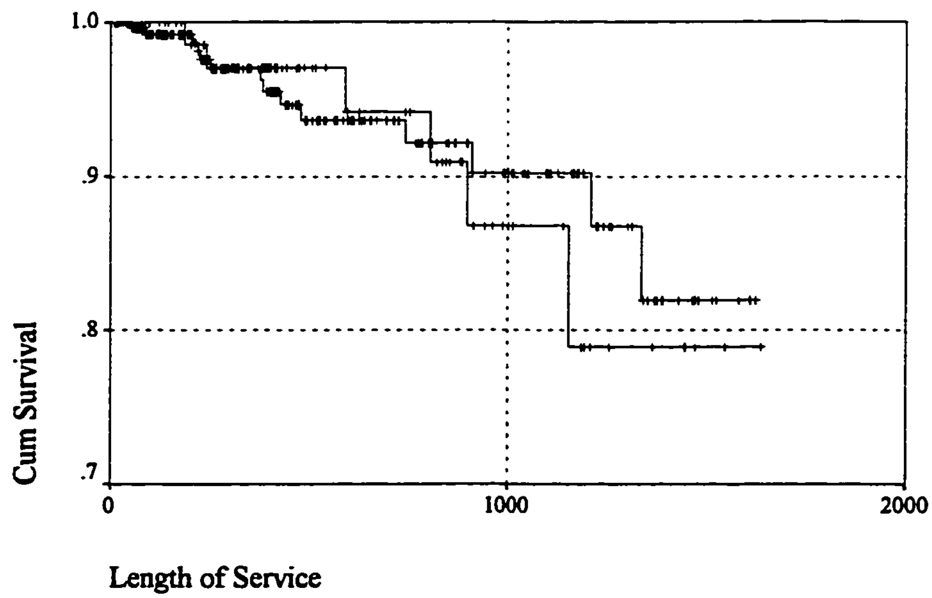
Survival Functions

HCS = 0



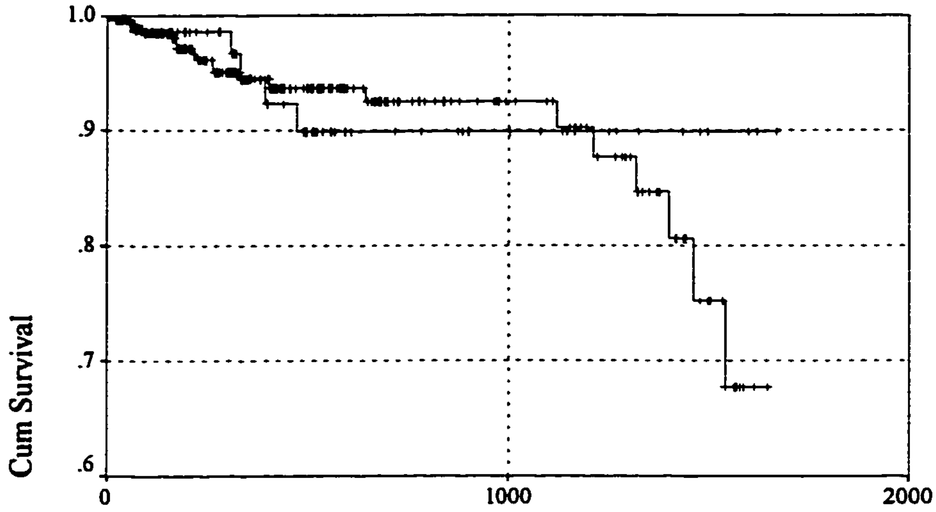
Survival Functions

HCS = 1



Survival Functions

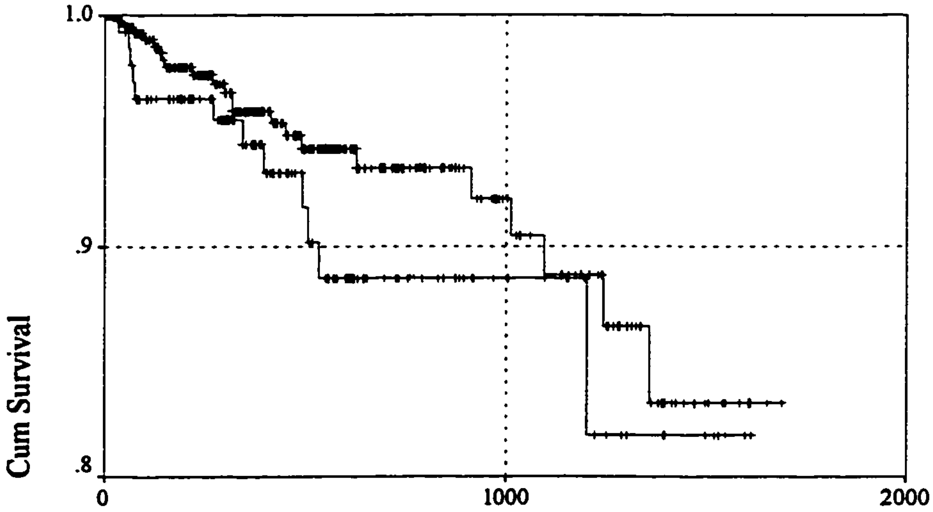
HCS = 2



Length of Service

Survival Functions

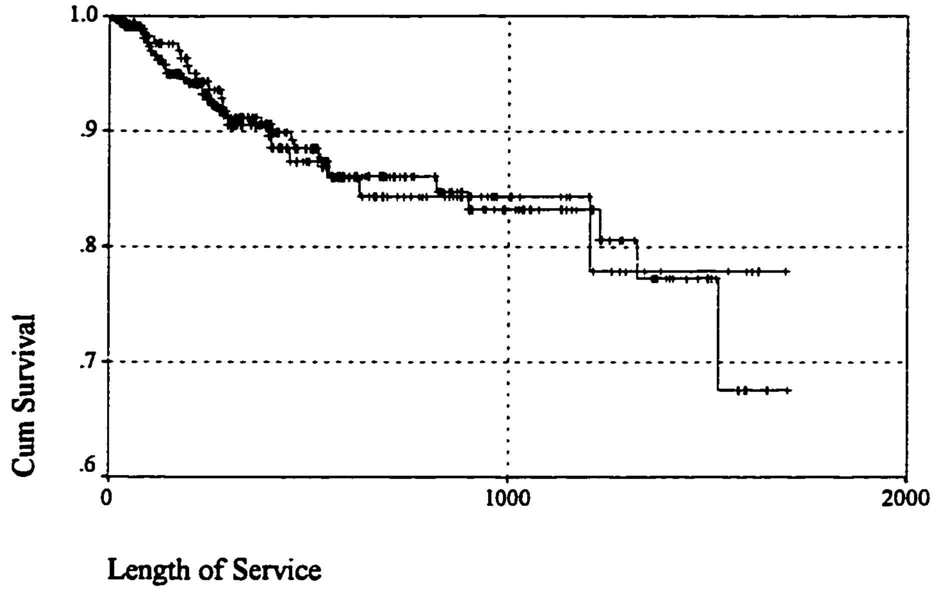
HCS = 3



Length of Service

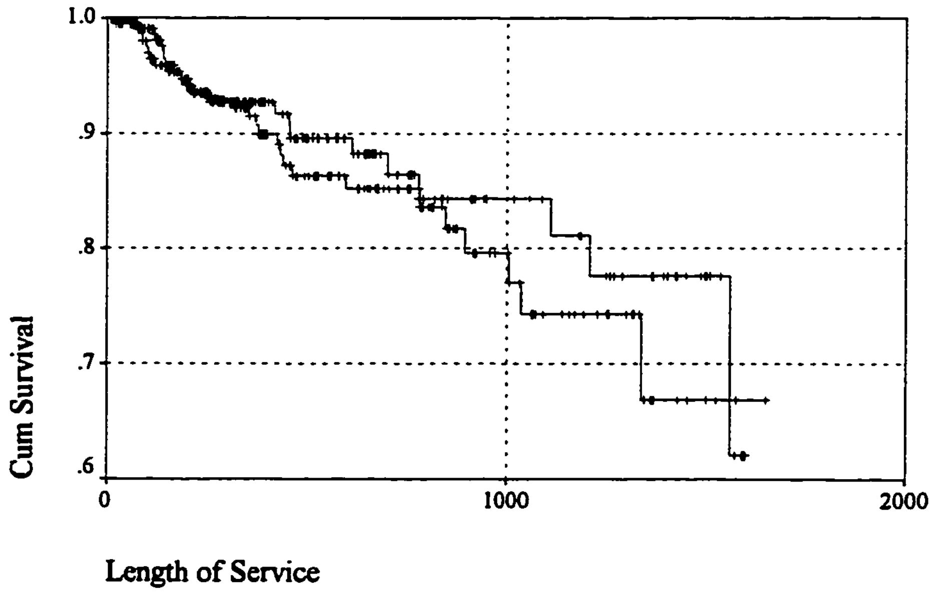
Survival Functions

HCS = 4



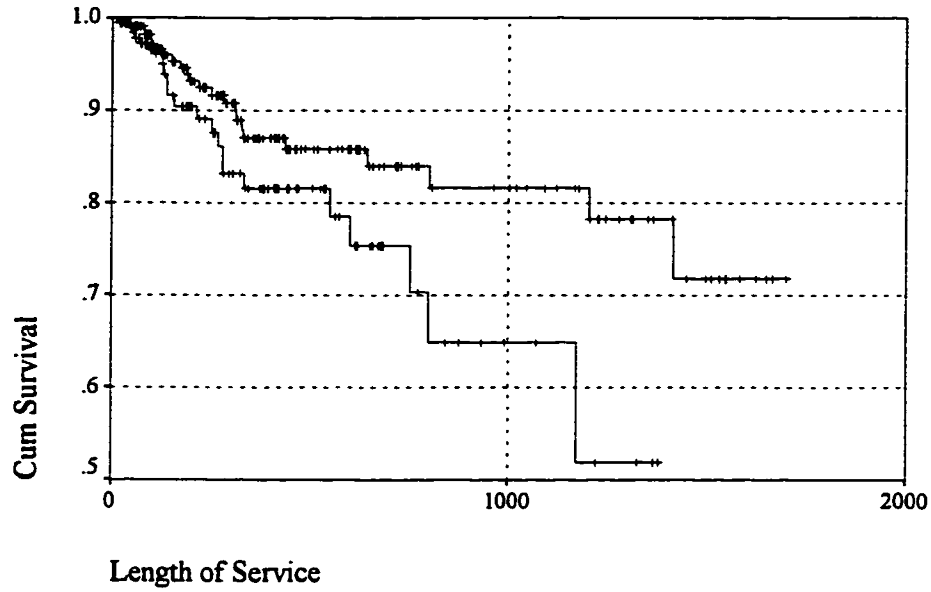
Survival Functions

HCS = 5



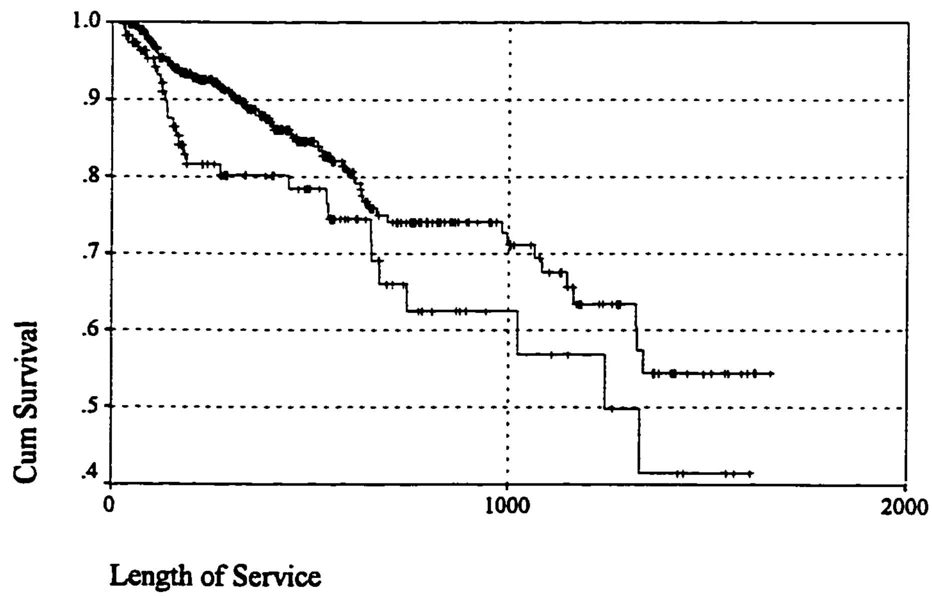
Survival Functions

HCS = 6



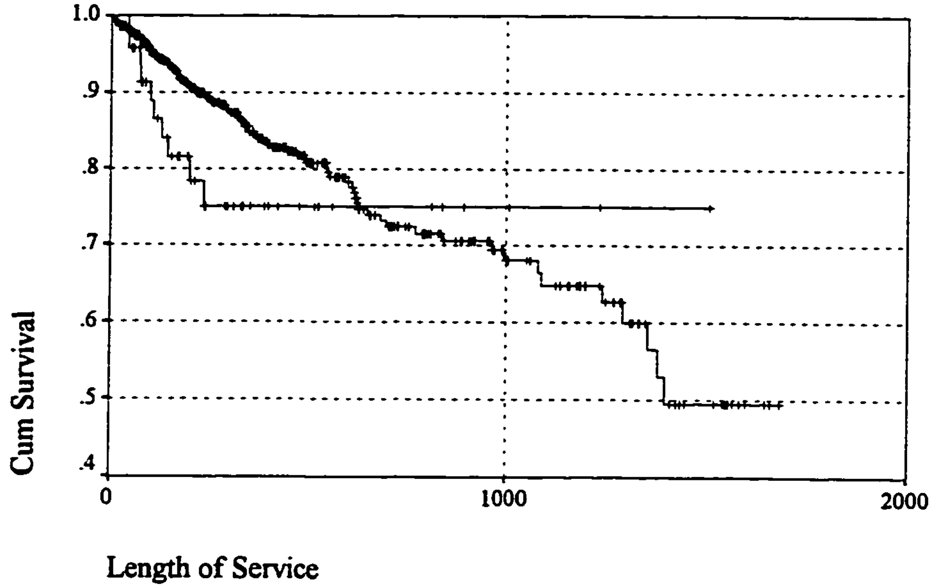
Survival Functions

HCS = 7



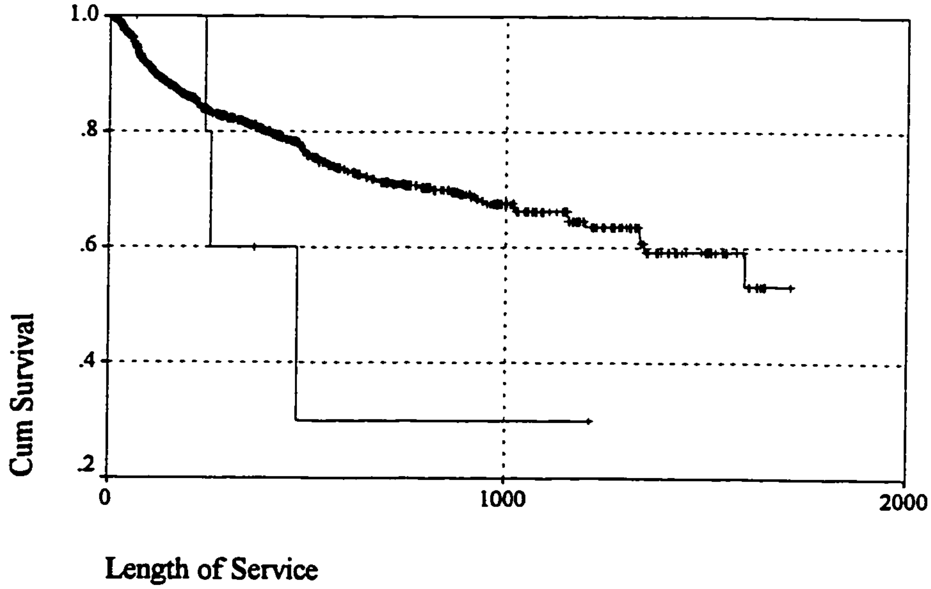
Survival Functions

HCS = 8



Survival Functions

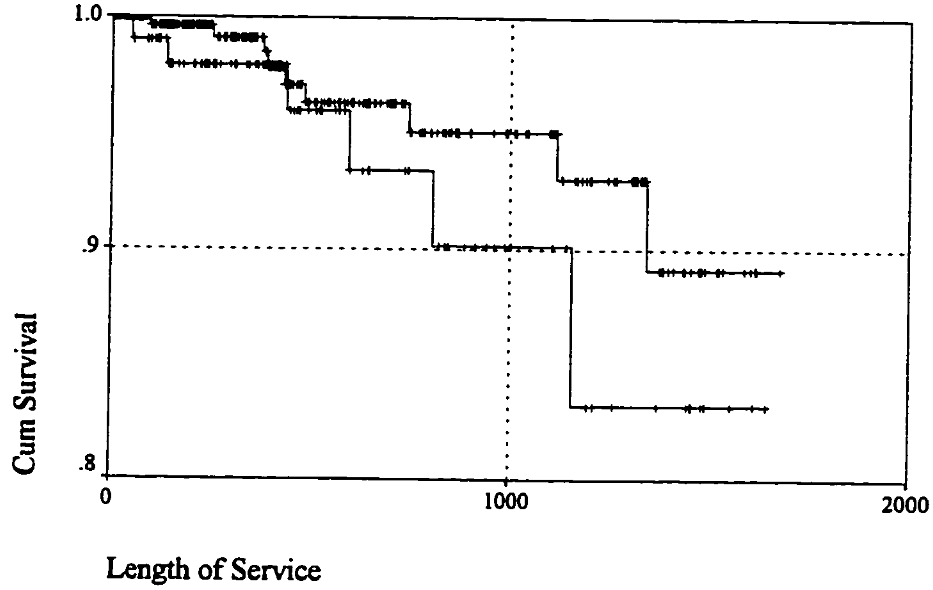
HCS = 9



(d) Informal Support Score

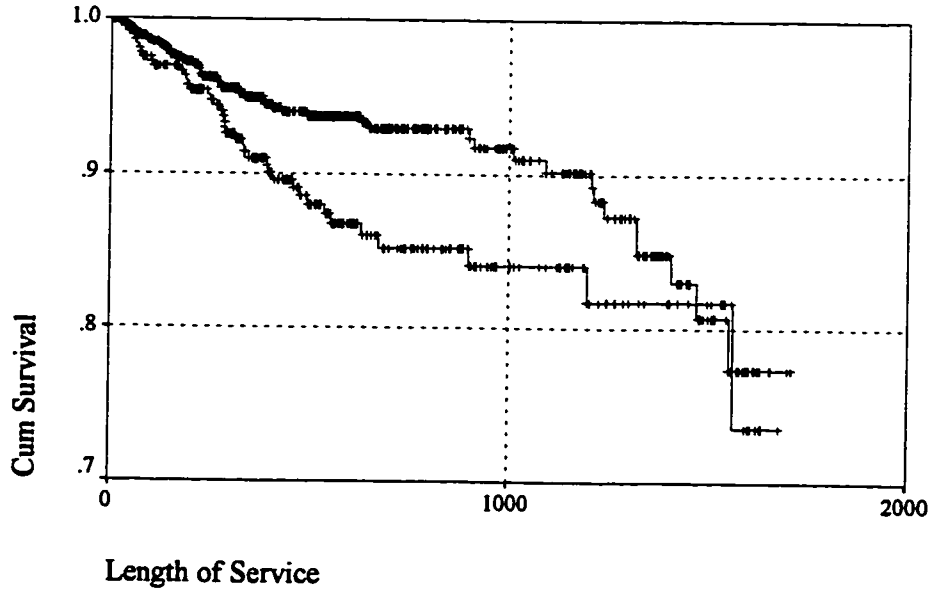
Survival Functions

HCIS = 1



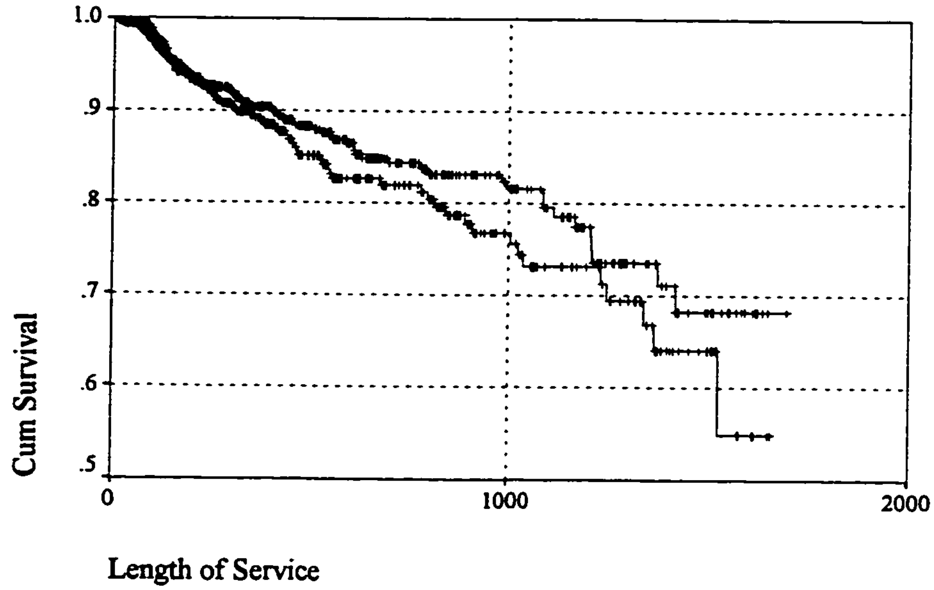
Survival Functions

HCIS = 2



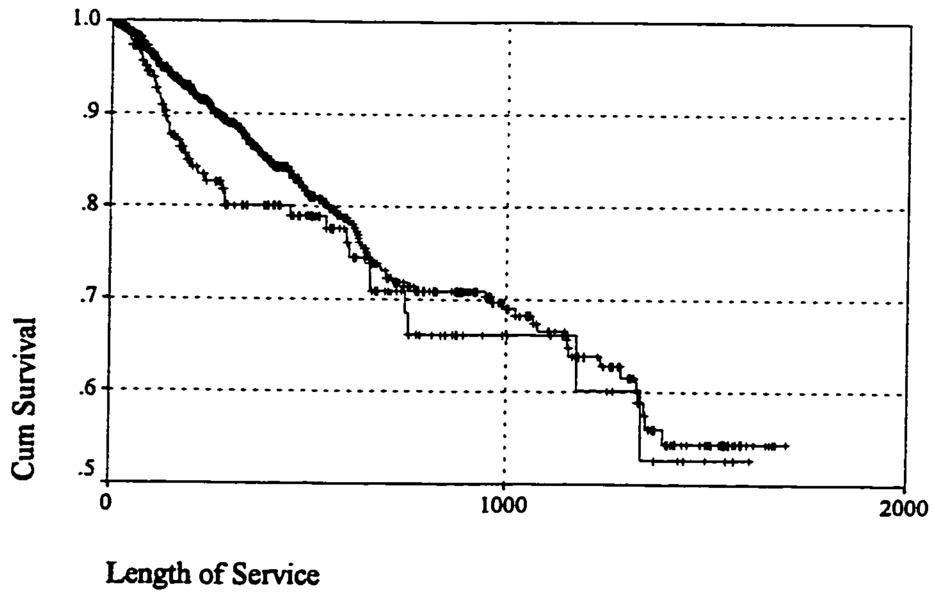
Survival Functions

HCIS = 3



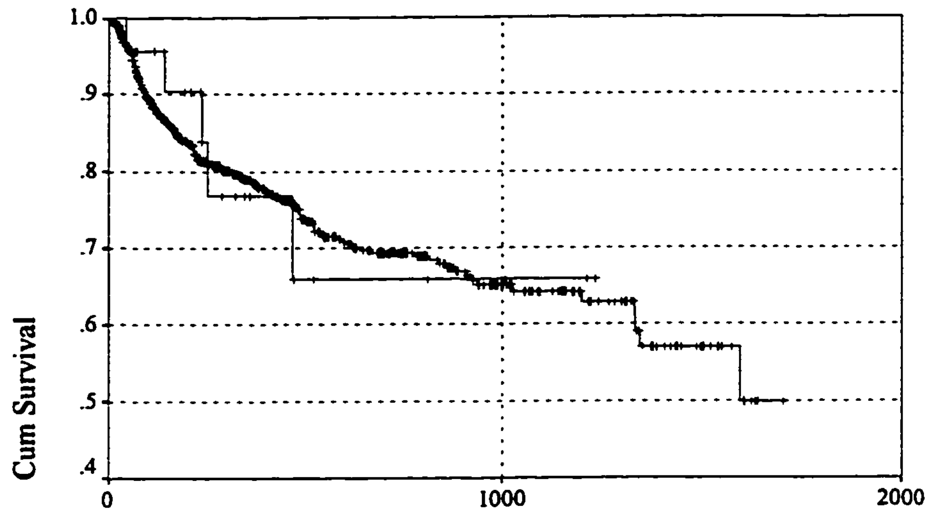
Survival Functions

HCIS = 4



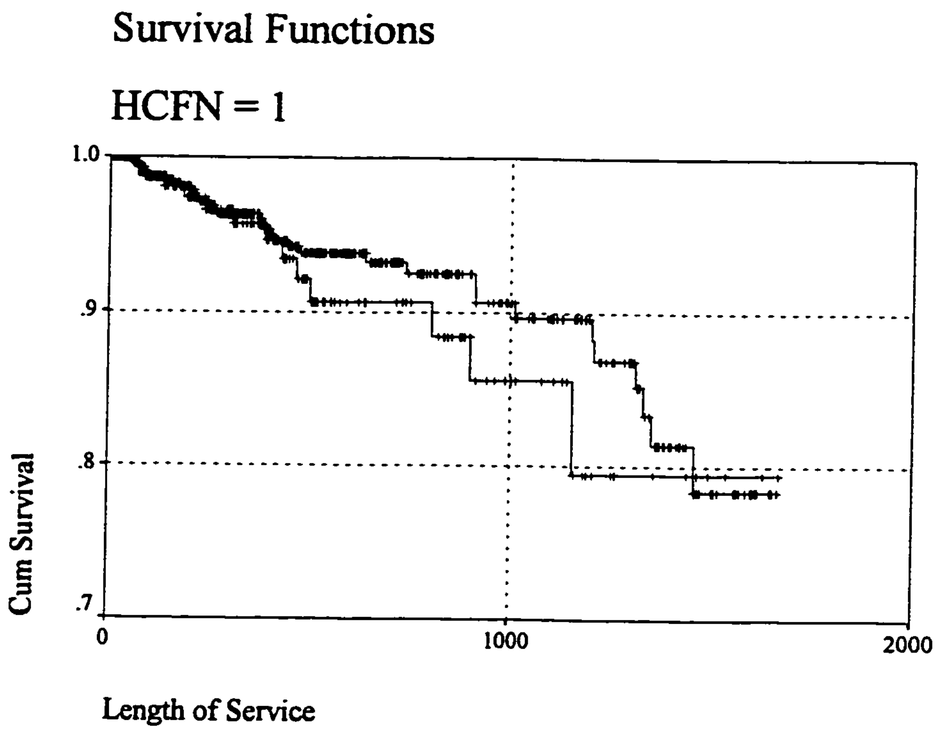
Survival Functions

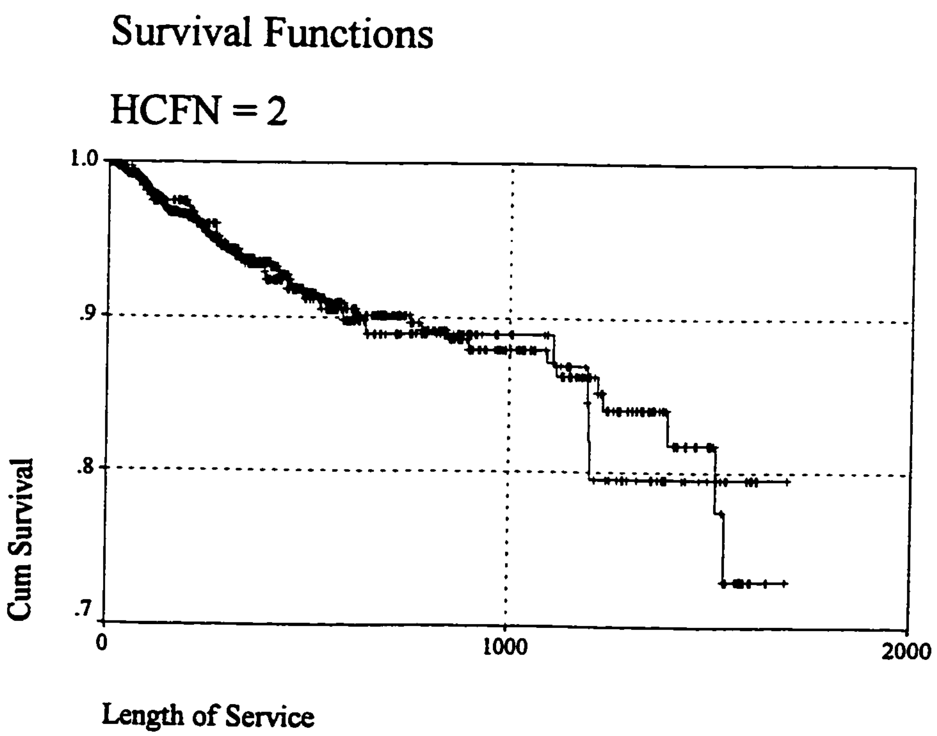
HCIS = 5



Length of Service

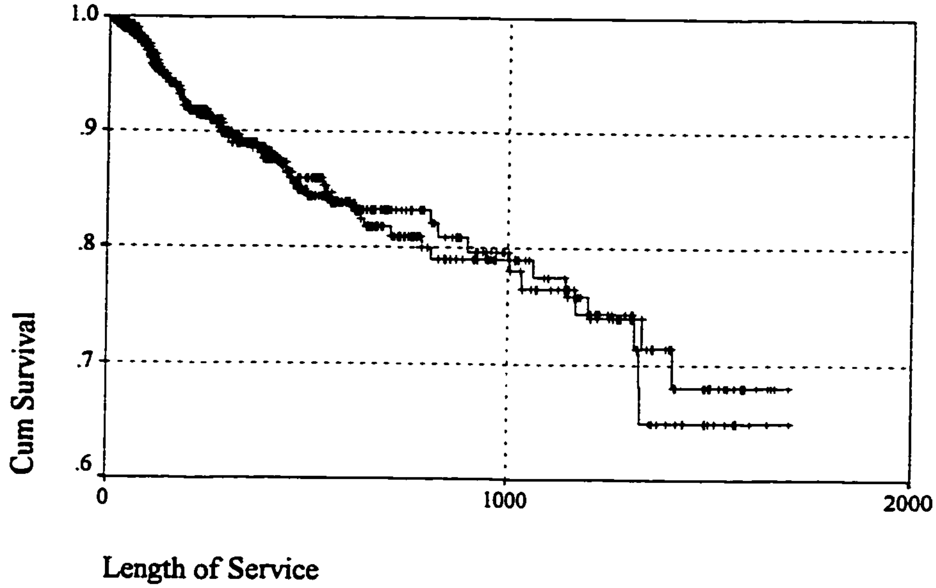
(e) Functional Need Score





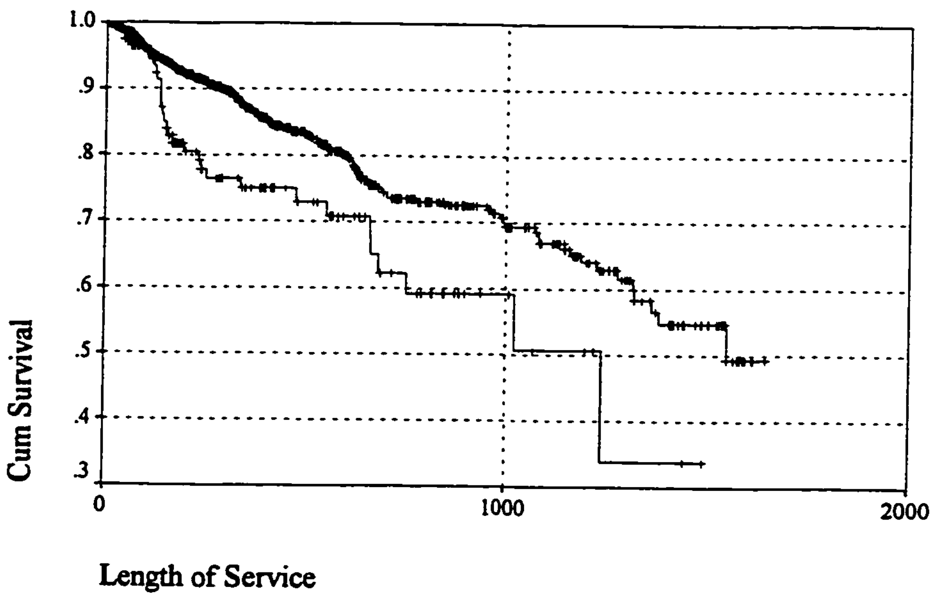
Survival Functions

HCFN = 3



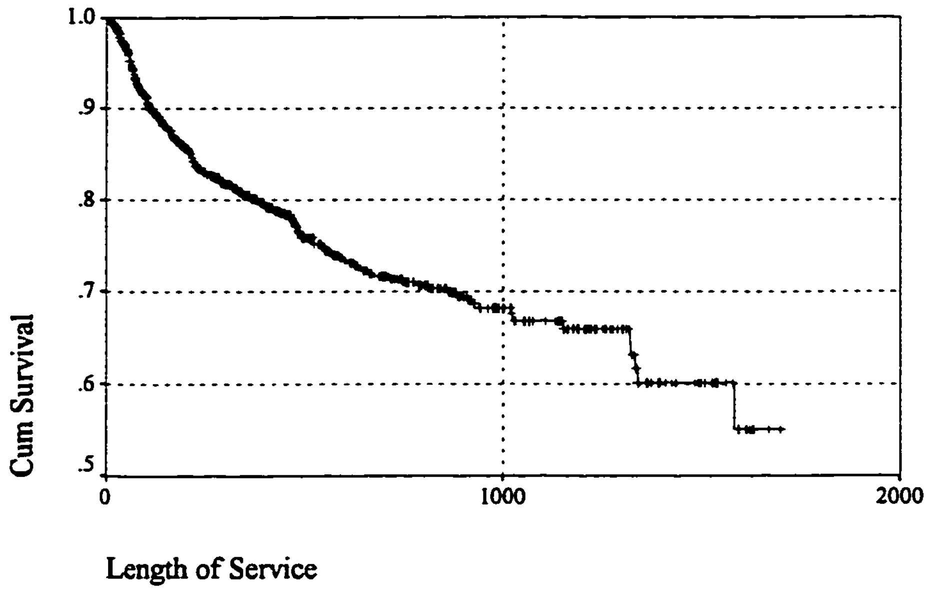
Survival Functions

HCFN = 4



Survival Functions

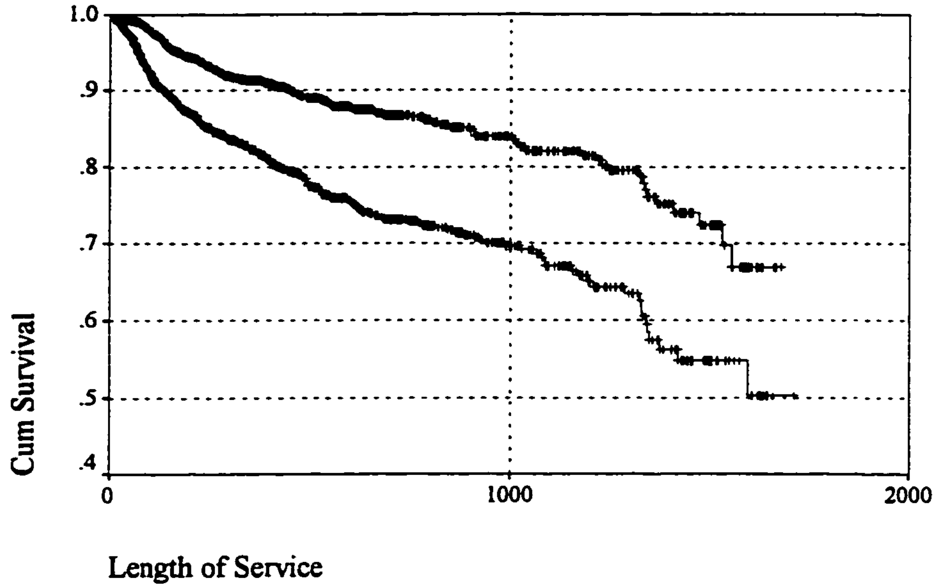
HCFN = 5



(f) Observed Informal Cost per Day

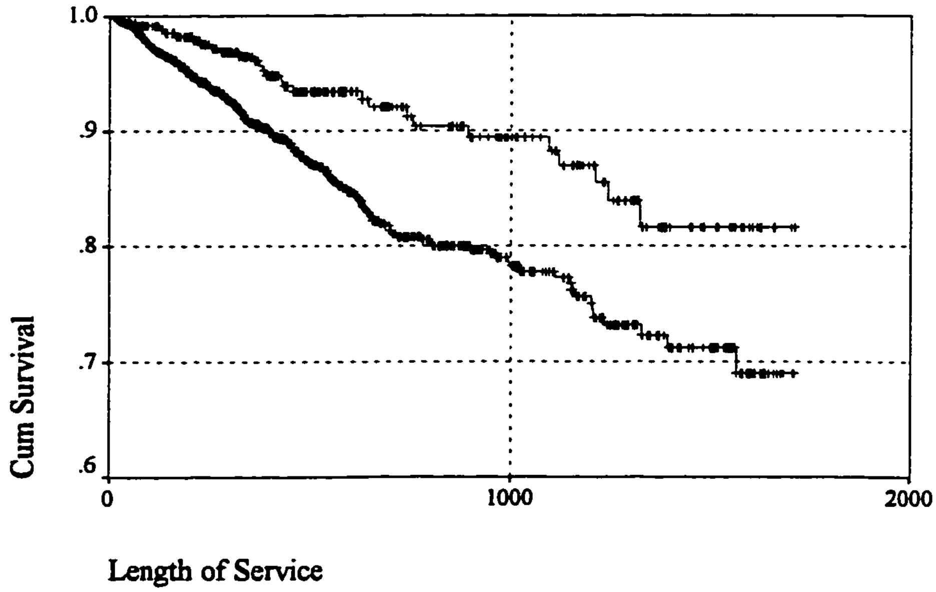
Survival Functions

DICPDAY = Informal Cost / Day < $2.50

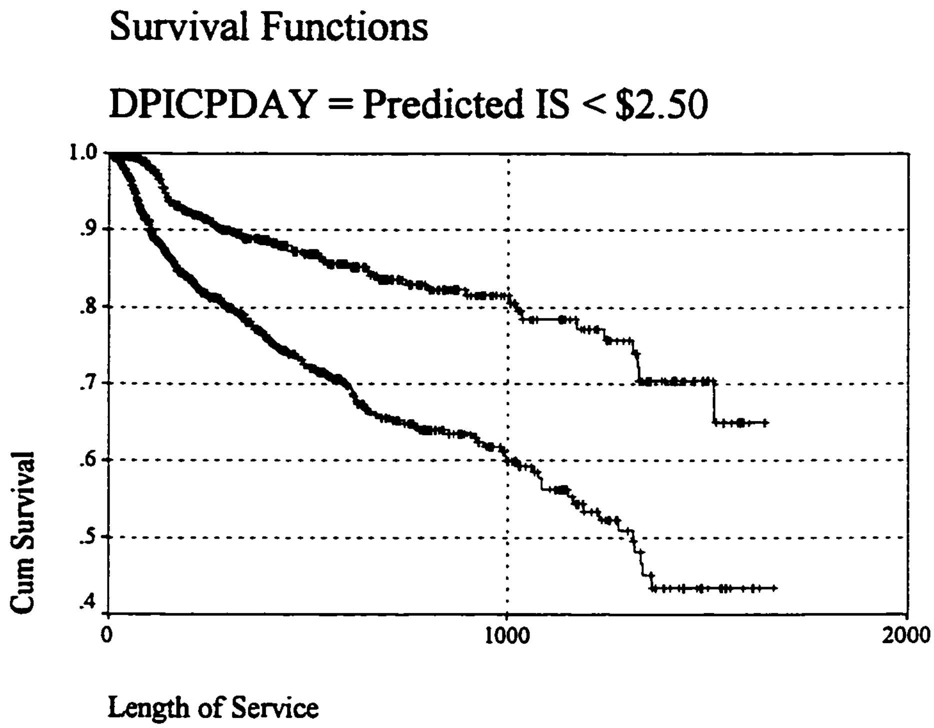


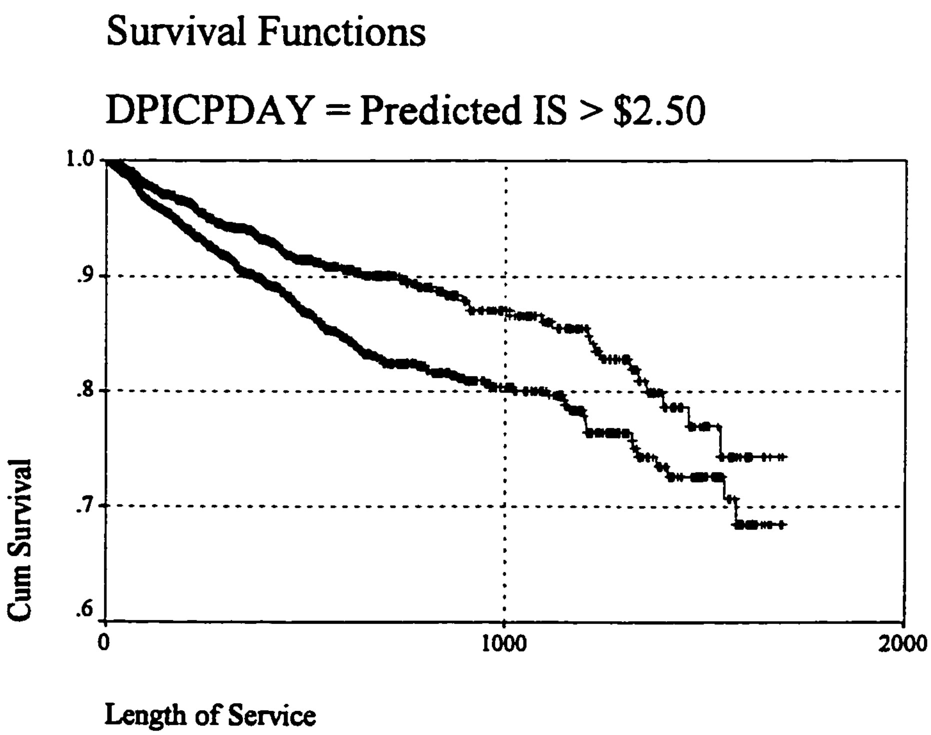
Survival Functions

DICPDAY = Informal Cost / Day >= $2.50



(g) Predicted Informal Cost per Day

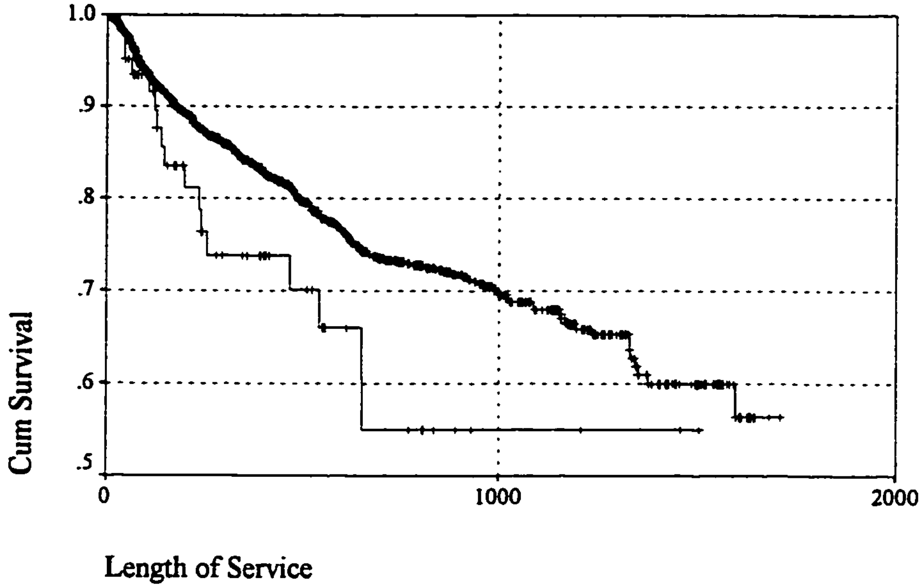




(h) Physical Function

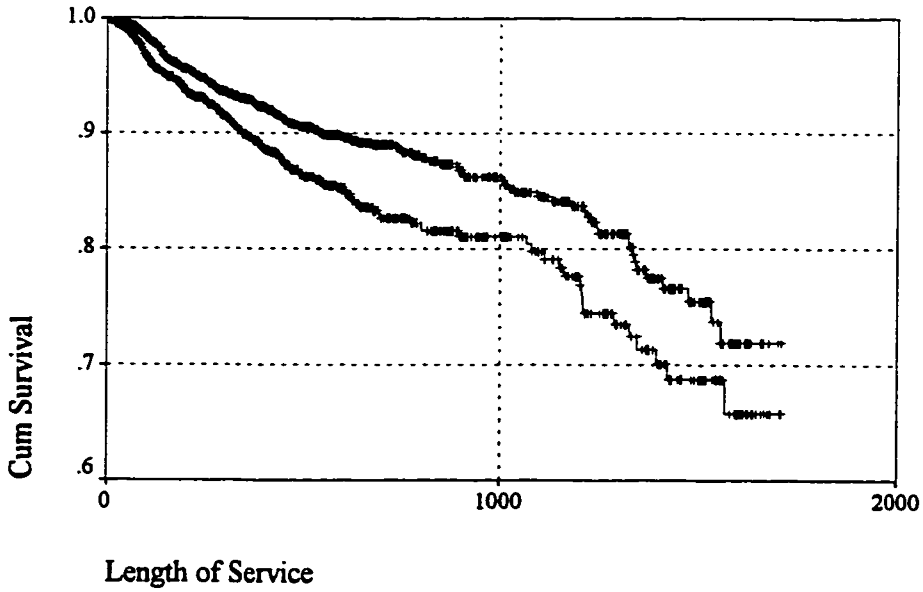
Survival Functions

DPFN = Below Average PFn



Survival Functions

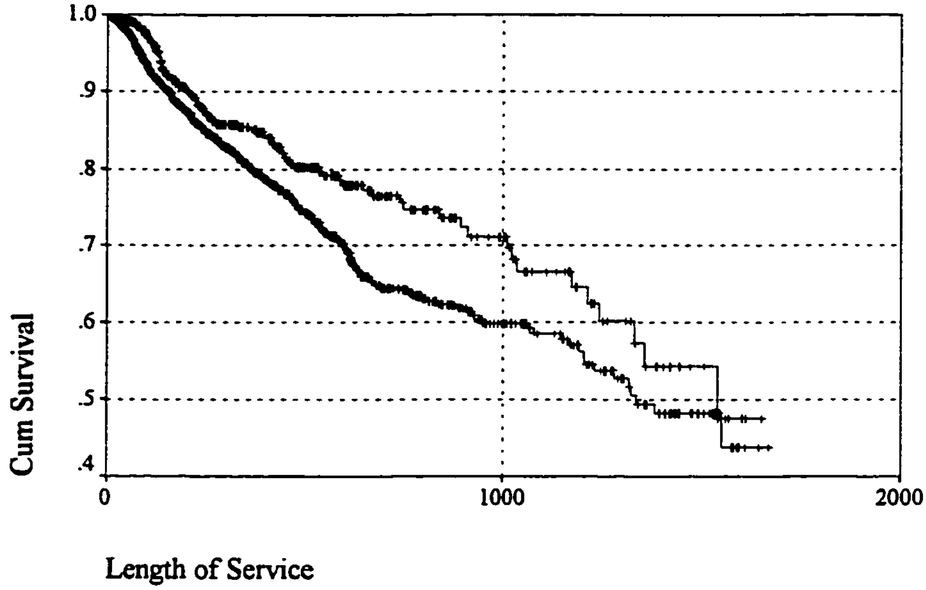
DPFN = Above Average PFn



(i) Mental Function

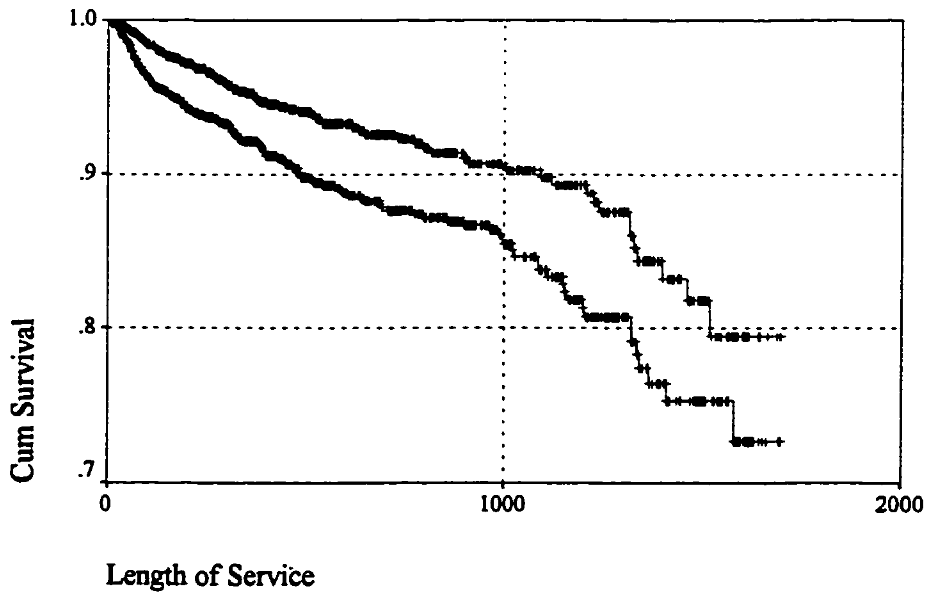
Survival Functions

DMFN = Below Average MFn



Survival Functions

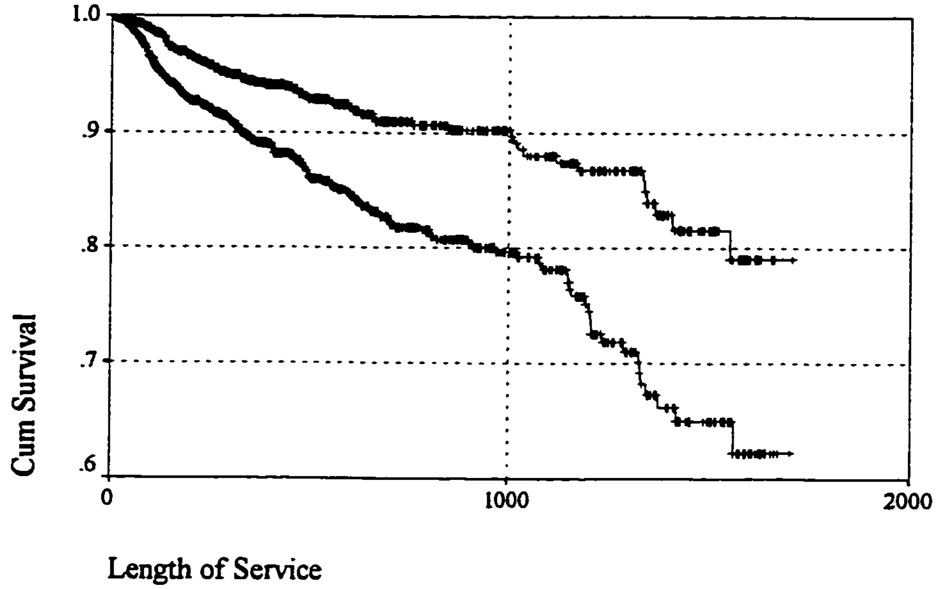
DMFN = Above Average MFn



(j) Weighted Comorbidity Index

Survival Functions

DWCI = Musculoskeletal / Injury



Survival Functions

DWCI = Cancer / Nervous System / Circulatory

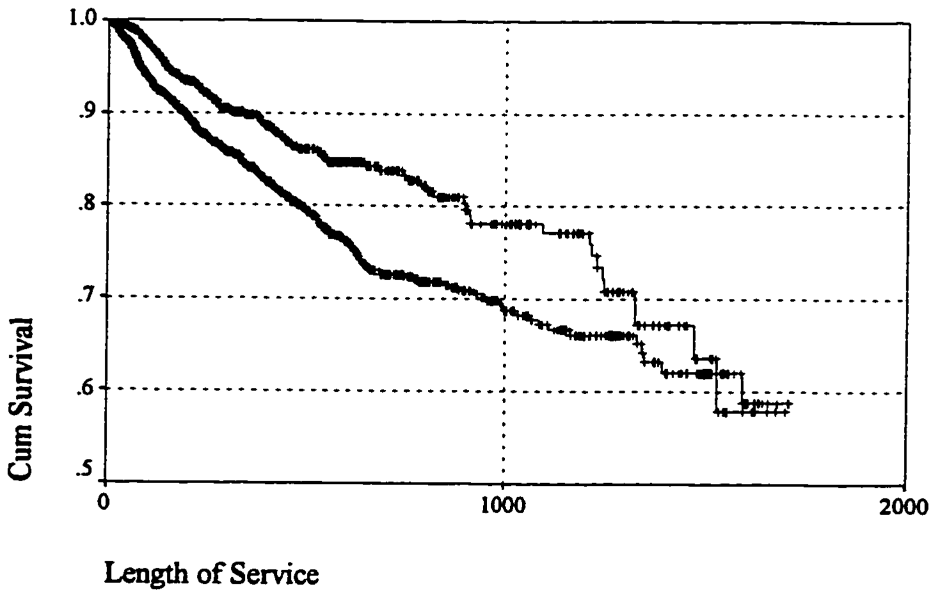
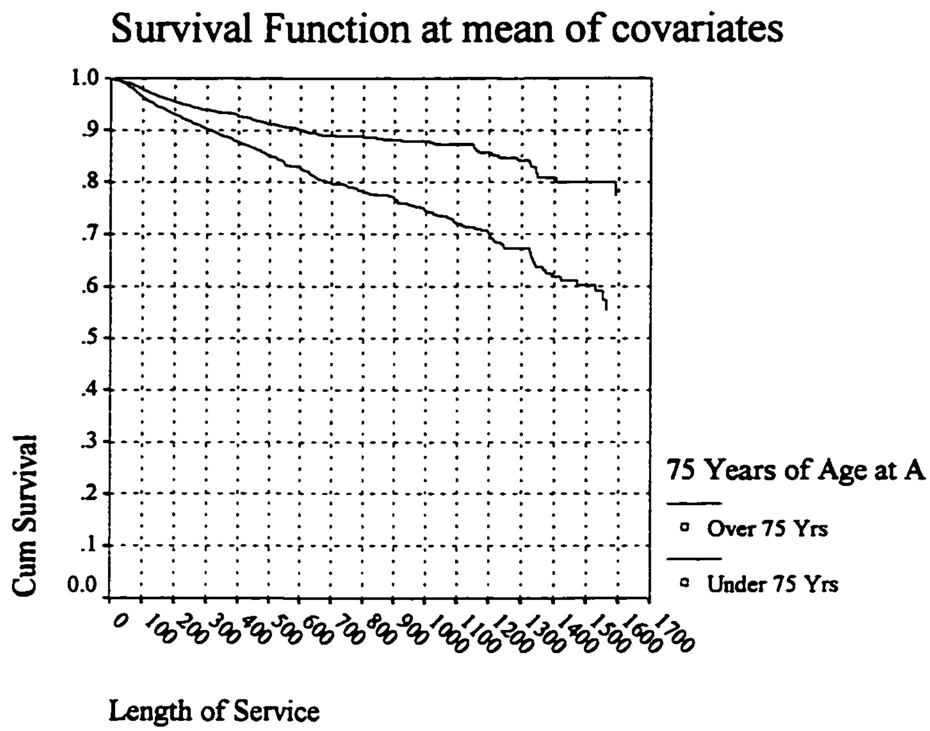
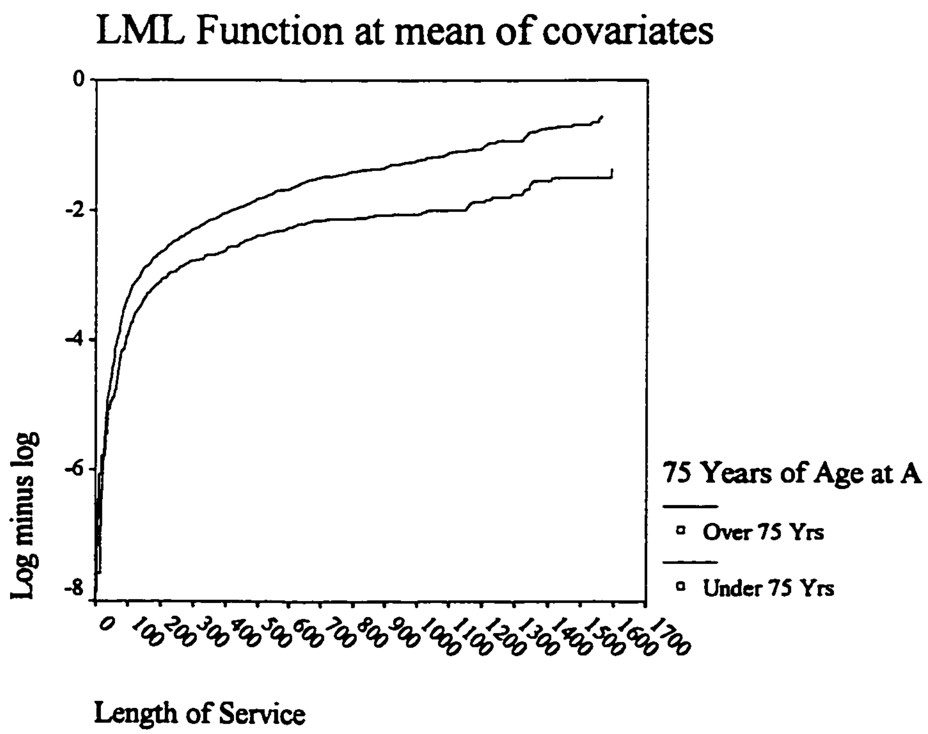


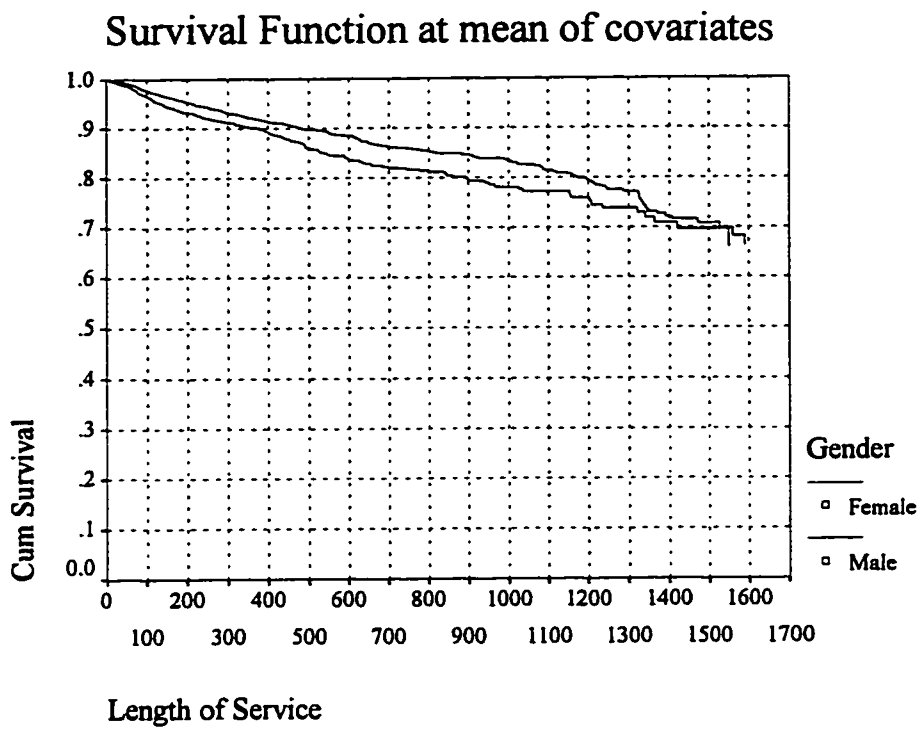
Figure 5.15: Stratified and LML Survival Curves

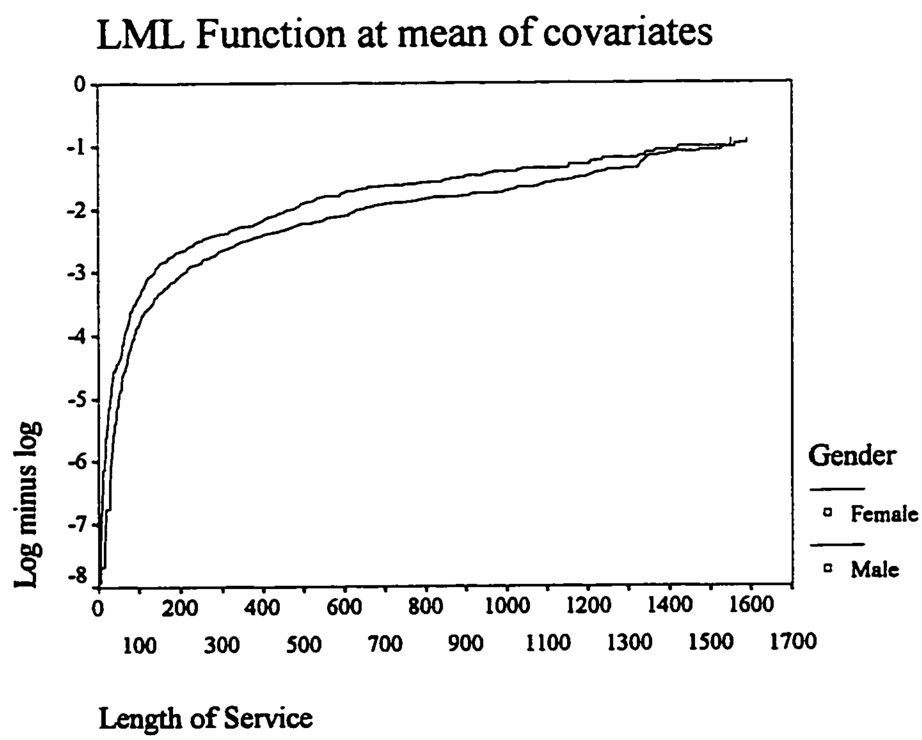
(a) Age





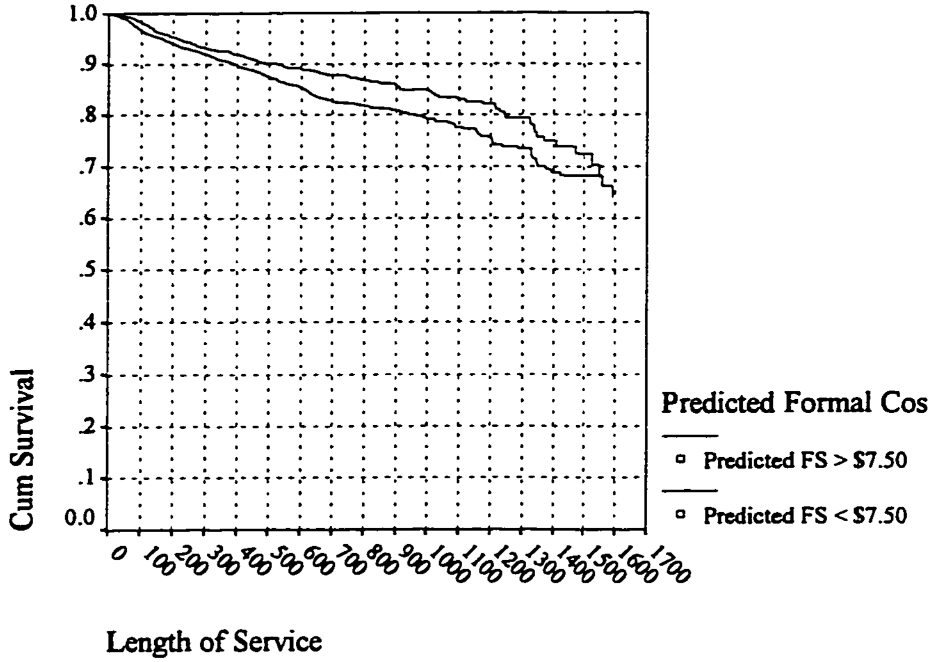
(b) Gender


Survival Function at mean of covariates
Cum Survival
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
0.0
0 200 400 600 800 1000 1200 1400 1600
100 300 500 700 900 1100 1300 1500 1700
Length of Service
Gender
Female
Male

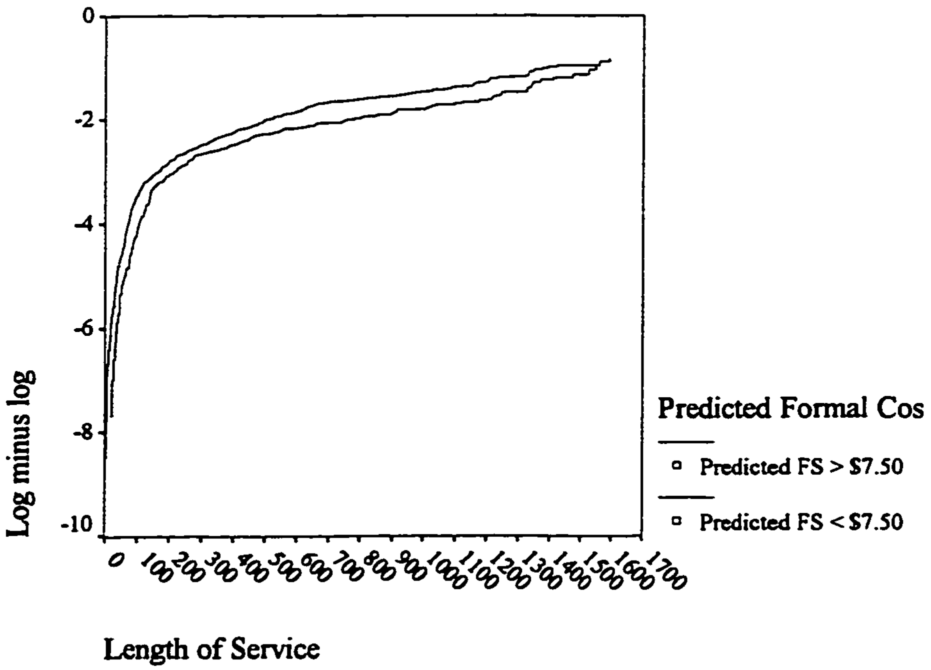

LML Function at mean of covariates
Log minus log
0
-1
-2
-3
-4
-5
-6
-7
-8
0 200 400 600 800 1000 1200 1400 1600
100 300 500 700 900 1100 1300 1500 1700
Length of Service
Gender
Female
Male

(c) Predicted Formal Cost

Survival Function at mean of covariates

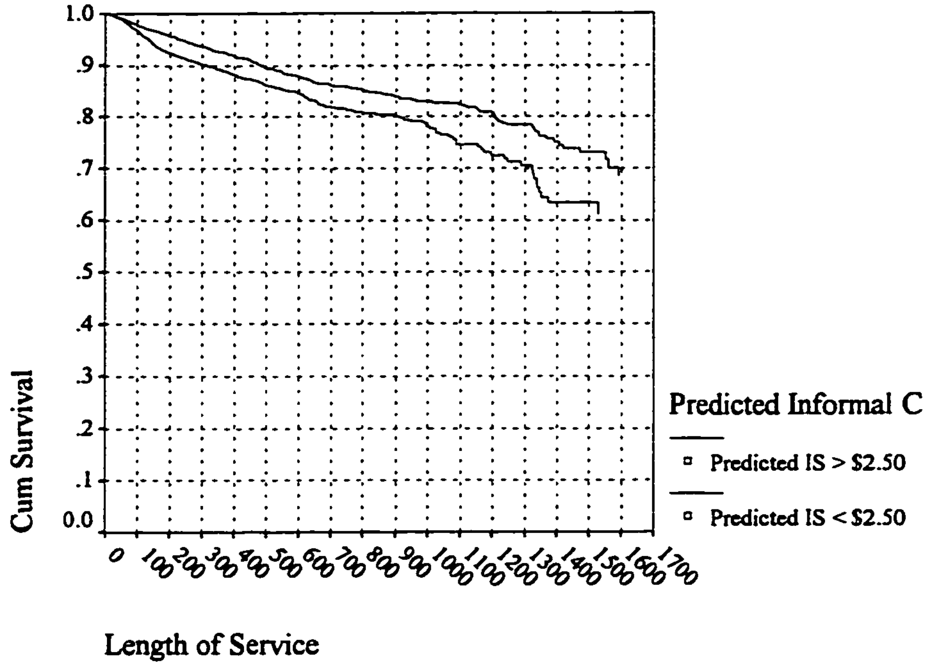


LML Function at mean of covariates

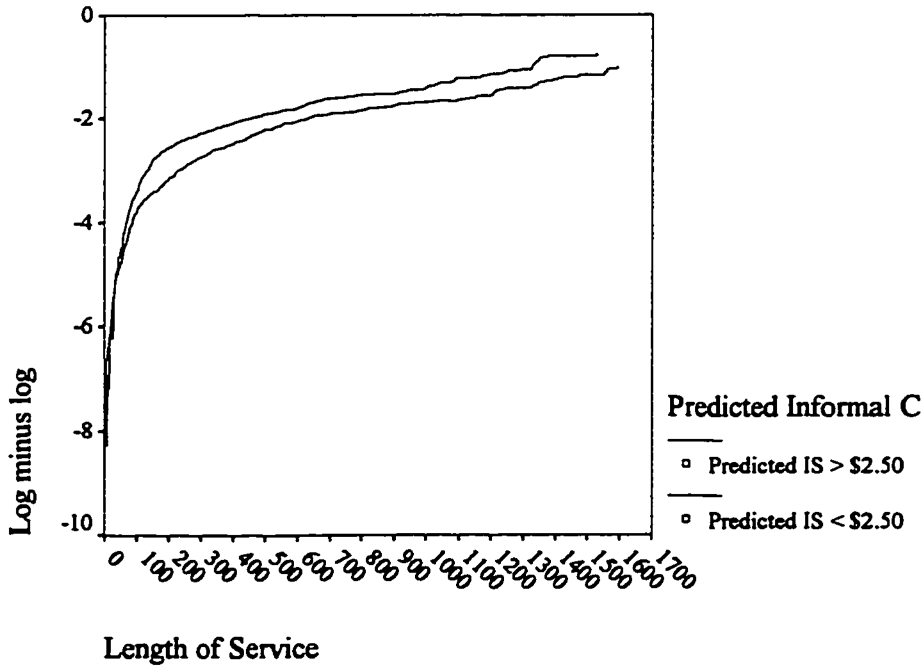


(d) Predicted Informal Cost

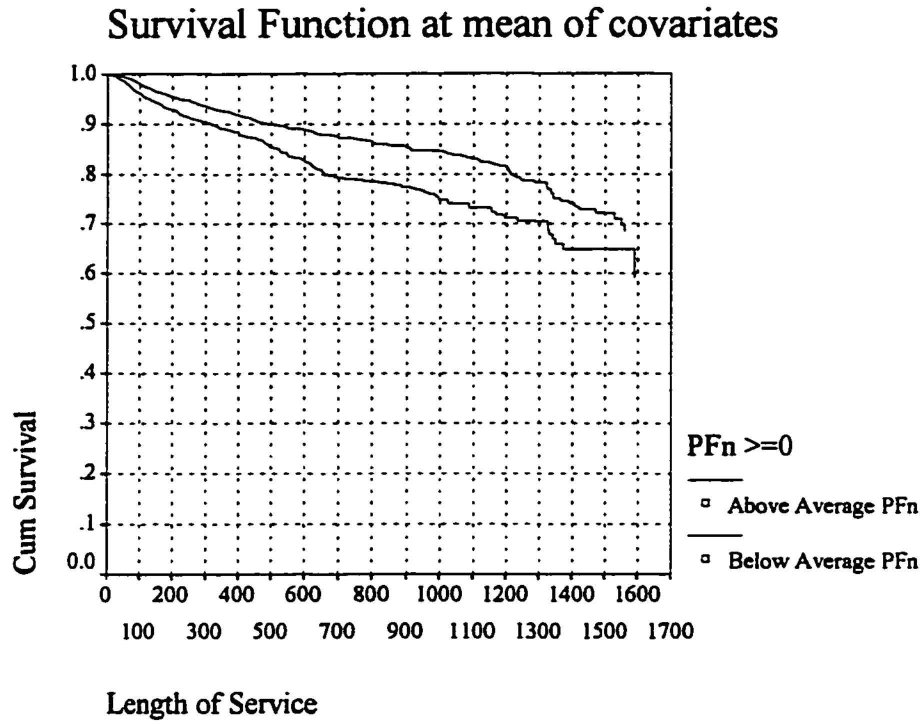
Survival Function at mean of covariates

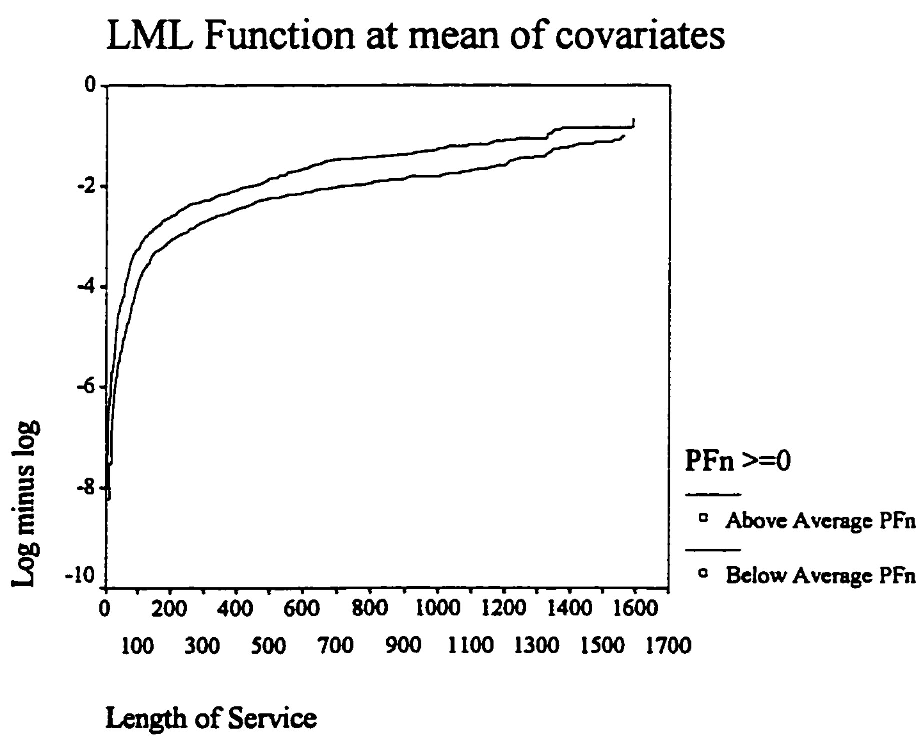


LML Function at mean of covariates



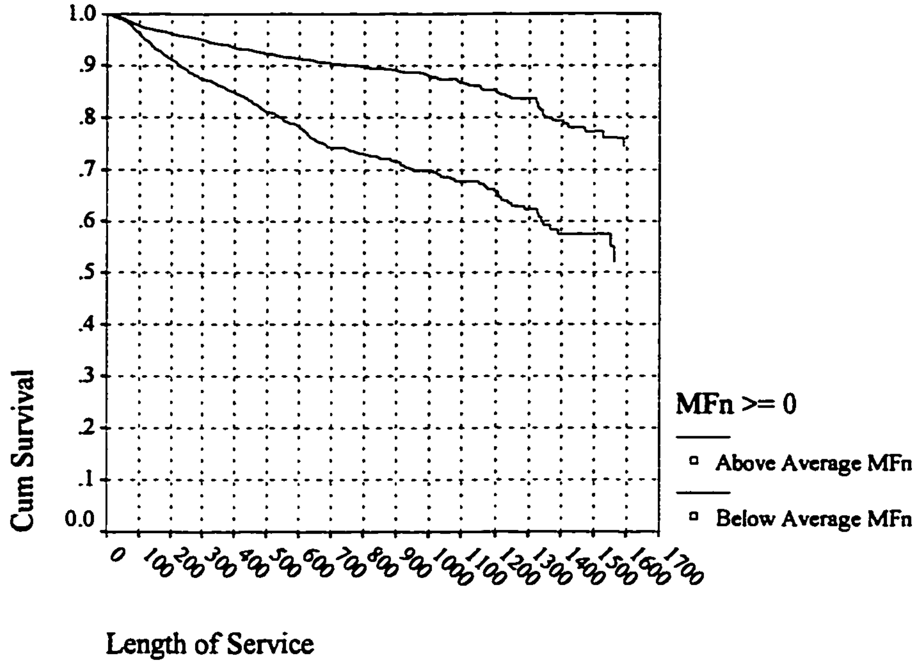
(e) Physical Function



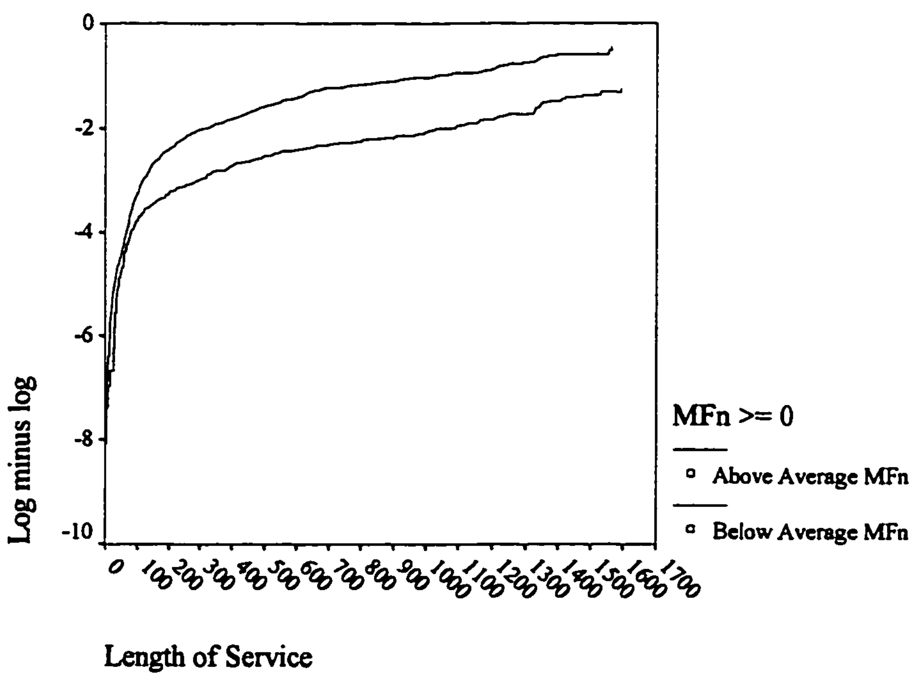


(f) Mental Function

Survival Function at mean of covariates

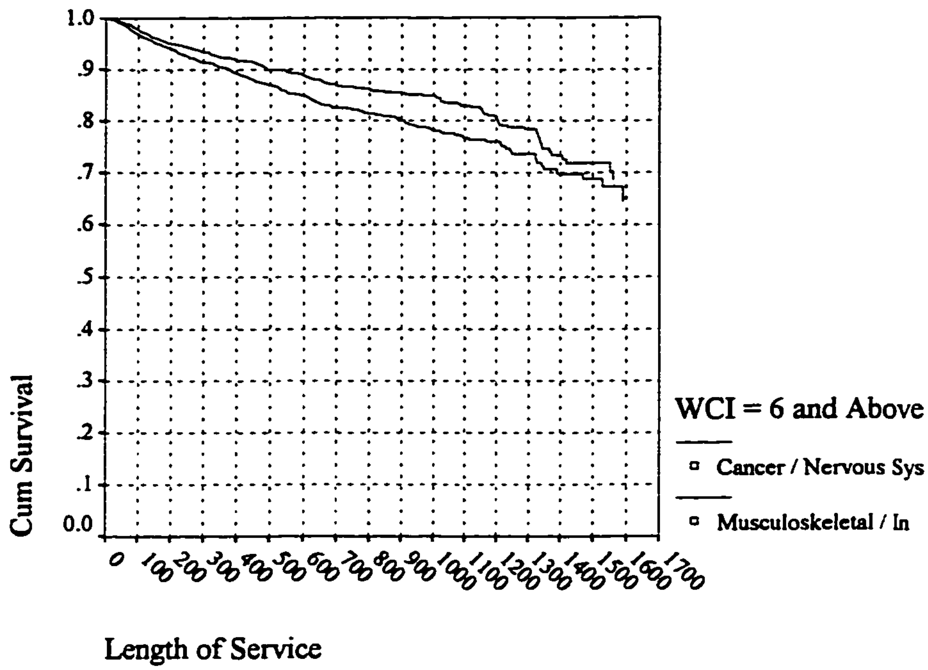


LML Function at mean of covariates



(g) Weighted Comorbidity Index

Survival Function at mean of covariates



LML Function at mean of covariates

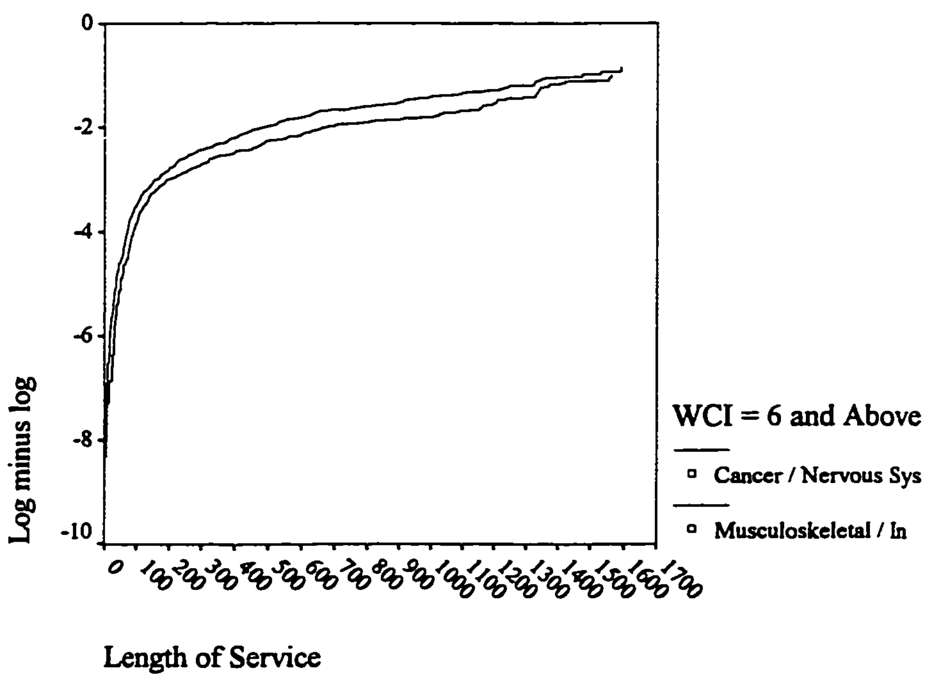
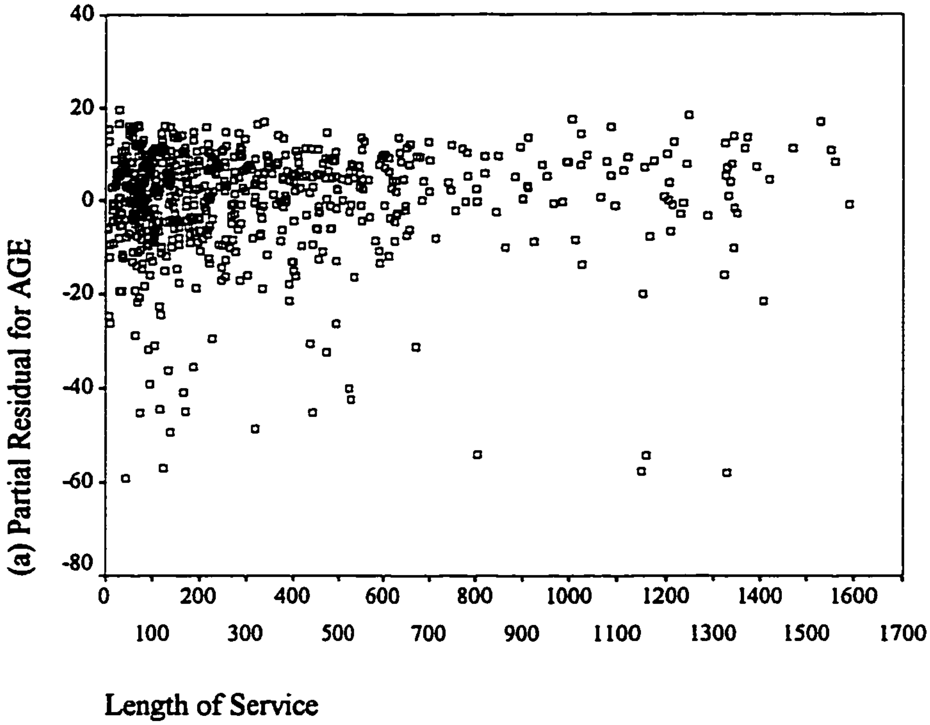
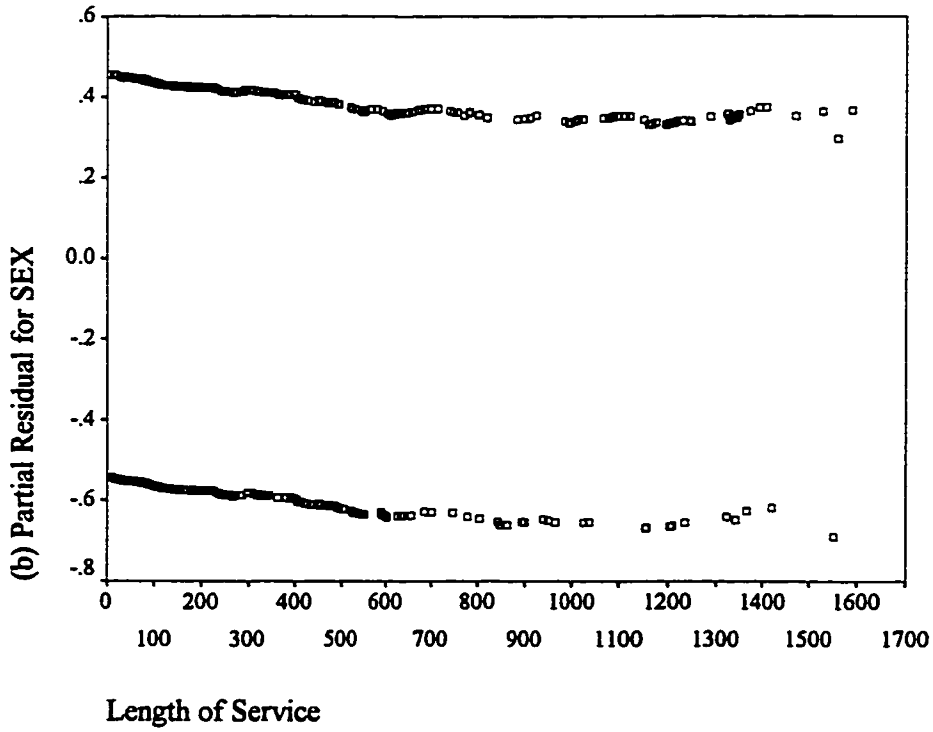
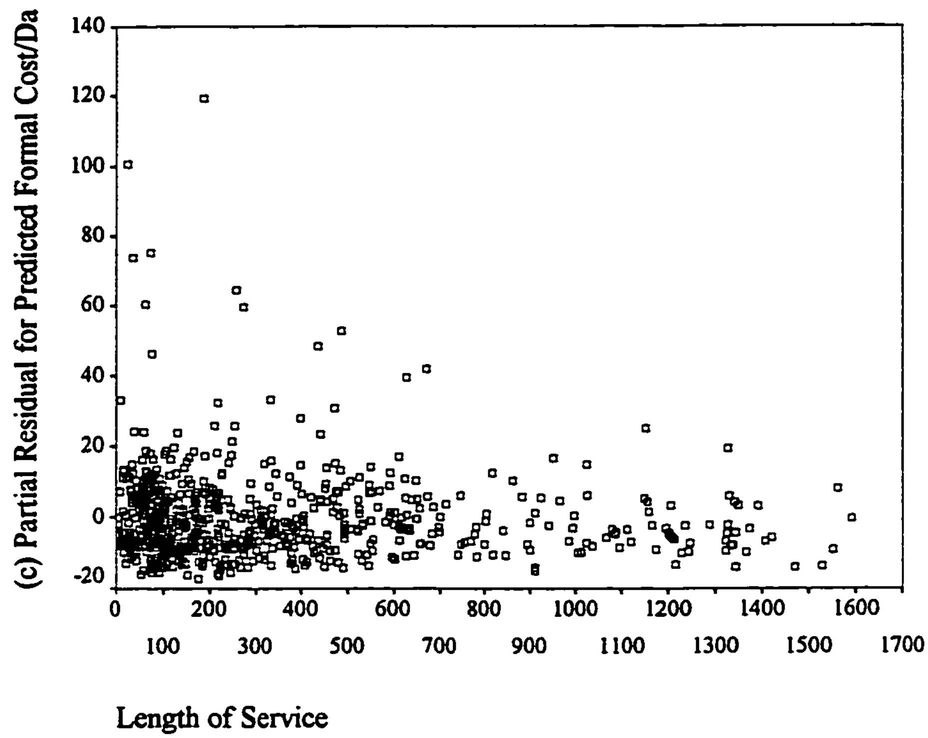
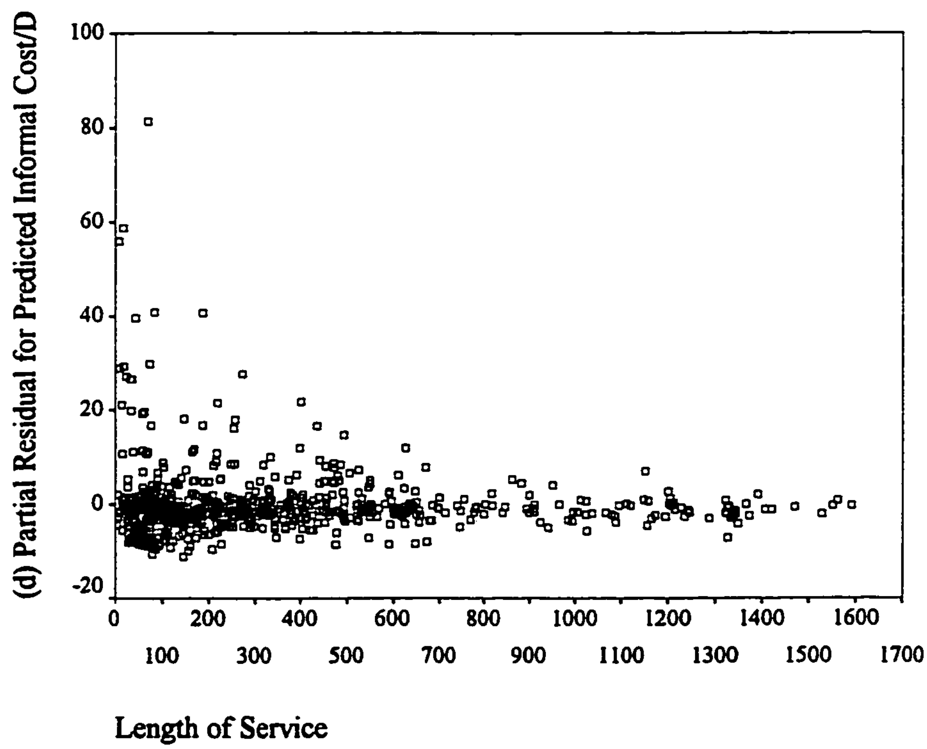


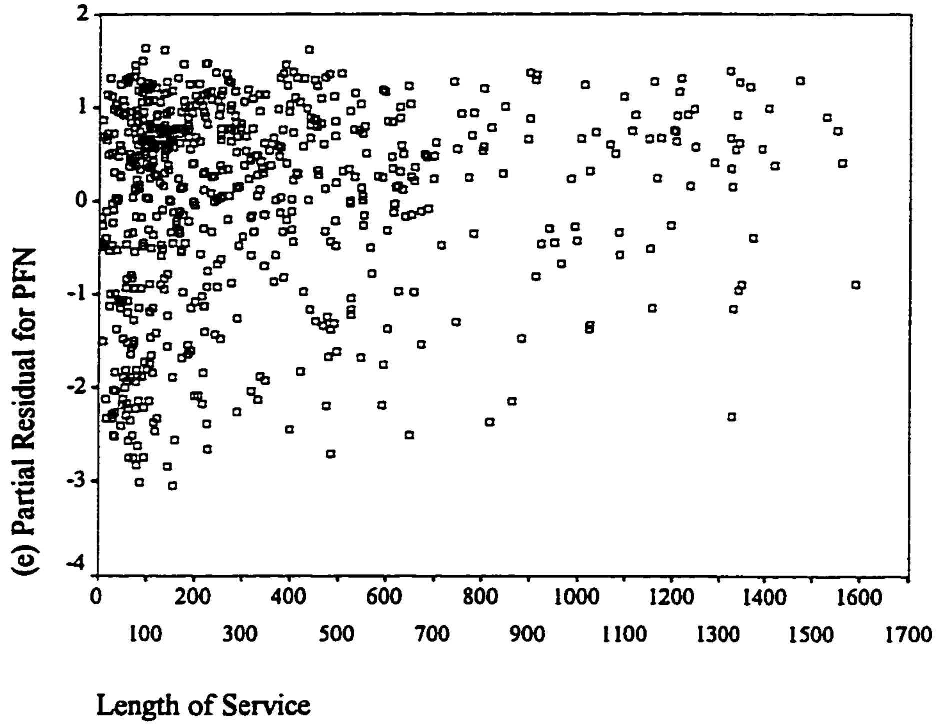
Figure 5.16: Schoenfeld's (Partial) Residuals

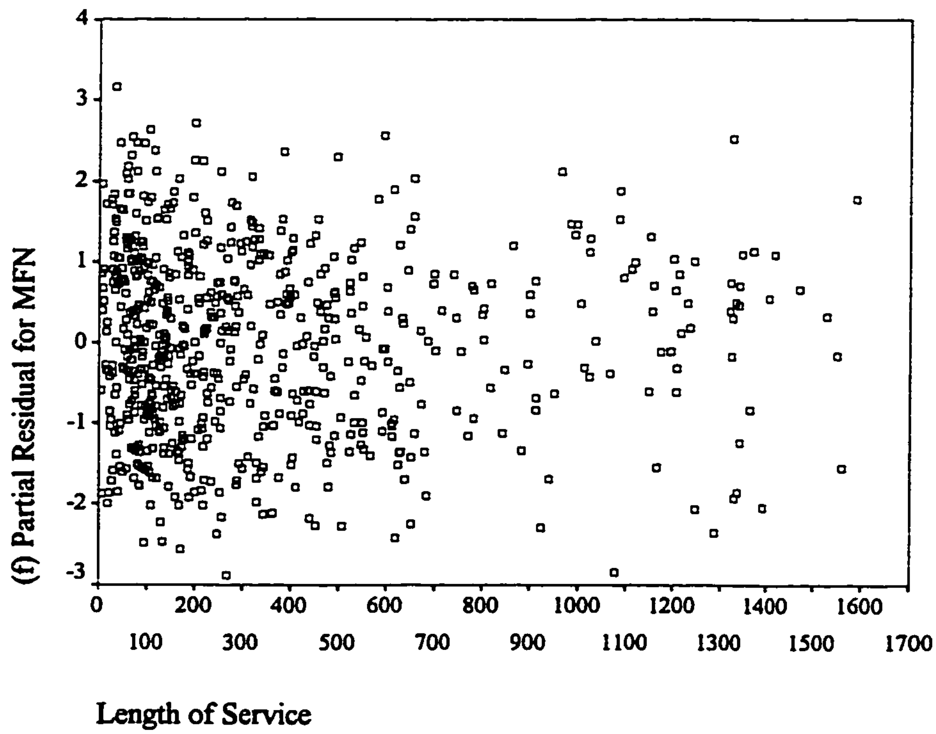





(c) Partial Residual for Predicted Formal Cost/Da
140
120
100
80
60
40
20
0
-20
0
200
400
600
800
1000
1200
1400
1600
100
300
500
700
900
1100
1300
1500
1700
Length of Service


(d) Partial Residual for Predicted Informal Cost/D
100
80
60
40
20
0
-20
0
200
400
600
800
1000
1200
1400
1600
100
300
500
700
900
1100
1300
1500
1700
Length of Service


(e) Partial Residual for PFN
2
1
0
-1
-2
-3
-4
0
200
400
600
800
1000
1200
1400
1600
100
300
500
700
900
1100
1300
1500
1700
Length of Service


(f) Partial Residual for MFN
4
3
2
1
0
-1
-2
-3
0
200
400
600
800
1000
1200
1400
1600
100
300
500
700
900
1100
1300
1500
1700
Length of Service

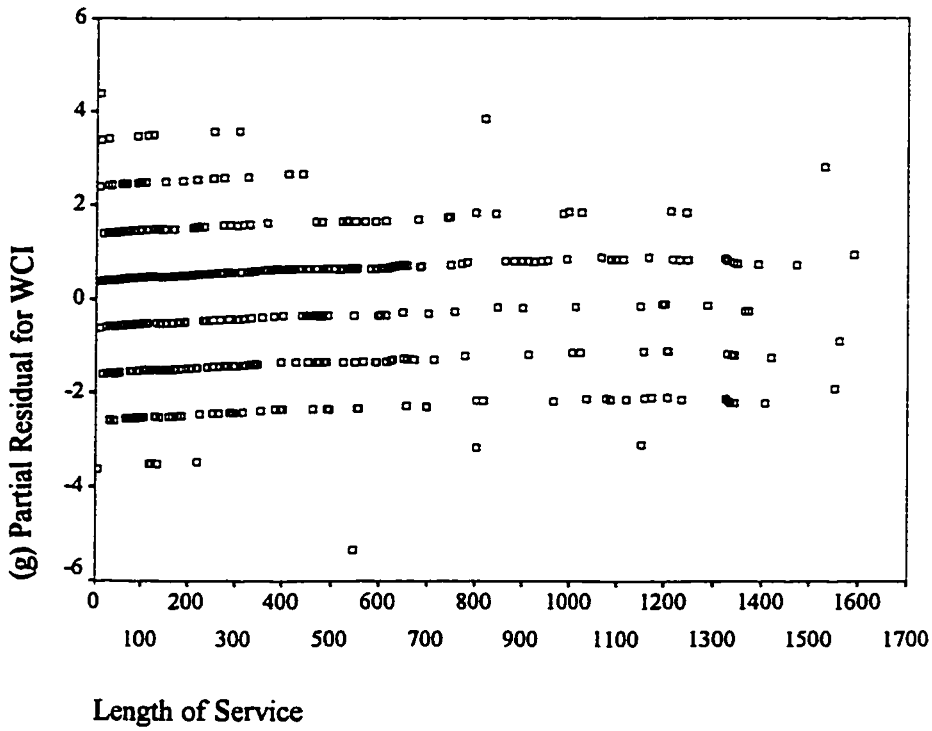
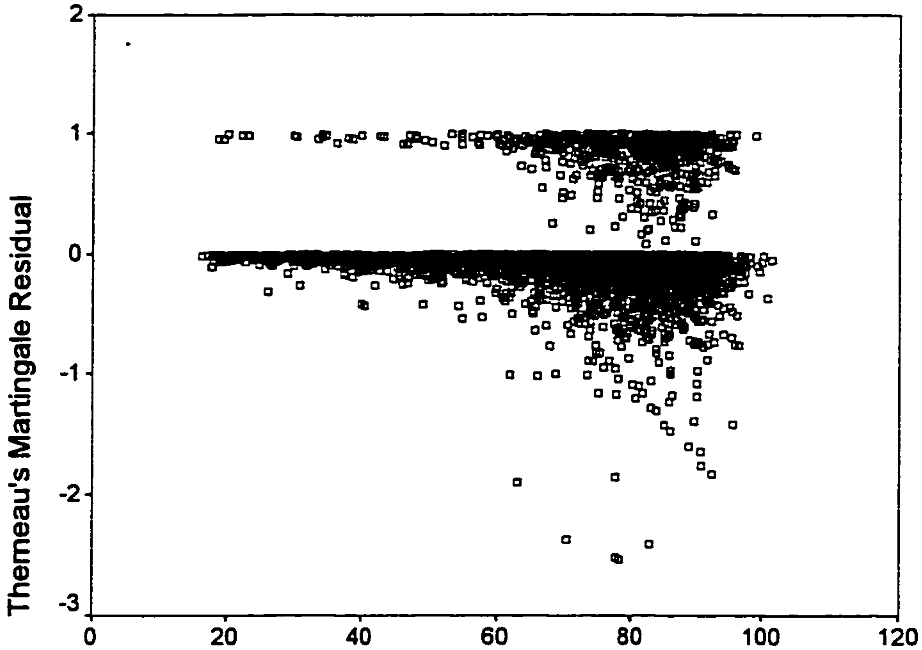
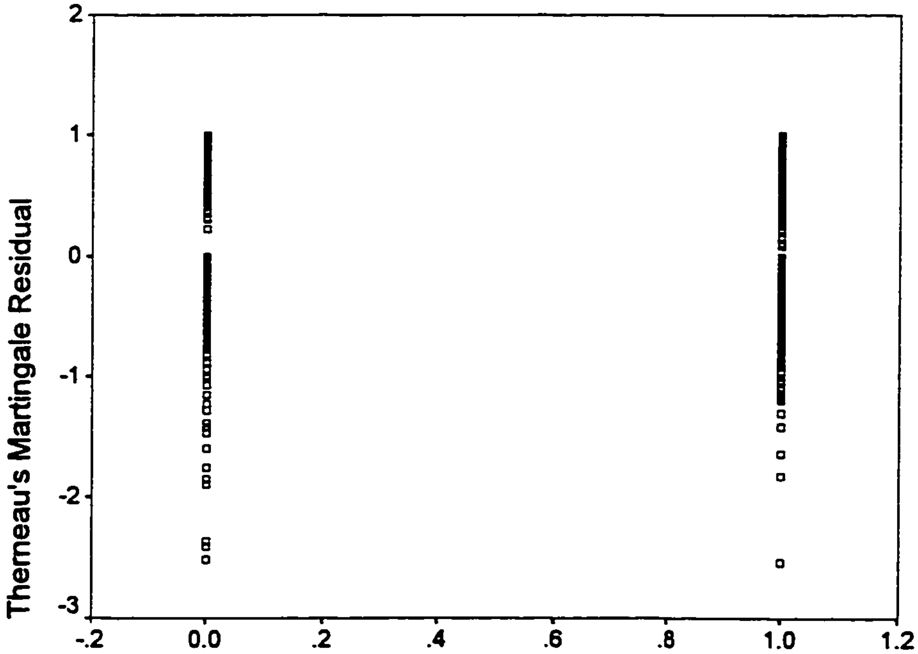

6
4
2
0
-2
-4
-6
(g) Partial Residual for WCI
0
200
400
600
800
1000
1200
1400
1600
100
300
500
700
900
1100
1300
1500
1700
Length of Service

Figure 5.17: Therneau's Martingale Residuals

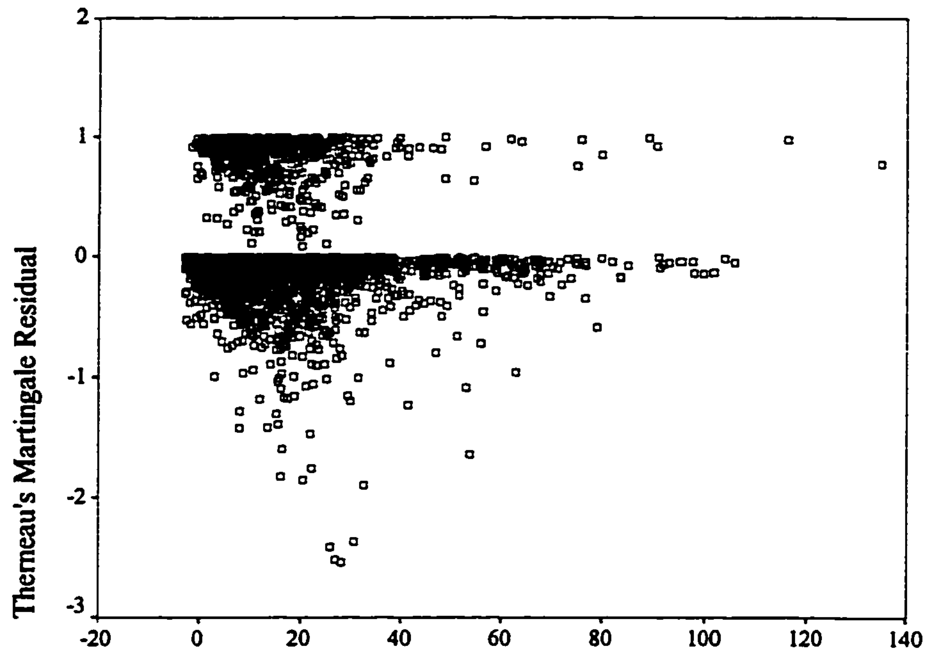


(a) Age at Admission

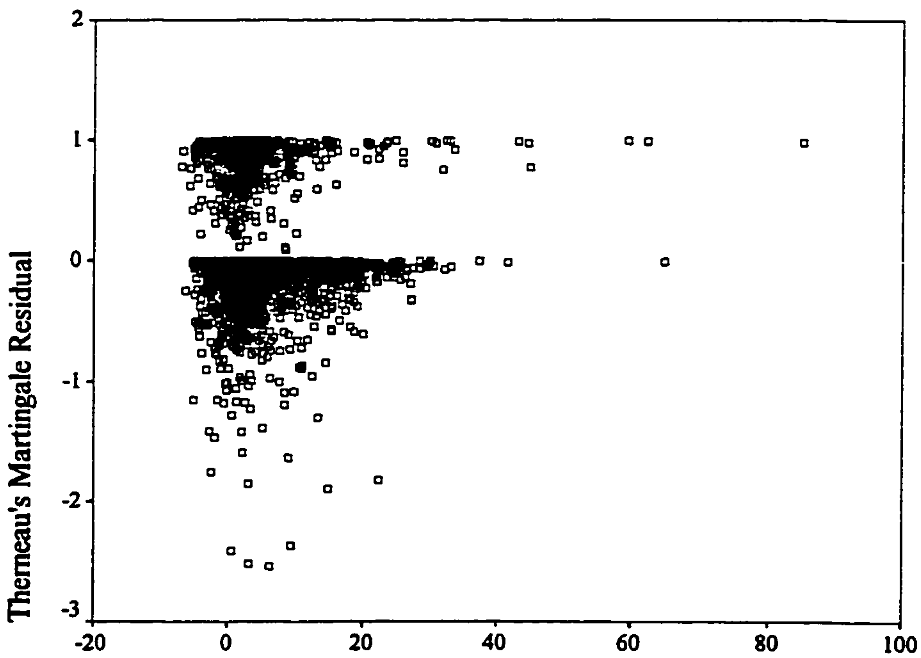


(b) Gender

Page 1

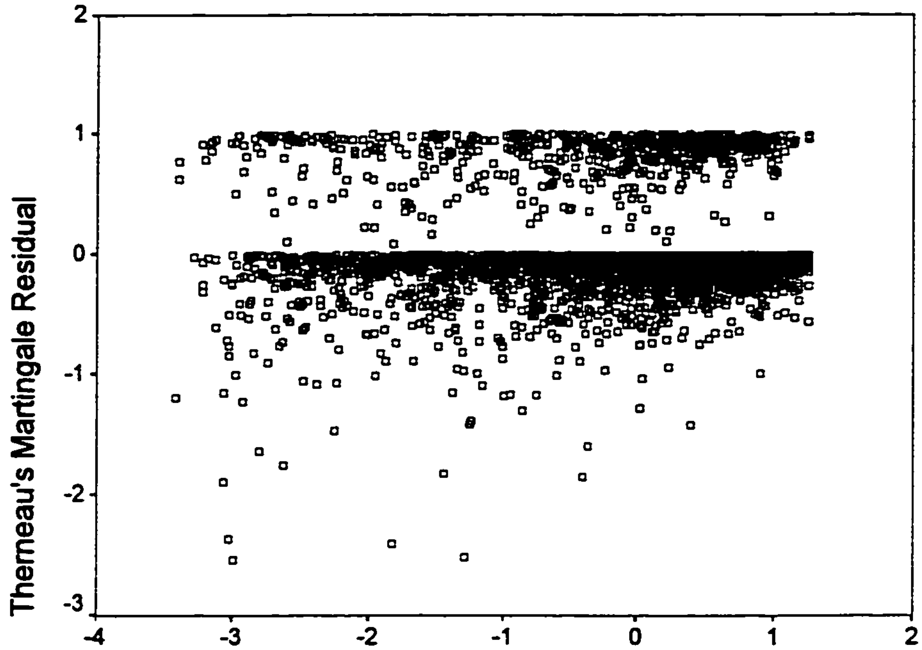


(c) Predicted Formal Service Intensity

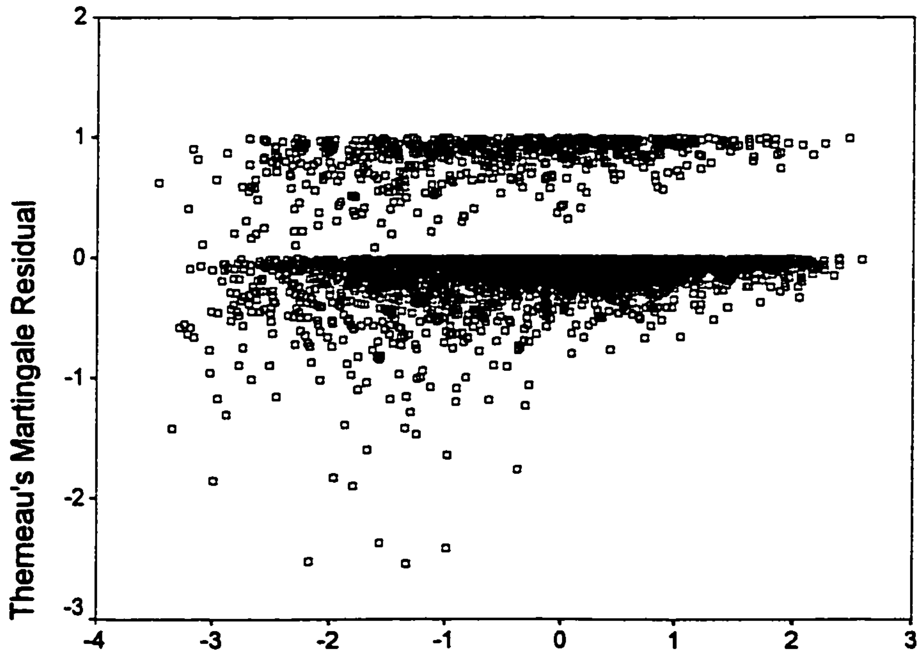


(d) Predicted Informal Service Intensity

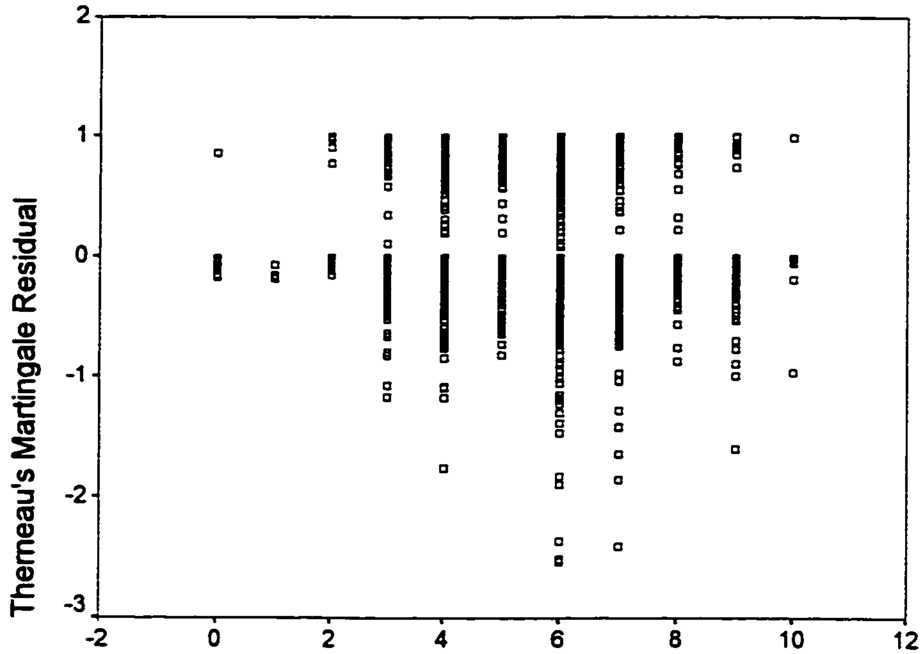
Page 2



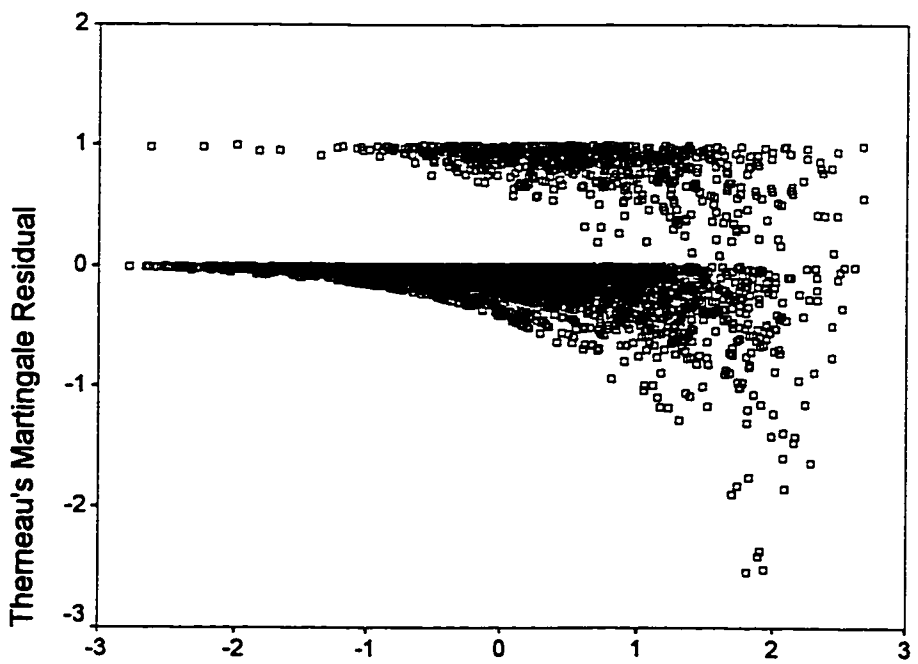
(e) Physical Function



(f) Mental Function



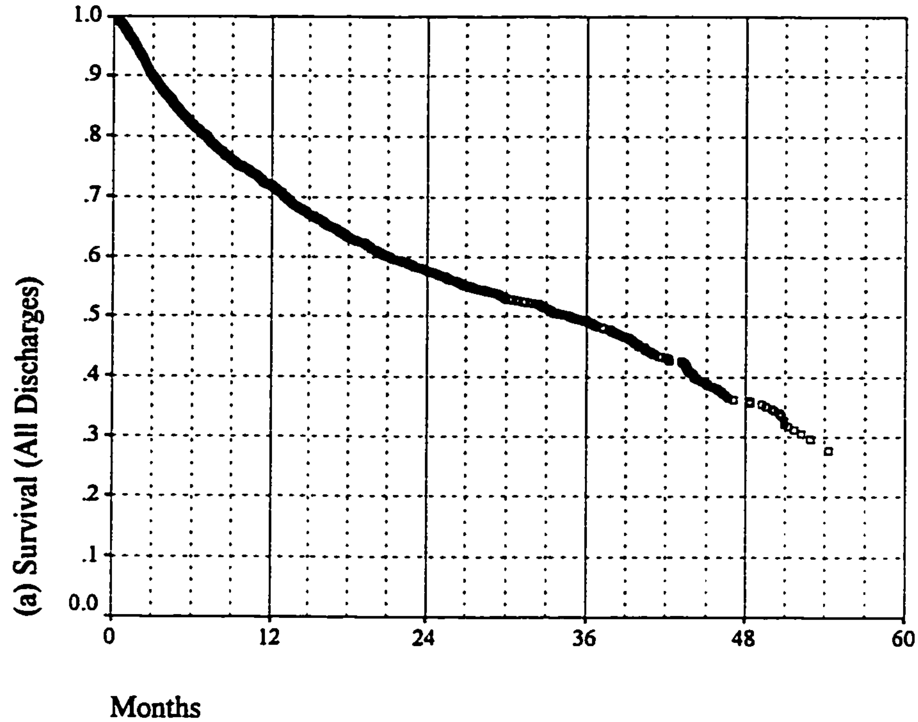
(g) Weighted Comorbidity Index

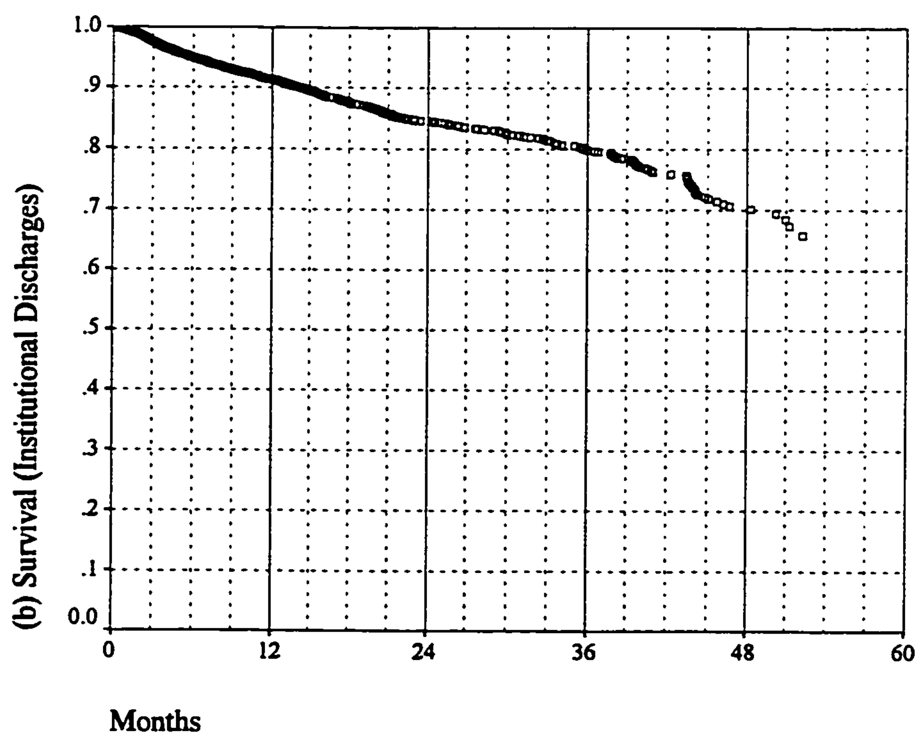


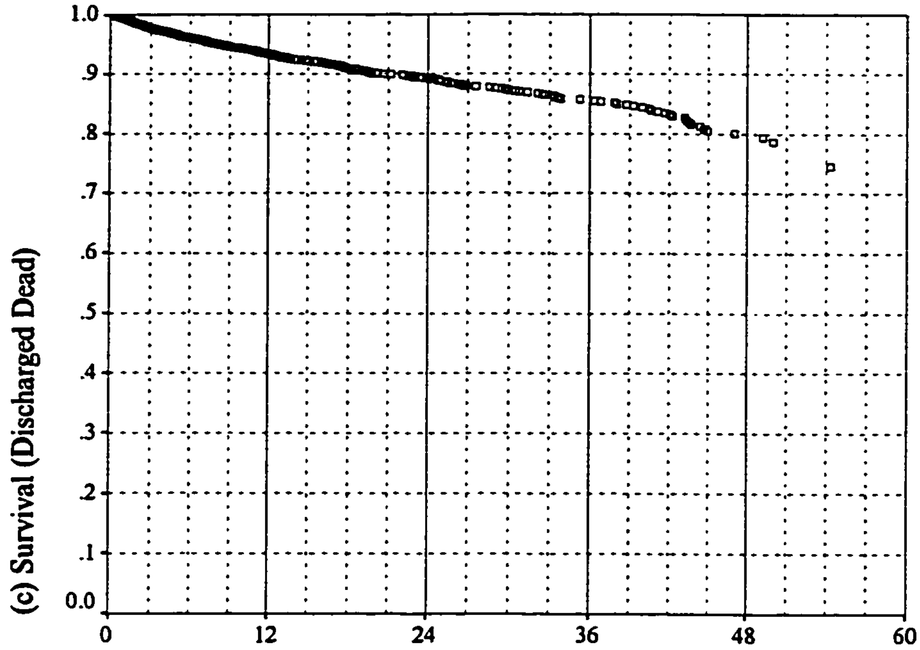
(h) X'Beta

Page 4

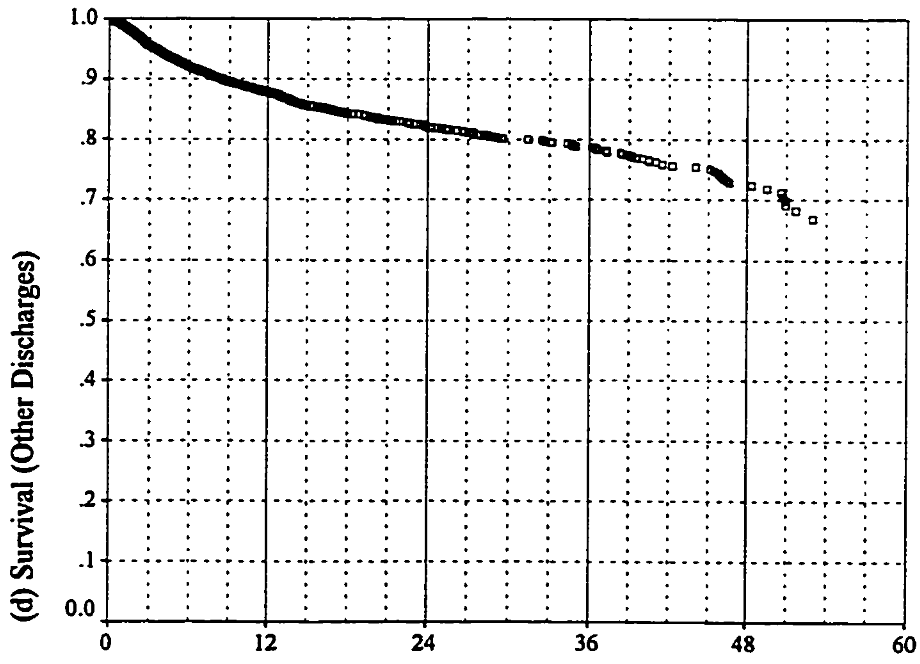
Figure 5.18: Survival Curves by Competing Hazard







Months



Months

Bibliography

Cox, D. R. and Oakes, D. (1984). *Analysis of Survival Data*, London: Chapman & Hall.

Goldman, A. I. (1992). Visualizing survival and other timed--events data". *American Statistician*, 46:13-18.

Kalbfleisch J.D. and Prentice R.L. (1980), *The Statistical Analysis of Failure Time Data*, Wiley and Sons, New York

Lin DY, Wei LJ, Ying Z. (1993) "Checking the Cox model with cumulative sums of martingale-based residuals". *Biometrika*, 80:557-572, 1993.

Schoenfeld D. "Residuals for the proportional hazards regression model". *Biometrika*, 1982, 69(1):239-241.

Schoenfeld, D. (1982). "Partial residuals for proportional hazards regression model", *Biometrika*, 69, 239-241.

Therneau, T. M., Grambsch, P. M., and Fleming, T. R. (1990). "Martingale based residuals for survival models", *Biometrika, 77*, 147-160.

VI. Summary and Conclusions

In this chapter the research results are reviewed and discussed in relation to the theoretical framework. In particular, the extent to which the results provide evidence in favour of the framework and the study hypotheses is examined. In part B, unexpected findings and their implications are summarized. Next, the health policy implications are discussed. Although this is a utilization study, the results are discussed with respect to their application in the development of policy and program management. Section D lists the limitations of the methodology and study in general. The last section of this chapter is devoted to implications for further research and practice. In essence, this section is the skeleton research agenda for the next steps that must be taken in to understand the determinants of home care.

A. Generalisation From Results to Theoretical Framework

The determination of utilization of home care services in an observational or non-experimental context requires specification of the confounding variables and relationships between them. The Andersen model is a framework that specifies the relationships between variables based on underlying behavioural concepts. In summary, this model provides some evidence in favour of the theory and accompanying framework.

1. Importance of Main Covariates and Model Specification

Model specification is the process of determining which covariates are included in the model and the relationship between those variables. The K-M univariate results demonstrate that all the variables are significant determinants of home care utilization. The only exception to expectations was that increasing levels of formal service intensity were associated with earlier discharge from the home care program.

The K-M bivariate results provide some evidence for the theoretical framework given the high number of insignificant results. As discussed in the previous chapter, a non-significant result in the stratified, bivariate setting indicates support in that formal service intensity is at least not of the wrong sign. In addition for three subgroups, the sign did reverse completely. It is not surprising that these three variable/categories were casemix and need variables.

Finally, in the multivariate setting, other than two predisposing variables, the model parameters were significant and of the correct sign. An exception occurs when defining an event as other discharges. Since other discharges encompass disparate reasons for utilization of home care resources, its evidence is not damaging to the framework. The only notable evidence against the framework is the insignificance of informal service intensity. However, observed informal support (HCIS) is significant and of the right sign for home care utilization as an alternative to death or institutionalization. Hence, the problem with the IS variable is the method of calculation. This will be discussed further in the limitations section.

In the OLS context, omission of relevant explanatory variables results in bias. It is not known to what extent this phenomenon plays a role in the reversal of signs in the Cox model.

2. Unimportance of System and Environmental Covariates

Results from the two auxiliary regressions did not support the inclusion of system and environmental covariates. With the exception of living with a spouse and/or others, all eight variables were insignificant.

The lack of evidence supporting the health care system variables (nursing home market and program budgeting) could be a result of the manner in which the variables were constructed and model specification. Construction of these data required specification of weights. These weights were used to specify whether all months within an episode were weighted equally (arithmetic weighting) or whether the last months were of more significance (geometric weights). It was hypothesised that the program financing and management exerted its influence throughout the episode of home care consumption. Perhaps the influence is determined at the time of admission. For example, the services may be 'set' at the beginning and constrained by the availability of resources at that time. Constraints later in the episode are then borne by the new admissions. Alternatively, the lack of evidence in this regard may signify excess supply. That is to say that utilization is primarily driven by demand and that program finances are increased to match demand.

It was hypothesized that the ratio of urgent to total admissions and the institutional waiting list were increasingly influential nearer the end of the episode. If a case manager discharges the client during a month when the waiting list is greater than usual, then a client may remain in the program for a longer period of time. Alternatively, if case managers anticipate a longer waiting list, then the intention to discharge the client may occur earlier.

The lack of evidence for the environmental variables can be explained by two very important limitations in the database. First, informal support is measured only at the time of admission. Construction of the variable assumes that the level remains constant throughout the episode. This is not very credible. Secondly, the quality and completeness of the environmental variables is not great. As a result averages were substituted for the missing observations. This is not ideal and the results may change with better data collection. Thirdly, the environmental variables are subject to change but are only recorded at the time of admission. For example, the marital status could change during the episode. Again, the results may be an artifact of the data quality.

In conclusion, the primary determinants of home care utilization that were specified at the onset are supported. The determinants of service intensity are not supported, however alternative hypotheses exist. Finally, several data quality issues need to be addressed before the evidence could generalized to all home care users.

B. Serendipitous Findings

In addition to reviewing the evidence in favour of the Andersen model, there have been some other findings as well.

1. Competing Hazards

Examinations of the results for the separate and combined competing hazards have provided interesting results. The fact that the covariates are significant when the alternatives are death or institutionalization supports the explanation of home care utilization as an alternative to skilled nursing care at a facility. As mentioned last chapter, the support for the model when death is the alternative is positive but not surprising as it could be hypothesized that the same underlying mechanism that results in a loss of functioning and independence is the same that results in death.

The lack of support for explanation of home care utilization when there are other reasons for discharge is not well explained. In part this finding is not surprising given the disparate reasons attached to individuals belonging to this group. The overwhelming support for the all discharges definition of an event (including a significant and correct sign for service intensity) could be an artifact of combining the three hazard categories. On the other hand it should be viewed positively as it will be of service to program planners.

2. Substitution Between Formal and Informal Services

The unexpected finding that substitution does not take place between formal and informal services was a result arising from the two auxiliary regressions. In both cases the coefficient of the complementary service intensity was positive and highly significant ($p<0.000$

in each case). This means that formal service intensity increases as a result of an increased level of informal support. Many hypotheses could be generated. For an advocacy explanation, the case could be made whereby a higher level of informal support will enable the client to demand more services. Another explanation may be that the presence of a caregiver may actually require more service provision in that the caregiver may also be an unintended recipient. For example, suppose that home support were supplied to an elderly woman. Shopping for groceries would naturally include her husband's groceries as well.

Likewise, informal service intensity levels are associated with higher formal service levels (complementary goods). It is important to recognize that informal service levels are not observed but imputed at the time of admission. Therefore the explanation of informal service intensity results requires greater caution. The bottom line is that these results were not expected.

C. Health Policy Implications of Research Results

These results cannot be discussed in a vacuum. The primary purpose of examining the determinants of utilization is that once identified they may aide in the development of policy and program management. Cost containment is the primary driver for many policy decisions in the market for health care services. As a result the question as to the appropriate level of funding is the most important implication. Following this line of inquiry leads one to ask whether certain modes of delivery of home care are effective or different service mixes. Would a copayment for some services be effective? Lastly, should services be provided to a select group of clients whose risk of institutionalization is greatest?

The Andersen framework allows for determination of equity and provides discourse for mutability (Andersen 1968,73). These concepts are summarized in Figures 6.1-2. In Figure 6.1 the shaded covariates which are significant are deemed to be inequitable. In other words correlation of home care utilization and demographic or need variables is equitable. However if insurance, income or other enabling factors dictate utilization, then home care is inequitable.

Figure 6.2 categorises variables into low, medium or high mutability. For policy purposes these are factors which can be altered by the public health system. Highly mutable covariates are the enabling factors of course (eg provision of insurance, lowering of co-payments). Health beliefs are medium mutability while other predisposing and need variables cannot be altered.

In terms of questions facing governments, how do the findings from the dissertation impact on expansion of home care, identification of cost-effective services, and targeting of high-risk (of nursing home admission or death) populations.

1. Expansion of Contraction of Home Care?

Rising home care expenditures are predicated on the fact that they are a cost-effective alternative to nursing home care. Therefore the primary question is to determine the appropriate level of financing home care. That is to say whether to expand or to contract funding of home care services. If we look at the marginal costs of formal service provision (derived from the Cox model results), then we notice that for models A to D the marginal costs are $5000, $500, $500, and -$128 per day respectively. (A negative marginal cost indicates that expansion of financing would increase the probability of a nursing home admission.)

Since the marginal costs for nursing home care is approximately $120 per day, marginal costs of $500 to $5000 would mean that the level of home care is too great and that contraction of services is warranted. However, this is not an effectiveness analysis and self-selection bias is not completely controlled so conclusions of this nature are premature.

2. Program Effectiveness Considerations

The utilization study however provides the foundations for effectiveness and eventually, a cost-effectiveness analysis. Effectiveness may refer to the level of services, provider-, and service-mix as well as the mode of delivery. For example, the data allows for sub-analysis of provider and service types. It appears for the descriptive data that the service mix has remained constant and that the service mix has shifted in favour of professional nursing services. The mode of delivery is a popular area of analysis, as it is becoming increasingly popular to offer some services at a centralized community setting as opposed to the home. Also, the frequency and timing of visits may be of consequence. The mode of delivery is not recorded (as most services provided here are in the home) and the visit variable is calculated so conclusions in this regard are premature as well.

3. Identification of High Risk Populations

If provision of home care services follows the trend in other health sectors (especially hospital care), then it may become necessary to ration these services in the future. In order to ration home care services effectively, it is important to identify sub-populations that are at higher risk of requiring nursing home care. This model would enable the case manager to calculate the exact probabilities for each hazard and to determine the consequences of withholding services.

D. Limitations of the Study

In addition to the methodological limitations discussed in chapter IV, some concerns remain. This section will summarize the general concerns that are inherent to this study setting.

1. Self-selection Bias

Despite the anticipation and statistical methodology employed for self-selection bias, there is no formal statistical test as to whether the problem remains or whether the predicted service intensity variables have been in fact purged of their correlations with the error term in the Cox model. The procedure developed here is limited by the lack of rigorous, statistical methods concerning unobservable, censored data. Furthermore, they are limited by the data themselves. For example, longitudinal data on the level of functioning, disease burden, and external environmental variables would aid greatly in addressing the issue of whether self-selection bias can be adequately dealt with in this context.

2. Data Quality

Although not discussed at length, the issue of administrative data quality is of importance. Because of political constraints resulting in the inability to conduct randomized trials in the provision of home care services and the inability to monitor crossovers, analysis will continue to depend on observational data. Completeness, consistency problems associated selected variables, and the fact that the data remodelling (i.e. conversion to episode and visit) exercise was extensive should result in a recommendation for improved funding of the home care program's information service department.

3. Absence of Other Resource Consumption Data

The home care clients consume other formal services. For example, they may visit their doctor or be admitted to a hospital. Likewise, there are a great number of community based (usually volunteer, e.g. meals on wheels) programs which provide services to these clients and may help to explain some of the additional variation. The absence of this data in the database means that service intensity is incomplete.

E. Implications for Further Research and Practice

Despite the limitations in the study, a number of alternative hypotheses have been proposed. In no particular order the research agenda should include outcome measurement, cost-effectiveness analysis using this model, and linkage with other databases.

1. Outcome Measurement

Home care clients in this population are scheduled for an assessment of physical function once per year. In fact only 20% of clients have received more than one assessment. Home care is meant to provide independence from limitations arising in physical and mental functioning. It would be prudent to measure functioning more frequently (6 months?) and especially at the time of discharge to ascertain the effect of home care services on the rate of decline in average functioning levels.

To this point the analysis implicitly assumes that one day at home equals one day at an institution. Measurement of alternative health status indicators would enable a weighting of days at home versus days I a nursing home. Since the weighting is equal, the sole criterion for efficiency is the dollar. For example, $120/day in a nursing home may equal $200/day or $50/day at home. Both functional assessments and alternative health status measurement have important implications for evaluating program efficiency.

2. Cost-Effectiveness Analysis

At the end of the day, the question of efficiency is to be addressed. Conducting a cost-effectiveness analysis with observational data requires two things. Firstly, a clear understanding of which covariates explain resource consumption is necessary to control for this effect. Secondly, an understanding of self-selection bias is necessary to control for the unobserved covariates. These are the two challenges in conducting a cost-effectiveness analysis.

3. Prospective Linkage With Other Databases

In order to get a full picture of the effects of the health care system on the home care client, the role of other health services must be addressed. In addition, acute care services may provide a mechanism to address the self-selection problem. If we compare two clients with the same disease classification and functioning level but differing levels of acute care requirements during an episode of home care, then variations in formal home care service intensity may be explained more adequately.

Also, with increased pressures to enter clients into home care, and with the presence of caregiver stress, it is important to address the extent to which other community services are relied upon. This will allow the researcher to more fully understand the effects of substitution of services between the formal (health care system) and informal (family and community).

Figure 6.1: Andersen's Behavioral Model: Equity

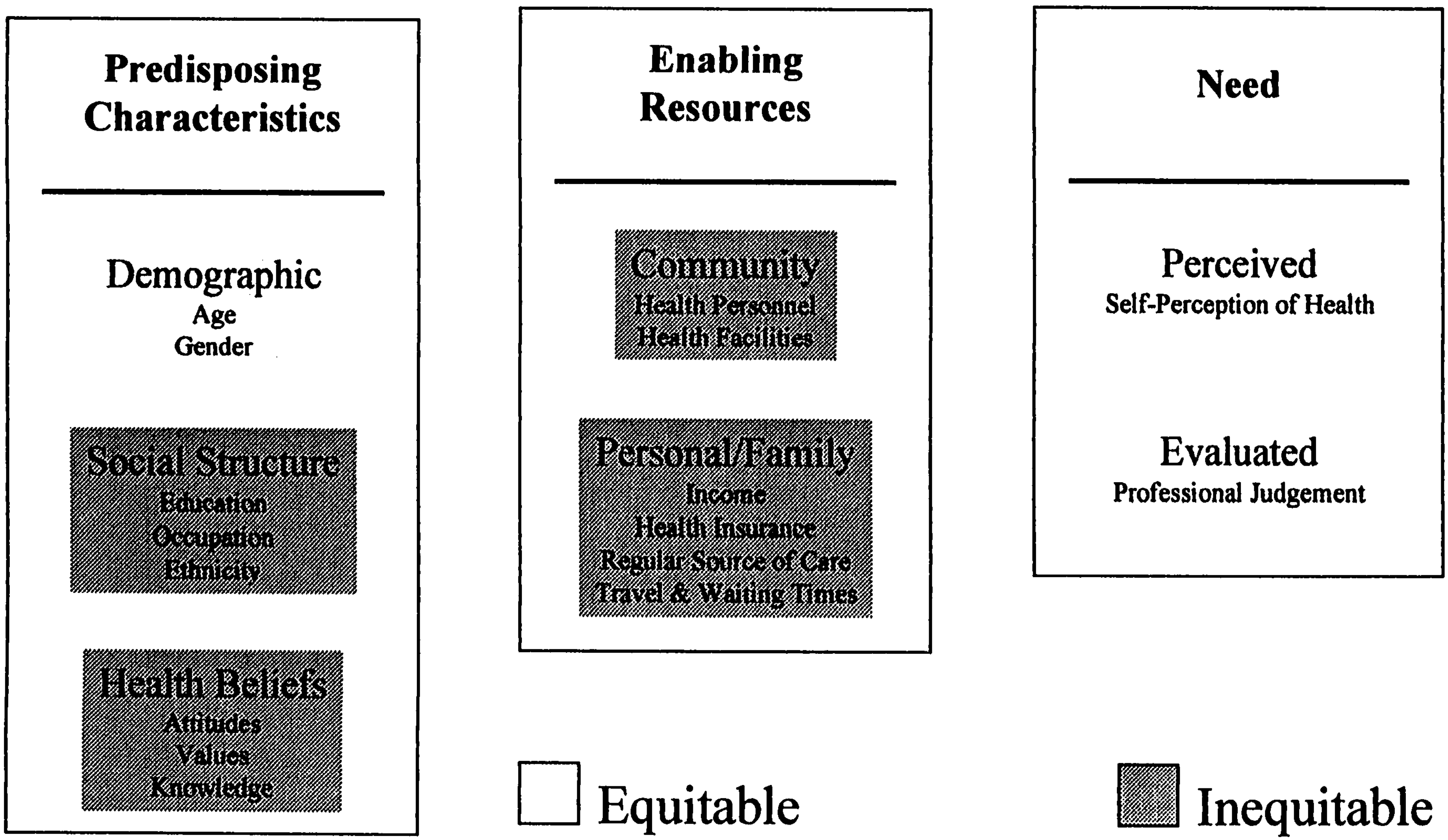
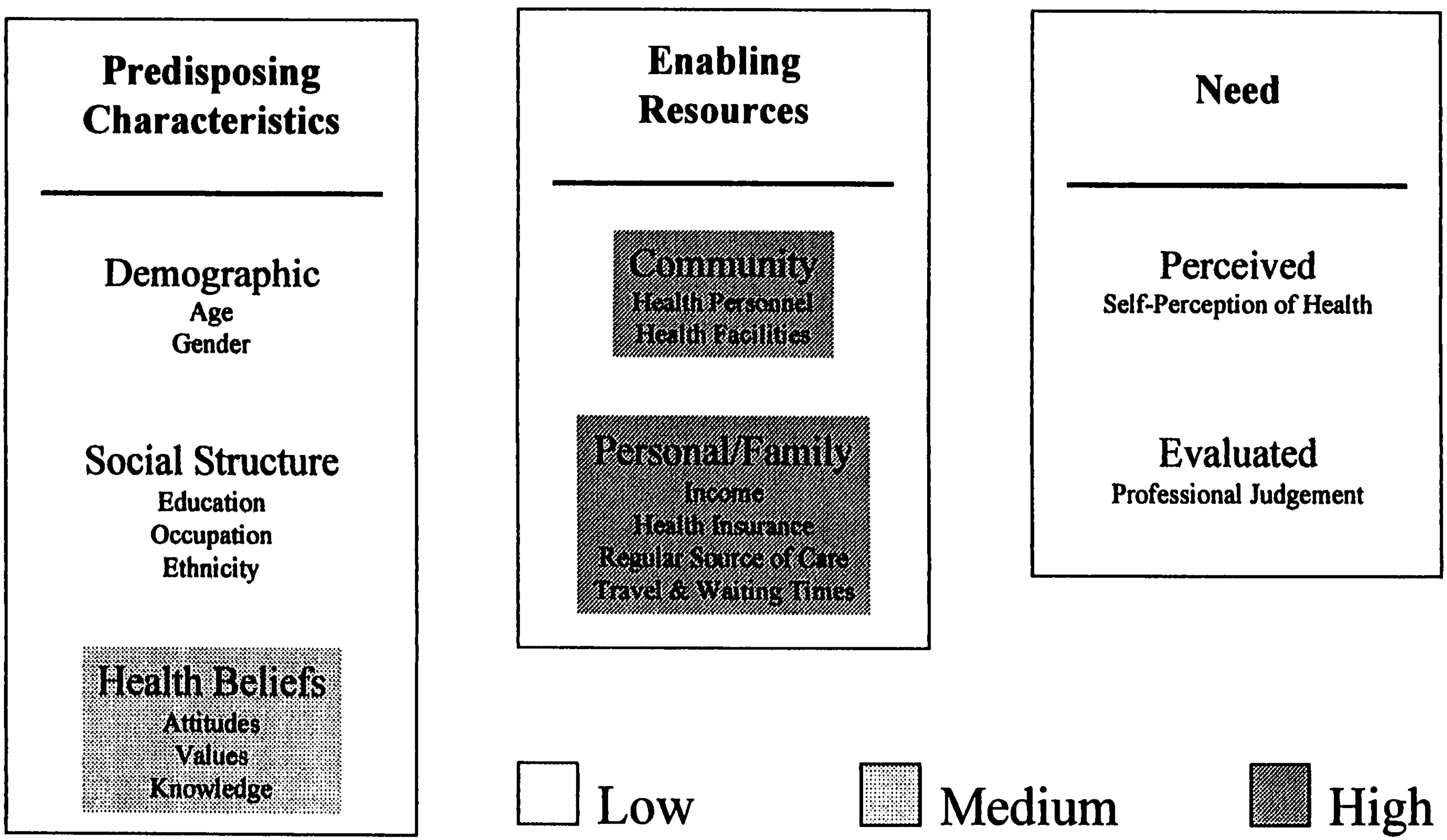


Figure 6.2: Andersen's Behavioral Model: Mutability



Bibliography

Andersen, R.M. (1968) Behavioral Model of Families' Use of Health Services. Research Series No. 25. Chicago, IL.: Center for Health Administration Studies, University of Chicago.

Andersen, R.M. and Newman, J.F. (1973): Societal and Individual Determinants of Medical Care Utilization in the United States. Milbank Memorial Fund Quarterly Journal, 51:95-124.

Greene VL, Lovely ME, Ondrich JI. The cost-effectiveness of community services in a frail elderly population. Gerontologist 1993;33(2):177-89.

Liu, K., McBride, T.D., and Coughlin, T.A. (1990): Costs of community care for disabled elderly persons: the policy implications. *Inquiry*, 27:61-72.

Rivlin AM. Simulating policy alternatives for long-term care: an example of the cochrane-orcutt approach. Journal of Economic Behaviour and Organization 1990;14: 53-64.

Scanlon, W.J. (1980): Nursing home utilization patterns: implications for policy. *Journal of Health Politics, Policy & Law*, 4:619-641.

Shaughnessy, P.W. (1985): Long-term care research and public policy. *Health Services Research*, 20:489-499.

Weissert WG. Seven reasons why it is so difficult to make community-based long-term care cost-effective. Health Services Research 1985;20(4):423-33.

IMAGE EVALUATION TEST TARGET (QA-3)

150mm

6"

APPLIED IMAGE . Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989