The Dissertation Committee for James Frederick Lukenbill Certifies that this is the approved version of the following dissertation:

A Proportional Hazards Model for the Prediction of Psychiatric Rehospitalization

Committee:

Gary D. Borich, Supervisor

Barbara G. Dodd

Edmund T. Emmer

Ron P. Jemelka

Michael T. Johnsrud

A Proportional Hazards Model for the Prediction of Psychiatric Rehospitalization

by

James Frederick Lukenbill, B.A., M.A.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

December, 2003

Dedication

To my wife Julie, and my son Aaron

Acknowledgements

First of all, I would like to thank Dr. Gary Borich who has been a tireless advocate for me during the dissertation process and also in securing employment in the statistical crafts. I would also like to thank the members of my committee whose tireless efforts and patience with me allowed me to bring this process to fruition.

I would also like to thank my parents, Dr. Bill and Shirley Lukenbill who encouraged me to press on with my education and supported me both financially and emotionally. Special thanks go out to my mother whose ability to edit long documents accurately will always exceed my own.

Dr. Ron Jemelka also played a pivotal role in bringing this project to a successful conclusion. Ron gave me my first job dealing with “big” data and made available the data used in this study. Without his support, this project would have never come to be.

Dr. Michael Johnsrud played a significant role by allowing me easy and timely access to the vendor drug data which played an important role in this study.

Finally, I want to thank my wife for putting up with my squandering sunny days which could have been spent in the park with our son Aaron. I'd also like to thank Aaron who always had a smile for his dad when the goals seemed too lofty or the paperwork too deep.

A Proportional Hazards Model for the Prediction of Psychiatric Rehospitalization

Publication No._____________

James Frederick Lukenbill, Ph.D.
The University of Texas at Austin, 2003

Supervisor: Gary D. Borich

Cox regression and logistic regression models were used to explore the predictors of psychiatric rehospitalization in 558 Medicaid managed care beneficiaries. The Chronic Disability Payment System (CDPS) was used to create measures of psychiatric morbidity based upon the subjects' diagnostic history the year prior to the study year. Younger subjects, who possessed a larger number of co-morbid psychiatric diagnoses, as well as a diagnosis of schizophrenia in the year before rehospitalization behavior was measured, were found to be rehospitalized at a greater rate and to have less community tenure from discharge to rehospitalization than subjects that did not have these characteristics. Subjects who had outpatient psychiatric care in the interim between discharge and rehospitalization were also found to be at greater risk for rehospitalization. No significant interactions were discovered. Theoretical as well as practical implications of these findings are discussed.

Table of Contents

List of Tables

List of Figures

CHAPTER 1: INTRODUCTION

The essential task of this dissertation is to describe and to explore the factors which influence the time between inpatient behavioral health visits in Medicaid managed care. In order to contextualize the environment in which these services are delivered, it is necessary to explore the mechanisms that society has built to care for the mentally ill and also to explore the evaluation techniques and approaches which have evolved with these systems in order to gauge the effectiveness and the appropriateness of the care that is delivered. In order to do this, I will first offer a brief look at the experience of a large Southwestern state (Texas) in dealing with spiraling Medicaid costs. Then I will explore at length how American society as a whole has attempted to hold the line on health care costs, while presumably maintaining the quality of medical care delivered in both public and private settings. After that, I will look at the evolution of medical evaluation in general and the evaluation of managed care in particular with a brief look at how statistical thought has, in effect, "grown-up" with the evaluation of medical care. This discussion will then lead to the primary purpose of this study, a discussion of the evaluation approaches brought to bear against the problem of measuring quality in Medicaid behavioral health service delivery. I will focus especially on the measurement of recidivism in inpatient care and the various pros and cons of the quantitative evaluation techniques that can be brought to bear against the problem. Although many authors have struggled with the problem of modeling rehospitalization, few have attempted to use the latest statistical techniques in extracting all of the information from the data available. Many authors have chosen logistic regression; although far superior to chi-square tests, this and other traditional regression techniques expose the practitioner to biases from censoring and discard much of the information potentially available in the data (Allison, 1984; Allison, 1995). Besides the lack of what this author feels is an appropriate choice of statistics, the field has also suffered from a lack of an appropriately derived variable which is a proxy for psychiatric morbidity. A possible remedy for this problem can be found in actuarial science in the form of diagnosis grouping technology. I will show that, although designed to predict costs, these technologies can also be useful as risk-adjustment strategies for empirical studies in health services research. This dissertation will attempt to advance the field of rehospitalization prediction by applying Cox regression to the question of recidivism while also controlling for psychiatric morbidity and a host of other variables found to be relevant from a review of the literature.

CHAPTER 2: LITERATURE REVIEW

A Short History of Public Behavioral Health Finance

Providing behavioral health services to the indigent is a very expensive proposition and fraught with complexity and fragmentation. Lutterman (1994) states that the number of state behavioral health agencies operated in the United States were almost 3,000 consuming approximately $14 billion dollars of taxpayers' dollars (Frank & McGuire, 1996). The latest figures available for Texas from the Texas Health and Human Services Commission (THHSC) report that for fiscal year 1998 the Texas Department of Mental Health and Mental Retardation alone consumed $744,372,125 (THHSC, 2002). Total Texas Medicaid appropriations for 2002-2003 are about 25 billion dollars with 10 billion of this coming directly from state general revenue (Dunkelberg, 2002). This represents a substantial drain on the budget of Texas government, representing 15% of state general revenue and 72% of THHSC funding (Dunkelberg, 2002). Of the $106.8 billion dollars of state revenue generated annually, about half ($54.5 billion) is generated from taxes. Sales taxes make up approximately $14 billion (State Fiscal Year [SFY] 2000) or 28% of state revenue with $14.8 billion (SFY 2000) coming from the federal government (Dunkelberg, 2002).

State budget officials have had considerable difficulty in correctly forecasting Medicaid caseloads and expenses due to failures both in the Medicaid forecasting mechanism and in the budgeting process which does not allow for inflationary increases in the two-year budgets that are submitted by agencies except for special items. The two-year budget cycle and budgetary idiosyncrasies force the state Medicaid agency, the THHSC, to submit a budget which, almost by default, will be inaccurate. This budgetary process also obfuscates the real inflationary trend of Medicaid spending (Dunkelberg, 2002).

Texas enjoyed significant declines in Medicaid enrollment during the period from 1996 through 2000, due in part to welfare reform and a roaring economy. Monthly Medicaid enrollment decreased 13% overall from January 1996 to December 2000, but spending also grew by 26%. Texas projected a 17% increase in Medicaid enrollment in 2001-2002 (Dunkelberg, 2002). Currently, Texas, like many other states, is facing a substantial (5 billion dollar) budget deficit for SFY 2003 (9/01/2002 – 8/31/2003) (Associated Press, 2002). A substantial portion of this shortfall is related to the confluence of a flagging economy, growing Medicaid rolls, and relaxed Medicaid eligibility requirements (Dunkelberg, 2002). The current budgetary climate has made the delivery of behavioral health care services to the neediest Texans challenging, with lawmakers looking for ways to cut costs while ostensibly not cutting services.

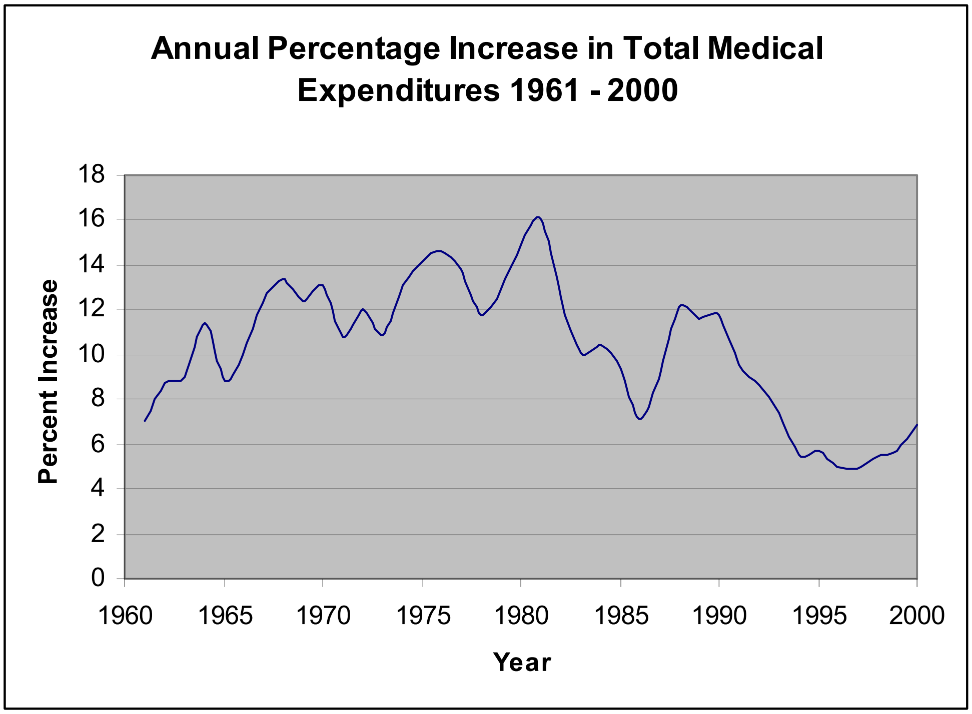
One of the major ways Texas and other states have chosen to do accomplish this goal through the expansion of Medicaid managed care (THHSC, 2002).

THE HISTORY OF MANAGED CARE

For many years traditional insurance, often called "indemnity" insurance, was the norm. Starting during World War II, employers used the new tool of health benefits to lure workers into their employ in an increasingly tight labor market. The Internal Revenue Service made the practice of compensation by benefit even more attractive in the 1940s by making the employee's portion of the health benefit tax-free. Therefore, employees were able to use 100 percent of their wages to purchase benefits before the federal government took a cent.

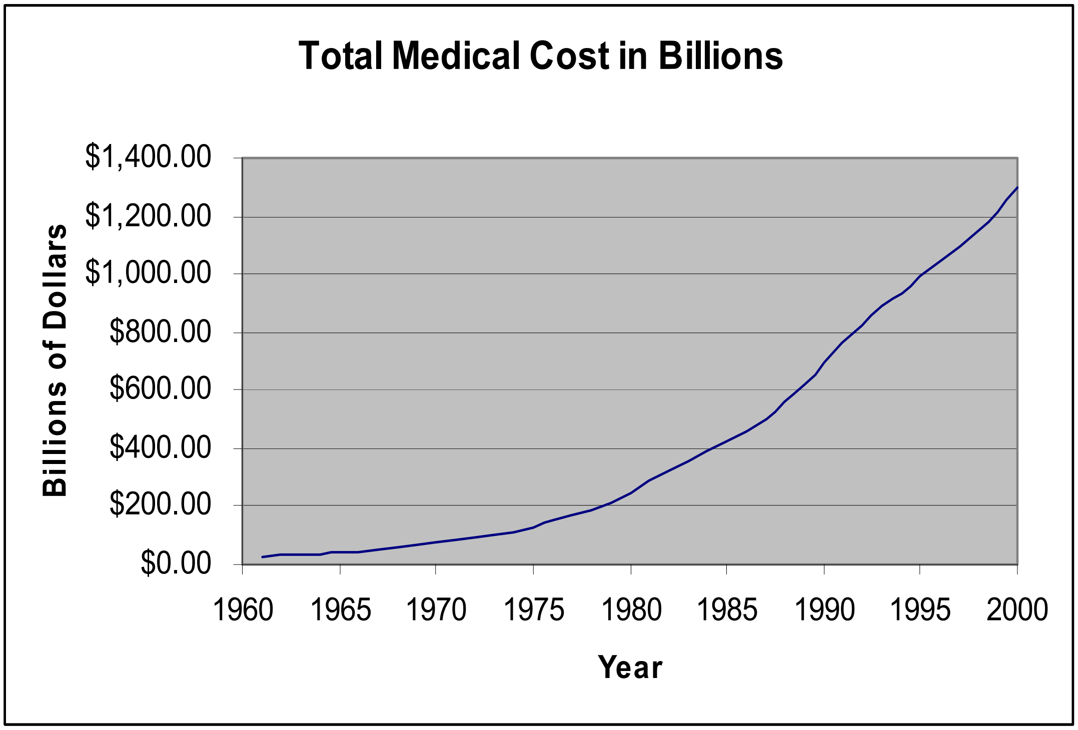
This was, of course, extremely attractive and also served to dilute the relevance of increasing health care costs to the average employee (Zelman & Berenson, 1998). The distinct lack of relevance that rising health care costs had to the average insured consumer and the putative inability of government or business interests to take effective action resulted in an annual 12.5% annual growth in per capita expenditures—fully 5% greater than the underlying rate of inflation (Merril, 1985). The Kaiser Commission on Medicaid and the Uninsured (2002) notes that attempts at controlling the growth in health care spending have resulted in short-term reduction in the growth of spending, but no long-term fixes have emerged for this vexing problem. In the early 70s wage and price controls reduced costs; in the late 70's voluntary efforts had some effect; finally in the 80's and 90's managed care began to decrease the inflationary rate, but only for a time.

Figure 1: Annual Percentage Increase in Total Medical Expenditures, 1961-2000



Source: United States Centers for Medicare and Medicaid Services. (2002)

Figure 2: Total Medical Cost in Billions, 1961-2000.



Source: United States Centers for Medicare and Medicaid Services. (2002)

By the late 1960s policy makers were beginning to take notice especially as the true expense of the new national health care programs of Medicaid and Medicare began to be known. In 1967 President Johnson summoned experts to a conference to discuss what could be done about the burgeoning drain of U.S. health care costs on the U.S. economy. The effect of rising health care costs was probably particularly salient at this time as the Johnson administration had virtually abandoned all hopes of curtailing costs in Medicaid and Medicare in concessions to the American Medical Association in order to ensure the speedy passage of Medicaid and Medicare, the crucial pillars of his “Great Society” program (Zelman & Berenson, 1998).

Although the alarm was sounding in the late 1960s about spiraling medical inflation, no real abatement in the trend was felt until the mid-1990s (Zelman & Berenson, 1998). Why? The reasons are many: first and foremost is the fact that physicians at the time were in the enviable position of controlling both the cost of care and the supply of care. Consumers with indemnity insurance or public insurance through Medicare or Medicaid could simply choose any physician they wished to see, the physician could then provide any and all care he or she deemed necessary with no concern that the costs for the care--however exorbitant--would be paid. Consumers also bore little of the cost increases directly because the costs of their insurance and medical care were subsidized by the government in the form of pre-tax benefits to the employees themselves and the health benefits the employers paid which were deductible business expenses.

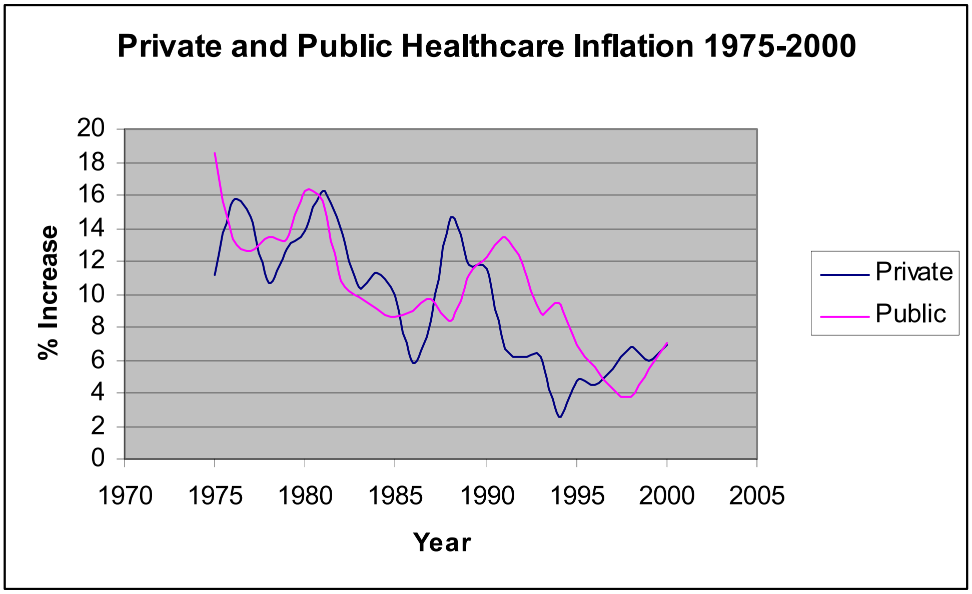
Curiously, the largest government interventions in this deepening quagmire of increasing costs occurred during Republican administrations. In the early days of his administration, Richard Nixon spoke of the nation’s health care inflation situation as a “Crisis” (Zelman & Berenson, 1998). In 1971-1973 general wage and price controls were imposed in an effort to curb what economists have termed “stagflation” (Rockoff, 2002). As mentioned previously, these cost controls had no long-term impact on controlling the rise of health care expenditures. Realizing there was no end in sight to the spiraling costs, the Nixon administration attacked health care costs; the pillar of its legislative campaign was the HMO Act of 1973. This Act passed with broad support from both Democrats and Republicans. It was forecast that by the end of the decade there would be over 1,000 HMOs serving over 60 million people (Zelman & Berenson, 1998). This estimate was grossly optimistic. The Carter administration attempted to control hospital costs but was unable to force a legislative agenda effectively through Congress (Zelman & Berenson, 1998). It was not until the Reagan administration that something substantive was done to reduce costs.

In 1983 the Reagan administration passed legislation that revolutionized the manner in which inpatient Medicare services were reimbursed. Called the prospective payment system, this system reimbursed facilities a set amount based upon diagnosis, abandoning the fee-for-service system that had been in place since 1965. Under the previous fee-for-service system, facilities and doctors were paid more for doing more. Thus, more tests and more time in an inpatient facility would result in higher reimbursement from the federal government. After the introduction of a prospective payment system, the government would only pay a set amount based upon a diagnostic classification no matter how long the stay in the facility. This new reimbursement scheme possessed the main mechanism employed by the now emergent managed care organizations: financial risk borne by the provider of services. By implementing this revolutionary change, the federal government accomplished two things: 1) it took care of its own problem with the program for which it had direct fiscal responsibility; 2) it sent the message to health care purchasers that at least some measure of control could be exerted to rein in spiraling health care costs (Zelman & Berenson, 1998). This would prove to be one of the pivotal events in American health care finance.

THE "OVERNIGHT FALL" OF INDEMNITY HEALTH INSURANCE

Throughout the 1980s the tremendous transfer of wealth from private citizens and the government to health care providers continued to increase until the cycle of greed reached its apogee in 1986. During 1986, private employee-sponsored premiums increased from an astounding 24 percent (Sullivan & Rice, 1991) to 14 percent according to data from the United States Centers for Medicare and Medicaid Services (USCMMS) (2002).

Figure 3: Private and Public Healthcare Inflation, 1975-2000



Source: United States Centers for Medicare and Medicaid Services. (2002)

Although the range of estimates for inflation for indemnity insurance expenditures are obviously quite wide, using either estimate, it was becoming clear to many that the cycle of constant medical inflation was destroying the competitiveness of U.S. products in world markets and eroding the wages of U.S. workers. Employers were no longer absorbing the costs of employee health insurance plans and handing workers skimpier paychecks; instead they had started to pass the inflation straight to the employee. This development was perhaps the straw that broke the camel's back. For the first time the American worker was now staring straight down the barrel of the result of his/her unintentional complicity in a system that rewarded unnecessary tests, treatments, and hospital time, all in the name of physician autonomy and quality of care. What happened after employees began to complain is truly striking.

By the end of 1997, almost 80 percent of employees enrolled in an employee sponsored health plan were enrolled in managed care (Zelman & Berenson, 1998). The old system of paying providers whatever they asked with no measurement of quality proved to be so inefficient at delivering care that it unraveled quickly for such a large institution. The epoch of capitation was at hand. Capitation is the practice of paying an insurer, or in some cases a provider of medical services, a flat rate for providing care for an individual. This abrupt shift in private insurance pricing in the mid to late 1980s was followed by another shift in the early 1990s in public insurance which served to reduce medical inflation in private and public insurance by 50% to 70% (see Figure 3). These abrupt reversals of price trends directly correspond to the time when managed care was introduced into private and public settings. By 1996 annual employer-sponsored premium increases had decreased to .5 percent (Zelman & Berenson, 1998) or about 2 percent using USCMMS data (see Figure 3). Managed care had proven its ability to turn around the American system of health care finance, apparently with no decrease in quality, despite numerous media blitzes to the contrary. With the coming of managed care also came the demise of many independent providers--they rapidly coalesced into provider organizations to gain bargaining clout with HMOs. The effectiveness of Medicaid managed care in battling medical inflation in the private sector also caught the attention of Medicaid directors in many states; this resulted in the introduction of managed care to many of America's neediest citizens.

MEDICAID JUMPS ON THE BANDWAGON

In the first two years of the 1990s, Medicaid expenditures increased at a 27 percent rate. Some of the increases were due to expanded federal eligibility requirements, while some were due to the pernicious effects of general medical inflation (Ku et al., 2000). Whatever the underlying reason, states were eager to try revolutionary reforms in the delivery of Medicaid services in an attempt to bring costs down. With managed care's notable cost-containment advantages beginning to surface in the privately funded sphere, plus the putative ability to measure the quality of care delivered, states began to adopt mandatory Medicaid managed care programs during the mid-1990s under section 1115 research and demonstration waivers and 1915(b) waivers (Ku et al., 2000). Section 1915(b) waivers are granted by the federal government to states for a two-year period, after which time they may be continued indefinitely. Jordan, Adamo, and Ehrmann (2000) describe the key provisions which attract states to the 1915(b) and 1115 demonstration programs. The key provisions of the 1915(b) program, which allow some requirements of section 1902 of the Social Security Act to be waived are: 1) Freedom of Choice 1902(a)(23) -- allows states to mandate enrollment in a Medicaid managed care plan; 2) Statewideness – allows states the ability to choose specific geographic areas for managed care implementation while leaving other areas out; 3) Comparability of Services 1902(a)(10) – allows states to mandate managed care for some eligibility groups while allowing others to remain under the care of the traditional indemnity system. The section 1115 program goes even further than the 1915 waiver program allowing states to waive the aforementioned provisions that the 1915 act provides plus it allows states, at their discretion, to expand services, require co-payments from beneficiaries, and avoid reimbursing some safety-net providers at fee-for-service like rates. Boden (2000) describes a rapid rise of Medicaid managed care under the federal waiver system:

> In 1990, only a handful of States had waivers under 1915(b), and only two (Arizona and Minnesota) had 1115 waivers to implement Medicaid managed care. As of June 1999, 34 states had at least one 1915(b) program, and 16 states were operating research and demonstration projects under 1115 that involved managed care in some way. The percentage of Medicaid eligibles in managed care increased from approximately 10 percent in 1990 to 55 percent in 1999, of which 42 percent were enrolled in capitated managed care, and 13 percent were assigned to a primary care case manager. (Boden, 2000, p.1.)

States participating in 1115 waiver projects in the early to mid 1990s tended to have the following characteristics: 1) compulsory enrollment in a capitated managed care plan; 2) participating plans agreed to abide by often onerous state regulations; 3) enrollees had to have several choices of health plans (Ku et al., 2000). All of the initial programs also faced problems with a number of issues: 1) insuring the financial viability of so-called "safety-net" providers, providers which historically served a large number of Medicaid recipients; 2) setting capitation rates (establishing the payment per person to a health care entity, usually and insurer) ; 3) making the Medicaid managed care market sufficiently lucrative for private insurers to guarantee participation by several insurers in a given market without breaking the state budget; 4) updating state data systems or hiring contractors to evaluate the quality and access to care delivered under Medicaid managed care; and 5) ensuring access to behavioral health services.

The states tended to take varying approaches to many of these issues. For the "safety-net" providers and behavioral health providers, debate still rages as to whether significant financial harm has been done to these providers. States have had varying success with attracting and keeping managed care plans, although plans that exited the programs tended to be smaller commercial plans instead of larger plans set up specifically to serve the Medicaid market. Most states took conservative approaches to setting capitation rates in the early 1990s with only Maryland adopting a diagnosis-based risk adjustment system, adapted from Johns Hopkins University's Ambulatory Care Groups risk-adjustment system (Ku et al., 2000).

Related to this point is the problem that states encountered in the early 1990s and from which they still continue to suffer: getting timely and reliable encounter data. Without quality encounter data, diagnosis-based risk adjustment systems cannot work properly and can lead to disastrous under-capitation of plans, leading to large losses and eventual abandonment of Medicaid insurance products. Many plans in the early 1990s tended to "carve-out" behavioral health services to distinct behavioral health insurance services. Evidence is ambiguous, and debate still rages as to whether these Behavioral Health Organizations (BHOs) provide better, more cost-effective care than packaging behavioral health benefits with the larger Medicaid managed care organization.

The debate as to whether or not managed care, and in particular Medicaid managed care, has increased the quality of care still remains vibrant. Some authors maintain that at least a mechanism to measure quality is now in place, which is a vast improvement over the fragmentation of traditional indemnity insurance. Under traditional indemnity insurance, there was virtually no system, just a loose collection of doctors where any attempts at care coordination and outcomes measurement were probably more a result of personal commitment than health policy (Zelman & Berenson, 1998). Other authors (Geller, 2000) are openly critical of the system, and believe that the rise of HMOs has wrought a slew of new problems including fragmentation of care, "revolving door" syndromes at psychiatric inpatient facilities, and dubious improvements in quality of care.

SUMMARY OF U.S. BEHAVIORAL HEALTH FINANCE

The provision of behavioral health services to indigent individuals in the United States echoes the struggles that state and federal governments experience to provide a broader array of health services. Medicaid managed care has been one of the major weapons that policy makers have chosen to wield against the almost constant strain of continued medical inflation (Dunkelberg, 2002; THHSC, 2002; Kaiser, 2002; Zelman & Berenson, 1998). Unfortunately, the remedies to the medical inflation perils often see short term reductions in the annual rate of medical inflation, but long term relief has not been realized (Zelman & Berenson, 1998; Kaiser, 2002). Although traditional indemnity health insurance was dealt a quick and fatal blow in the early 1990s (Zelman & Berenson, 1998), the palliative effect that the HMO revolution brought appears to be exhausted (USCMMS, 2002) and some influential authors wonder if all HMOs have brought is a fragmented system with poor behavioral health care (Geller, 2000).

Medicaid managed care has also taken the country by storm with states opting to embrace the concept of capitated private insurers vying for state contracts under the 1915(b) or 1115 waiver programs. This system has been far from stable, with many private insurers dropping out of Medicaid programs because of onerous regulation or the failure of states to pay them equitably (Ku et al., 2000) perhaps due to a lack of knowledge, or a hesitation to implement diagnosis-based risk adjustment technologies (Ku et al, 2000). The lack of financial stability in insurers delivering behavioral health services to Medicaid consumers is clearly cause for concern and may lead to further fragmentation in the publicly funded behavioral health network (Geller, 2000).

Measuring "Quality" in Medicine

THE IMPOSSIBLE DREAM: COST SAVINGS AND QUALITY

The measurement of quality of care in medicine has had a traditionally rocky road of acceptance. Paradoxically, providers who are viewed by many to be "men of science" did not enthusiastically support the idea of being measured or compared in any meaningful manner and have in many cases sought to ostracize those who sought to measure their performance (Iezzoni, 1997). With the coming of managed care, there was also increased clout to impose the measurement of quality upon providers, providers who were not always very interested in being measured and compared to other providers. Fortunately, statistical techniques that even the playing field for providers who have sicker patients (have more "illness-burden") vis-à-vis another provider had been developed years earlier and were waiting in the wings with health services evaluation professionals who realized that the only way to improve the system was to measure the outcomes the system produced. As we shall see, this was more easily said than done.

THE EVOLUTION OF QUALITY MEASUREMENT AND THE HEALTH PROFESSIONS

The practice of tracking the outcomes of patients undergoing medical procedures, and further still, the practice of statistically adjusting the "illness burden" of these patients in order to fairly compare outcomes produced by one treating entity vs. another can still be regarded as in its infancy. The English collected "Bills of Mortality" to track the advance of the plague as early as the 1500's (Walker, 1929); and the industrial revolution hastened these advancements as the impact of packing large numbers of factory laborers into cities under squalid conditions began to become apparent. It is perhaps under the crushing weight of this profound change in the population distribution and centers of economic activity that the statistical thought focusing on public health began to percolate, first as a method that attempted to precisely delineate the movements of populations, and then to measure the health effects of concentrating a large number of people in a small geographic area. Two distinct lines of thought can be seen as converging to create the modern study of outcomes measurement and risk-adjustment as we know it today: the attempts to improve the horrid state of English hospitals in the 1800s and the rise of the "statists" in 19th century England.

Before the advent of the statists in the early 1800s, studying large numbers of people, or for that matter, large numbers of anything had been particularly challenging. Before this point, the belief that large-scale regularity could be discovered by the aggregation of many objects was largely foreign and based upon the concept of "political arithmetic." Political arithmetic is described by Porter (1986) as the practice of English noblemen enhancing the values of their land holdings by using imperfect census figures and placing a value on each inhabitant of their holdings based upon some schedule. Indeed the value of the holdings of a king or nobleman was a function of the number and character of the inhabitants of the nobleman's land.

Political arithmetic was increasingly supplanted with a more formalized version of statistics as European bureaucracy became more centralized and as the French Revolution destabilized European thinking to such a degree that the essential essence of man began to be questioned by the power-elite (Porter, 1986). Although this new energy that flowed from the French Revolution and the burgeoning industrial revolution showed great promise of sparking a statistical movement in France originating from a new concern for public health, it is to the English statistical societies forming in 1833-1834 where modern statistics can trace its roots.

These private societies were more of an outlet for noblemen than a true forum of scientific exchange but some revolutionary thought can be traced to them. One of the key individuals of this epoch is Adolpe Quentelet who drew from error laws developed in astronomy in order to impose a scientific order upon social behavior previously thought to be random. This belief in an underlying order to social phenomena and events is a theme which permeated many of the movements of the time and allowed the first reform of hospitals to occur in England.

In the early 1800s, English hospitals had the unfortunate distinction of being barely distinguishable from workhouses for the infirm. Records kept for the benefit of those who supported the hospitals gave some of the first health services researchers a chill; one's chance of surviving was actually better at home than in an English hospital of the time (Iezzoni, 1997). The state of hospital care was not much different in the military. In the mid-1880s Florence Nightingale was busy working to improve the sanitary conditions of English military hospitals. In one instance she was able to make such a difference that the mortality rate was reduced from 42.7 to 2.2 percent at one installation in Turkey (Cohen, 1984). Upon her return to England, she sought help from persons involved in the immature statistical movement in order to improve her work in military hospitals. She partnered with William Farr who promised to help Nightingale with her reforms of military institutions if she would focus some of her energies upon the urban hospitals in England. Together the team discovered that the major hospitals in England had an appallingly high mortality rate of 90.84 percent (the denominator leading to this statistic is highly suspect) (Iezonni, 1997). They were also able to uncover the causes of some of the mortality. The discovered links between mortality and location of hospitals, e.g., hospitals located close to open-pit sewage dumps had higher mortality, and links between facilities management and outcomes, e.g., facilities with poor sanitation often experienced higher mortality. Although Nightingale had made significant advances in outcomes management by pairing the then new statistical techniques with good sanitation practice, Iezzoni (1997) notes that many of the problems health services researchers battle with today would be very familiar to Nightingale: the paucity of outcome statistics; the lack of comparability between patient populations (e.g., one hospital's population may be sicker than that of a comparison hospital); and hospitals discharging patients only to be readmitted to yet another hospital where they expire and count deleteriously in the hospital's mortality statistics.

In America the study of hospital outcomes was vigorously promoted by a pair of Boston area surgeons, Ernest Codman and Harvey Cushing. Codman's interest in outcomes measurement arose from a challenge between the two classmates. After experiencing the unfortunate death of the first patient Cushing anesthetized, and the concomitant lackadaisical comments of the senior surgeon to the effect that such events were quite common and unavoidable, Cushing and Codman challenged each other to keep track of their patients' outcomes throughout their training. From this endeavor arose Codman's "End Results System" where he proposed that each patient be tracked to establish if the treatment was successful, and if not, the reason for the treatment's failure. For his trouble, Codman grew increasingly at odds with the medical establishment as he grew more and more committed to the discipline of tracking outcomes. Finally, he resigned from Massachusetts General Hospital where he was practicing to establish a small ten-bed facility where outcomes measurement would be rigidly applied. This facility soon fell on hard times due to lack of referrals from Codman's more conservative contemporaries who took a dim view of Codman's outcomes measurement initiative. Fortunately, Codman's effort was not in vain as the combined force of the English statists paired with Nightingale on one side of the Atlantic and Codman on the other, was enough to start the tracking of outcomes inexorably in a positive direction. No longer would patient deaths be simply dismissed as bad luck or fate (Iezzoni, 1997).

One of the challenges that both Codman and Nightingale understood was the fact that comparisons among hospitals and practitioners were necessary to understand the strengths and weaknesses of the care delivery system (Iezonni, 1997). Another fact which did not escape them was that not all patients had the same level of "illness burden," meaning that some patients were sicker than others and, therefore, more likely to increase the mortality figures of one hospital through perhaps no fault of that particular medical facility. Although the appropriate mathematics to adjust patient loads statistically between hospitals had been developed by astronomers before Nightingale's time, the possibility of applying similar mathematics to the problem of risk adjustment between hospitals and other medical facilities was just beginning to appear possible shortly before Codman's arrival on the scene.

The connection that occurred to make this possible appears to be traceable to Adolphe Quentelet, an astronomer and a social scientist. Quentelet made the leap of studying the error curve used commonly in astronomy, and realizing the entire error distribution mimics the objective truth of humanity; thus when thought of in anthropometric terms, the very center of the curve would reflect the average man or *l'homme moyen.* Quentelet's average man was controversial, but the concept allowed Quentelet the previously unavailable freedom to compare all manner of statistics of "average men" from differing countries and localities. Quentelet's studies tended to focus upon the moral end of the spectrum, with the construction and comparison of statistics related to drunkenness, insanity, suicide, etc. Quentelet viewed the average man as a way to smooth-out the random error in society and reveal the underlying regularities. And for this he is recognized, at least by some, as being the "gatekeeper" to mathematical social science (Stigler, 1986). From this point, the well known Francis Galton applied the error theories borrowed from astronomy to study examination scores to develop theories of inheritance. Later, Francis Edgeworth and Karl Pearson would extend these theories with the concept of least-squares approximation to develop multivariate regression analysis which stands today as the lynchpin of statistical adjustment of outcomes data.

SUMMARY OF THE EVOLUTION OF HEALTHCARE QUALITY MEASUREMENT

The measurement of the delivery of health care can still be considered in its infancy even though the statistical tools with which reliable and valid measurements can be taken have been in existence for many years (Stigler, 1986; Iezzoni, 1997; Cohen, 1984). Resistance by physicians to the measurement of the quality of services has traditionally been met with a less than enthusiastic response as physicians have become accustomed to leading a professional life which affords autonomy second to none (Iezzoni, 1997; Zelman & Berenson, 1998). This rejection of "meddling" by outsiders, and economic reasons, have led to initiatives which use existing claims data to measure the quality of care in delivery systems, thereby reducing the burden on physicians to provide medical records and support for quality measurement activities (Iezzoni, 1997; National Committee for Quality Assurance, 2000).

Measurement of the Quality of Behavioral Health Care

THE CURRENT "STATE OF THE ART"

The measurement of the quality of behavioral health services delivered to managed care clients is a topic which is bathed in confusion and controversy (Geller, 2000). Nowhere is this controversy more evident than in the measurement of the quality of care of inpatient services, ostensibly because of the high cost of these services and their disproportionate effects on health care provider budgets. In a recent article, the cost of a bed and nursing services alone in a hospital cost $615 to $633 a day, whereas individual psychotherapy ranges from $98 (M.D.) to $55 (M.A. or L.C.S.W) per session (Fagan, Schmidt, & Cook, 2002).

Clearly there is a profound incentive for state governments to reduce inpatient admissions, reduce length of stay, and ensure that those who are admitted do not return quickly. In an attempt to address these costs, state policy makers have tended to shift more of the burden of inpatient care to private hospitals so that the federal government will pay more of the charges through the Medicaid program (Geller, 2000). Treatment in general hospitals is also more appealing to some clinicians and policy makers as it removes the locus of treatment from the state hospital setting which is viewed by many as stigmatizing (Okin, 1983).

Unfortunately, an unintended consequence both of the continuing move toward "dehospitalization" and the move toward capitated Medicaid managed care funding mechanisms is increased fragmentation of the care of individuals. This has resulted in the most chronically ill indigent being stuck in what Geller (2000) refers to as a revolving door mechanism whereby individuals who would have had long lengths of stay at state hospitals before dehospitalization now stay for shorter spells but return more often (Geller, 2000). Often the individuals must also travel for many miles in order to secure services because of the Medicaid managed care hospital network. Also, these individuals visit numerous hospitals where they have no “history”; this results in longer lengths of stay (Geller, 2000). Clearly the incentive is great to develop a robust methodology for evaluating the quality of inpatient services delivered to managed care clients. Although a de facto standard for measuring behavioral health access to care has been established, this author finds it less than compelling.

Unfortunately, the statistical advances that have benefited social science, clinical research, and innumerable other disciplines have somehow mostly eluded widespread application to the measurement of behavioral health outcomes in real-world retrospective studies. In fact, perhaps the largest organization that produces products to measure quality in health plans -- the National Committee for Quality Assurance (NCQA) -- measures quality much the way Nightingale and Farr did over one-hundred years ago, by the calculation of simple percentages. Only a small percentage of the physical health HEDIS measures employ substantive risk-adjustment, and none of the behavioral health HEDIS measures employ risk-adjustment, in fact. The Texas Health Care Information Council (THCIC) cautions consumers of its reports using Health Employer Data Information Set or HEDIS measures that the lack of risk adjustment in HEDIS measures may lead to erroneous results (THCIC, 2002). HEDIS has become the standard data collection instrument of managed care with fully 80 percent (77 million) of people enrolled in HMOs participating in HEDIS data collection through their health plans (NCQA, 2002). One of these HEDIS measures, Follow-up after Hospitalization for Behavioral Illness, is particularly relevant to this discussion for several reasons: 1) it is the most widely used method of appraising the quality of behavioral health care in both commercial and Medicaid settings (NCQA, 2002); 2) it incorporates no risk adjustment thus demonstrating the weaknesses of non-risk-adjusted approaches; and 3) this measure, along with the other HEDIS measures, is being promoted as the standard for quality measurement in managed care. Clearly most professionals would be somewhat concerned about the widespread adoption of an instrument that has questionable validity for the measurement of the traits it is marketed to measure; and it is somewhat surprising that more of an outcry hasn't been registered, especially considering evidence from methodologically sound studies which suggest that non-risk-adjusted indicators of behavioral health quality may be seriously flawed (Hendryx, Dyck, & Srebnik, 1999; Hendryx, Moore, Leeper, Reynolds, & Davis, 2001; Foster, 1999).

SUMMARY OF THE CURRENT STATE OF BEHAVIORAL HEALTH QUALITY MEASUREMENT

Healthcare maintenance organizations, states, and other entities at risk for paying the costs of inpatient behavioral health services are under pressure to find ways to treat persons in less expensive settings and shorten the lengths of stay in inpatient psychiatric facilities (Geller, 2000; Fagan et al., 2002). Some authors believe that when policy makers and business leaders attempt to save on inpatient psychiatric care what results is poorer behavioral health care and the continued fragmentation of the behavioral health delivery network (Geller, 2000). The rise of the health care maintenance organization (HMO) has led to an increased incentive to decrease inpatient behavioral health care (Geller, 2000; Fagan et al., 2002), and substitute this care with community alternatives, thereby continuing the move toward deinstitutionalization which started fifty years ago (Geller, 2000). Beyond a promise of less expensive care, HMOs also brought with them the promise of improved quality (Zelman & Berenson, 1998), and many attempts have been made to measure equitably behavioral health care (NCQA, 2002; Zelman & Berenson, 1998). Unfortunately, the de-facto standard of HMO behavioral health care quality, HEDIS, is not risk-adjusted which reduces its utility as a useful rubric of quality service delivery (Iezzoni, 1997; THCIC, 2002; Hendryx et al., 1999; Hendryx et al., 2001; Foster, 1999).

PROOF OF THE NEED FOR RISK ADJUSTMENT

In a 1999 article, Hendryx et al. investigated the impact of agency rankings when risk-adjustment methodology was employed to gauge the effectiveness of community behavioral health outpatient services. The authors constructed a survey packet which they administered to 289 outpatient clients. Maximum likelihood factor analysis was employed to reduce the multiple items in the survey packet to three theoretically supported dependent variables: functional status; quality of life; and satisfaction with services. The predictors of the dependent variables were sex, race, age, and presence of severe diagnosis, baseline levels of substance abuse, baseline functional status, baseline quality of life, and baseline satisfaction with services. The authors found that among the agencies compared in their analysis, there were sometimes profound differences in the performance of one agency versus another in the risk-adjusted versus non-risk-adjusted comparison schema. This article remains significant due to the authors' use of cross-validated multiple regression methodology as a risk adjustment tool. Unfortunately, this article does not directly address rehospitalization among psychiatric populations.

In another methodological article studying risk adjustment techniques, this time directly associated with rehospitalization rates, Hendryx et al. (2001) investigated the utility of stratification-weighting logistic regression without interaction effects and logistic regression with interaction effects among 1,616 subjects to determine if using these techniques changed the relative performance of the behavioral health agencies under study. The authors used a number of different variables as risk adjusters including, age, race, Hispanic/not-Hispanic, current residence, employment, primary diagnosis, secondary diagnosis, DSM-IV axis 4, DSM-IV axis 5, legal status, education, and diagnosis of serious behavioral illness. The authors conclude that risk adjustment does indeed change the relative ranking of agencies versus non-risk-adjusted rehospitalization rates. The authors go on to say that stratification-weighting risk-adjustment techniques are too difficult to implement, and recommend logistic regression for this sort of problem. Among the logistic regression models, they found that the agency rankings did not change when the authors compared the logistic regression interaction model results with those of the logistic regression model not incorporating interaction effect, although they stated that the interaction model was statistically superior. The authors found that predictors of rehospitalization in the main effects model were younger age, non-count-ordered involuntary admission, and presence of client economic problems, and black race. In the main effect plus interactions model the following were significant: client age, involuntary admission, and economic problems; the interaction of age, black race, and court order; the interaction age, economic problems, black race and court order. This article is probably one of the most methodologically sound of many of the articles due in part to the substantial sample size and use of multivariate statistics.

Measuring Psychiatric Rehospitalization

IS REHOSPITALIZATION A PROXY FOR QUALITY?

Besides the obvious methodological pitfalls of measuring service delivery without adjusting for nuisance variables, the very utility of measuring outpatient aftercare visits or inpatient rehospitalization as a proxy for behavioral health service delivery quality has been called into question by the work of several authors (Solomon & Doll, 1979; Lyons et al., 1997; Foster, 1999; Cuffel, Held, & Goldman, 2002). The quality of many of the empirical studies has also been called into question because of low sample sizes or lack or multivariate statistical techniques (Klinkenberg & Calsyn, 1996).

Solomon and Doll (1979) encouraged the field of recidivism research to consider that psychiatric recidivism is such a multifaceted phenomenon that it will probably always defy traditional attempts to find its determinants. The authors also emphatically discouraged the use of simple readmission percentages as indicators of system quality. The authors use a gatekeeper and pathway rubric (Polk-Walker et al., 1993; Gruber, 1982; Solomon & Doll, 1979) to review the literature on psychiatric recidivism. Pathway and gatekeeper factors discussed in the article are exhibited in Table 1.

It is worthwhile to note that never before nor since the work of Solomon and Doll (1979) have authors working in the realm of psychiatric recidivism seemingly even attempted to grasp the true complexity of the phenomenon like they did. Along with Solomon and Doll (1979) and other authors (Klinkenberg & Calsyn, 1996), this author has the opinion that many studies in the literature are based upon a group of variables that are readily at hand but may not necessarily capture the constellation of factors which determine psychiatric recidivism. In sum, authors writing in the area of psychiatric recidivism would do well to acquaint themselves with this work in order to avoid producing a study that obfuscates the true complexity of the recidivism phenomenon.

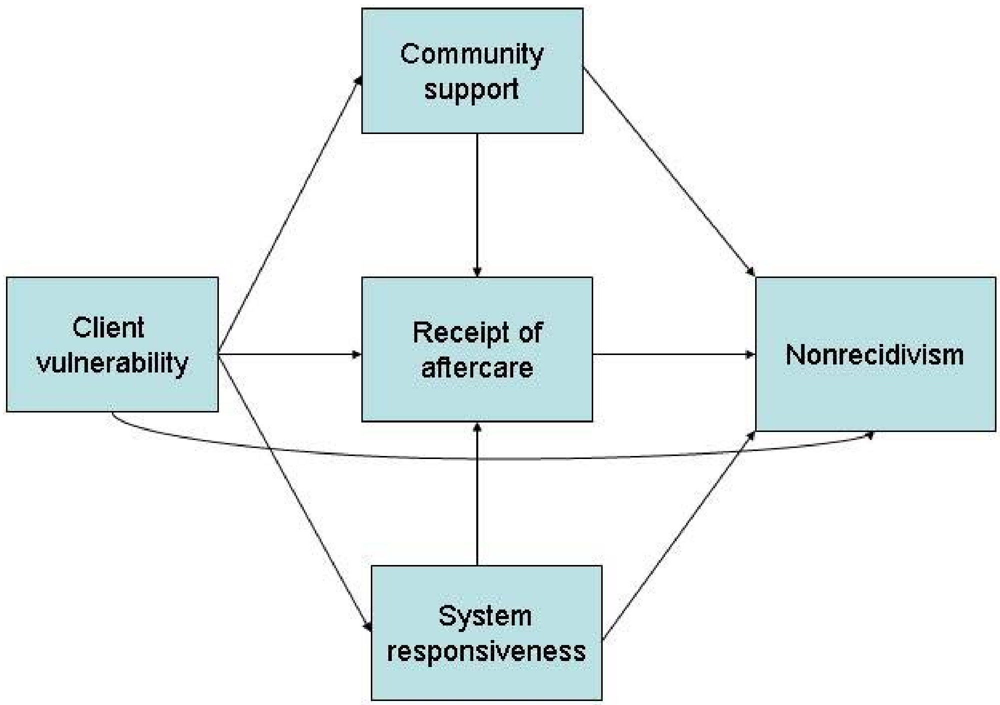
Table 1: Pathway and Gatekeeper Variables Leading to Rehospitalization

Pathway Variables	**Gatekeeper Variables**
Readmission as a patient's solution to:	**Readmission as a reaction to deviant behavior:**
1) Emotional Problems.	1) Behavioral symptoms.
2) The need for respite from daily pressures and stress.	2) Patient discomfort.
3) Family conflicts.	3) Social functioning.
4) The need to reduce stigma.	4) Familial and community pressures.
5) Financial problems.	
6) Housing problems.	
7) The need for companionship and recreation.	
8) The need to manipulate family.	
9) Protection from the weather.	
Readmission as a reflection of patient demographics:	**Readmission as a reflection of patient characteristics:**
1) Employment.	1) History of prior hospitalization.
2) Sex.	2) Demographics.
3) Social class.	3) Manipulability.
4) Race.	4) Likeability.
5) Marital status.	5) Accompanying individuals.
6) Age.	
Readmission as a reflection of variations in family tolerance by:	**Readmission as a reflection of the characteristics of admitting personnel:**
1) Demographics.	1) Clinical experience.
2) Nature of environmental tolerance.	2) Attitude toward community treatment.
3) Patient's financial or emotional role.	3) Awareness of community resources.
4) Solution to family conflicts.	4) Status within the hospital.
Readmission as a reflection of variations in community tolerance conditioned by:	**Readmission as a reflection of the nature of the behavioral health delivery system:**
1) Patient's visibility.	1) Hospital census.
2) Fulfillment of societal role.	2) Time/day of admission.
3) Attitudes of public and policy agencies.	3) Referral source to the hospital.
4) Attitudes and practices of referring sources.	4) Availability of community alternatives.
5) Behavioral health legislation.	5) Hospital policies.

TOWARD A MODEL OF PSYCHIATRIC REHOSPITALIZATION

In a comprehensive review of the literature published from January, 1974 to March of 1994 on both predictors of aftercare and psychiatric recidivism Klinkenberg and Calsyn (1996) advanced both a theoretical framework which complements that of several authors (Polk-Walker et al., 1993; Gruber, 1982; Solomon & Doll, 1979), and they also offered some criticism on the state of the literature in the recidivism field. The following figure illustrates the model of recidivism and aftercare predictors advanced by Klinkenberg and Calsyn (1996):

Figure 4: A Comprehensive Model of Variables Predicting Receipt of Aftercare and Rehospitalization of Persons with Severe and Persistent Behavioral Illness



Source: Klinkenberg & Calsyn, 1996, p. 488.

Klinkenberg and Calsyn (1996) also propose several functions which capture their model of recidivism and use of aftercare services. The likelihood of recidivism is captured by the following function: Recidvism = f(AC,CV,CS,SR) where AC is receipt of aftercare; CV is client vulnerability which captures such variables as low income, substance abuse, and previous psychiatric hospitalizations; CS is community support which captures variables related to the patient's living conditions and the support offered by his/her acquaintances; and SR is system responsiveness which captures formal actions taken by behavioral health programs in order to facilitate the recovery of the client. The authors note that no interactions are specified within the model, but interactions between client vulnerability variables, community support variables, and system responsiveness variables have an impact on recidivism through their effect on the aftercare component of the function.

Klinkenberg and Calsyn (1996) use their theoretical model as a framework to review the studies on recidivism and aftercare. Using this model, they find that system responsiveness variables are associated with recidivism more often than other types of variables. The authors also found that many of the community support variables were related to recidivism. Furthermore, the authors commented that the methodological character of many of the studies they found were less than satisfactory. Many of the studies used small samples that were not necessarily representative of the population. The authors also comment that a meta-analysis was impossible because the directionality, correlation, and regression coefficients were often not reported along with the p-values, making many studies of limited use. The authors propose that future researchers in the field of psychiatric recidivism make use of multivariate statistics to control for nuisance variables and explore interaction among the variables.

The present author believes that the Klinkenberg and Calsyn (1996) model serves as a useful heuristic for the further exploration of the subject of psychiatric recidivism. Thus the remaining portion of this review will be divided into those studies which find (regardless if it was the main focus of study or not) the main basis of psychiatric recidivism to fall within one of the spheres of influence outlined by Klinkenberg and Calsyn (1996): Aftercare, Client Vulnerability, Community Support, and System Resources. Only two of these variables are typically available for study in large-scale retrospective examinations of system performance using administrative data: aftercare and client vulnerability, thus the remainder of this review will center upon these elements.

A REVIEW OF THE PSYCHIATRIC REHOSPITALIZATION LITERATURE

Client Vulnerability

In a study by Polk-Walker et al. (1993), the authors built on the work of Solomon and Doll (1979) to explore the concept of pathway and gatekeeper variables which were advanced by these earlier authors. According to Polk-Walker et al. (1993):

> *Pathway variables, such as age, sex, social class, and number of dependents, are those that propel the potential patient toward the hospital. Gatekeeper variables, such as type of diagnosis, hospital admission policies, and patient's admission history, are those within the hospital environment.* (Polk-Walker et al., p. 163)

In this article (Polk-Walker et al., 1993), the authors use the pathway and gatekeeper variable framework to explore a number of potential predictors of readmission. The authors used chi-square tests to test the differences in frequencies between the group of patients with readmissions (n=233), and the group of patients without readmissions (n=83). The authors found significant differences between the groups on the following variables: race, marital status, marital history, number of children, residency of children, other admissions, and problems reported in the medical charts like marital, financial, social, work, sexual, impulse control, sleep, medical, central nervous system problems, compliance with medication, and compliance with treatment. The authors then entered the significant pathway variables into a log-linear model in order to control for the confounding effects of nuisance variables and attempted to discover the true drivers of readmission. The final model found that the residence of children dependent upon the subject, female sex, and history of previous admissions were the most predictive pathway variables. Among the gatekeeper variables entered into the model, only denial of existing financial problems, sexual problems, and impulse control problems remained in the model after the backward elimination procedure. The authors recommended that females, who have dependent children residing with another caregiver, have a history of previous admissions, and who have a documented history of denying financial, sexual, and impulse control problems should be targeted with special interventions following discharge to prevent readmission.

In a correlational study, Hoffman (1994) derides the literature for methodological shortcomings primarily in the area of the length that the patient is followed before readmission is established. He cites most studies as having a short 6-12 month follow-up period. In this study, Hoffman (1994) follows 50 schizophrenic patients for two years to establish if age, previous admission, and a number of other predictors were associated with rehospitalization in the two-year tracking period. The author found that previous admissions were the most influential predictor; age was a significant factor in readmissions and was also effective in reducing the influence of other variables when used in a partial correlation analysis. Hoffman (1994) found that age influences suicidal risk, antisocial behavior, and compliance to a prescribed drug regimen. The author also notes that he found no significant correlation between drug abuse and rehospitalization, or awareness of illness and rehospitalization. Besides the small number of subjects used in this sample, this author believes that the methodology could be improved by moving beyond spearman correlation and partial correlation analysis to a more suitable methodology such as logistic regression or survival analysis.

In a 1995 study, Haywood et al. used a small sample (n=135) of patients hospitalized with schizophrenic, schizoaffective, or affective disorder to attempt to find the determinants of multiple psychiatric readmissions. The authors used hierarchical regression to predict the log-transformation of the number of psychiatric visits the patients had experienced. The authors found the most important predictor of multiple hospitalizations to be the concomitant presence of an alcohol or drug problem followed by medication non-compliance. The full model developed by the authors accounted for 17% of the variance in the number of log-transformed visits observed among the patients. The authors were surprised to learn that schizophrenic patients did not have significantly more inpatient visits than did those with unipolar depression. Although this study did use multivariate statistics, it is still hampered by its small sample size and lack or representative sample reminiscent of the criticisms of the literature brought forth in the Klinkenberg and Calsyn (1996) review.

In a study of 255 psychiatric patients by Lyons et al. (1997), the authors question the utility of measuring the readmission rates as a proxy for psychiatric quality of care. The authors collected information from subjects using several instruments designed to measure the severity of psychiatric distress among psychiatric patients. They used a logistic regression model to predict the likelihood of readmission at one year. Among the variables studied, impairment in self care, severity of symptoms, suicide potential, and pre-morbid level of dysfunction were significant. The authors also found that the level of psychiatric illness based upon the Acuity of Psychiatric Illness scale did not differ at discharge between those readmitted within one year and those patients who were not readmitted. However, the authors did find patients who were readmitted were in more psychiatric stress at admission than those who were not. The authors also did not find any differences in length of stay at the index admission between those patients who were readmitted and those who were not, seemingly disproving a mantra of critics of managed care that early release promotes recidivism. Although this article is probably well founded in its assertion that readmission should not be studied as a proxy for quality, the present author believes that the study could have been much stronger if a range of aftercare variables could have been entered into the model. This would have helped separate the effects of severity of pathology among the patients from the effects of aftercare in the community.

In another 1997 study using a total of 1,575 Swiss patients, Huguelet explored the correlates of multiple psychiatric admissions to the only inpatient psychiatric facility in Geneva, Switzerland. The author used a database containing diagnostic information (including all five DSM-III-R axes) and demographic information. The author defined “revolving-door patients” as those persons who recorded three or more hospitalizations in a one-year period. The most compelling analysis in the article is a stepwise logistic regression in which the author found that patients with a diagnosis of psychosis were more likely to be hospitalized three or more times in one year if they had a co-morbid substance abuse disorder, longer duration of illness, were female, younger, and had suffered worse psychosocial adjustment in the preceding year. The author did mention that the study did not consider living conditions and other variables which have been shown to affect length of tenure in the community (Caton, Koh, Fleiss, Barrow, & Goldstein, 1985). To the credit of the author, the study used logistic regression which allowed statistical control of nuisance variables in the analysis, although inclusion of relevant aftercare variables would have improved the study.

In a methodologically sophisticated study, Foster (1999) used Cox regression to model time to rehospitalization in 204 children in the CHAMPUS (U.S. military health insurance) program at Fort Bragg. Foster (1999) found that individuals who received aftercare services were only 7 percent less likely to be readmitted than those who received no services after discharge from an inpatient setting before adjustment for predisposing factors and 18 percent less likely to be readmitted after adjustment for other factors. Both of these findings did not reach significance at .05. Interestingly, other variables primarily capturing the effects of personal characteristics such as gender and diagnosis did reach significance in the study: persons with a diagnosis of major depression were found to be almost two and a half times more likely to be readmitted than those who did not have this diagnosis; those diagnosed with oppositional defiant disorder were almost two times more likely to be readmitted; females were 18% more likely to be readmitted than males; and persons of white race were 30% more likely to be readmitted than non-whites. When Foster split out the four types of aftercare available in the system (outpatient therapy, case management, residential treatment center, and intermediate services) in an attempt to discern which types of aftercare were effective and which were not, none of the types of aftercare was statistically significant in the models, and only one (outpatient therapy) approached significance. In the end, Foster (1999) concludes that readmission, at least in his sample, is dependent upon personal characteristics.

Olfson et al. (1999) investigated the extent to which inpatient psychiatric readmissions within 3 months of the index admission were predicted among 262 patients. The authors studied the power of staff assessments on the chances of readmission as well as a host of demographic variables. The authors found that the most powerful predictor of readmission was four or more previous hospitalizations, followed by a diagnosis of co-morbid substance abuse, major depression, absence of a family meeting with inpatient staff, and the prescription of a conventional antipsychotic drug rather than a newer antipsychotic medication. The authors found that staff predictions of rehospitalization were little better than chance. Unfortunately, despite the large number of variables collected by these researchers, the study is weakened methodologically by the use of simple chi-square statistics to compare the non-rehospitalized vs. the rehospitalized groups. The use of multivariate approaches such as logistic regression would have enabled the researchers to control for nuisance variables within the specified models.

Another study, this time using a managed care sample of 370 patients as the locus of study (Moran, Doerfler, Scherz, & Lish, 2000), found that patients who were readmitted under managed care had previous admissions before the index admission, were unemployed, were recipients of Medicare or social security disability payments, and participated in day treatment. The authors also found that self-report symptomology was not significantly predictive of readmission. The authors point out in their discussion that even in systems designed to have continuity of care thus not delivering care in a fragmented manner, many patients return to the inpatient setting. The authors suggest a number of therapeutic approaches that they believe hold promise.

Craig, Fenning, Tannenberg-Karant, and Bromet (2000) also used a managed care sample (n=674) to explore rehospitalization. In their study the authors dichotomized their sample into patients who returned within 3 months of their first discharge, and those who were rehospitalized from 3 months to one year after their first discharge. These authors found that inadequate referral for aftercare services was not associated with less tenure in the community, and patients who fell into the group of patients who were rapidly readmitted tended to be symptomatic at their index discharge and have a diagnosis of affective psychosis without specific pharmacotherapy. Those who had greater tenure in the community tended to be compliant with their medication regimen after discharge. Perhaps this is another analysis that could have benefited from the use of survival analytic techniques instead of the simple dichotomization of readmission into 0-3 months and 3 months to one year. This article also depended upon a large number of chi-square analyses which, besides not controlling for other variables, also opens the possibility to false positive results due to type-1 error.

In another study using Cox regression, this time in an English cohort of patients, Hodgson, Lewis, and Boardman (2001) also found several personal characteristics to be predictive of readmission. Using the log-rank test, the authors found that single (unmarried) individuals had a greater chance of rehospitalization than married individuals. Persons with a diagnosis of affective psychosis had the greatest chance of readmission, with almost half (49%) being readmitted within the 5-year study window and those diagnosed with schizophrenia were found to be readmitted 42% of the time. Patients diagnosed with personality disorder had a readmission rate of 39%. Patients who had a length of stay longer than the median time were also found to be significantly more likely to be readmitted, as were those who were involuntarily placed under the English Behavioral Health Act. The authors also found that age and sex did not significantly affect readmission when tested using the log-ranks methodology. When the authors employed a forward stepwise Cox regression methodology, only four variables were retained in the model: gender, composite diagnosis (psychotic vs. neurotic), marital status, and behavioral health section (voluntary or involuntary). Of these variables, diagnosis was the most predictive of readmission, with psychotic patients having a 31% greater chance of readmission vs. neurotic patients. Unfortunately, outpatient services were not studied as a predictive factor in this article.

Aftercare

Solomon, Gordon, and Davis (1984) completed a discriminant analysis study involving 550 state hospital patients which were tracked for one year after initial discharge from a state hospital. The authors enlisted the help of a variety of agencies in order to build what they believed to be a comprehensive database of services received by the subjects. The authors found that patients who were not readmitted to a state hospital received a greater variety of aftercare services than those who were readmitted, although there was no difference found between readmission rates between those who received aftercare and those who did not. The authors concluded that this finding supports a more comprehensive aftercare model involving social and vocational rehabilitative services. Interestingly, the authors did not find significant differences in any demographic variables between those who were admitted and those who were not readmitted. The only variable associated with the client they found to be significant was the number of prior admissions; the authors attributed this finding as a variable which was in control of the practitioner. Although compelling at first, when the types of variables that are under study in this paper are examined further it appears that some critical variables may be missing. Variables serving as a proxy for morbidity are curiously absent, creating likely confounds in the authors' results.

In another early study using a small sample (n=119) of chronic schizophrenics, Caton et al. (1985) established links between discharge planning, community treatment compliance, interpersonal stress, and time in the community before rehospitalization. The authors also developed a mathematical model to predict the number of days that a patient was expected to be in the hospital. The unusual thing about this paper is that the authors use life-table analysis to model the time from index discharge to either: a) readmission to a psychiatric hospital; b) loss to follow-up; or c) an elapsed time of one year since the last readmission. This technique allows the authors to exploit the additional information offered by the variable of time to readmission instead of taking the simplistic approach of modeling a binary outcome of readmission/no-readmission with logistic regression or worse – Ordinary Least Squares (OLS) regression. The authors' life-table estimates yielded a 58% probability of readmission within the one-year period of surveillance. They also found a 24% probability of readmission in the first month after discharge. Significant differences in community tenure were found between those subjects who were compliant with their post-discharge plans, those patients who were released to high or low-stress community settings, and those who were rated by hospital staff as having a good or poor prognosis on the

Discharge Planning Schedule assessment instrument. This study stands out as it is one of the only studies in the field to use survival analysis techniques to fully exploit the information offered by longitudinal data. The authors could have exploited the data more if they had applied a Cox regression methodology, thereby allowing them to study numerous exploratory variables simultaneously instead of just different levels of the same variable, although it is somewhat debatable how accessible Cox regression techniques were to many researchers in the mid-1980s.

An Australian study by Owen, Rutherford, Jones, Tennant and Smallman (1997) attempted to link the 6-month rehospitalization experiences of 128 patients. The study included a logistic regression analysis which attempted to predict rehospitalization using standard patient demographic variables, clinical condition, and attitude toward follow-up after discharge, diagnosis, aftercare plans of the patient, and an unusual measure purported to capture patient likeability. The authors found that readmitted patients had more financial problems, difficulties with hygiene, and fewer leisure pursuits. The authors found no relationship between demographic factors, diagnosis, functioning at discharge, attitude of the patient to post-discharge follow-up, or patient likeability to readmission. Interestingly, the authors used a stepwise method only on those variables which were significant in a bivariate logistic regression equation. This may have prevented them from exploiting the full power of logistic regression to control for nuisance variables, allowing those variables which are truly related to the dependent variable to be identified.

Song, Biegel, and Johnsen (1998) completed a large sample, multi-year panel design which studied both the rates of predictors of rehospitalization as well as the predictors of the length of psychiatric hospitalization. The authors used logistic regression to model the probability of rehospitalization across three different consecutive years. In each of the three years, the authors found that the risk of rehospitalization in the current year was negatively related to the number of community psychiatric service hours consumed in the previous year. The authors also found that people who used larger numbers of hours in the current year tended to have higher rates of rehospitalization. A relationship was also found between psychiatric hospitalization in the previous year and rehospitalization in the current year.

Of these three variables, the number of hours of psychiatric community services consumed in the present year was found to have the strongest relationship with rehospitalization, controlling for the effects of hospitalization in the previous year and the number of hours of community psychiatric contact in the previous year. The authors also found that African-Americans were more likely to be rehospitalized the same year, as were schizophrenics, persons with a dual-diagnosis of substance abuse and a major behavioral illness, and persons who used a greater number of services measured in days (e.g., partial hospitalization, residential treatment, etc.). The authors note in their discussion that their study supports the literature which shows that prior hospitalization is the best predictor of future hospitalization. The authors state that they did not find a consistent predictor of length of stay over all their years of the study, although a diagnosis of schizophrenia, affective psychosis, or a dual diagnosis was related to longer length of stay in one year.

This article is significant in that it uses logistic regression to control for nuisance variables, has a large enough sample size to have adequate power to detect the hypothesized differences where they exist, and, probably most significant, uses four years of data which allow for three distinct analyses to take place; in a sense, the study is self-replicating. Of concern in this study is the authors' lack of separation of community service visits into visits that come before or after the hospitalization in question. This feature makes it difficult for the reader to discern if a patient uses more community services after the hospitalization in the current year as a result of follow-up release planning from the releasing institution, or if the subject became more distressed, sought more and more community services which finally culminated in a hospital inpatient stay.

The authors' other finding regarding community services, that previous use of services is inversely proportional to the probability of being hospitalized in the current year, appears solid and seems to be consistent with a model in which the individual decompensates, loses touch with the community behavioral health support structure, and eventually becomes acutely ill. It is unfortunate that these authors did not have prescription drug usage data to bolster this already strong study. It would have been interesting to note if the number of prescriptions for neuroleptics filled in the previous year was also inversely proportional with the likelihood of being rehospitalized.

A study by Nelson, Maruish and Axler (2000) may support the contention by some previous authors that the sample sizes of their studies were too small to find significant effects of aftercare on psychiatric readmission. In this article, the authors found that among their 542-subject sample, patients who did not keep at least one outpatient appointment were at twice the risk of readmission than those subjects who attended at least one outpatient appointment. The authors also found that those patients not attending at least one appointment had a change of readmission which increased with time, increasing from 15% to 29%; whereas those who attended at least one appointment had a constant risk of readmission. Unfortunately, this study did not describe controlling variables which were used along with the aftercare measure in the authors' model. Therefore, it is entirely possible that what these authors detected is the effect of personal characteristics such as race or gender along with diagnostic and morbidity characteristics masquerading as an aftercare effect. The authors did emphasize that they believe it to be important to track outpatient care beyond the industry standard (read HEDIS standard) of 30 days, stating that the 30-day measure may capture the quality of the hospitalization while more protracted periods of measurement would be more apt to measure the quality of the aftercare intervention provided.

In an interesting twist on the exploration of the correlates of psychiatric re-hospitalization, Dausey, Rosenbeck, and Lehman (2002) studied the effect of preadmission visits upon the length of stay in Veterans Administration psychiatric inpatient facilities. In this massive study (n=37,852), the authors found that patients experiencing at least one preadmission visit had shorter lengths of stay and tended to keep appointments with aftercare providers more often than did those subjects who did not get preadmission care. The authors also found that the number of preadmission visits did not seem to matter more than the presence or absence of preadmission care.

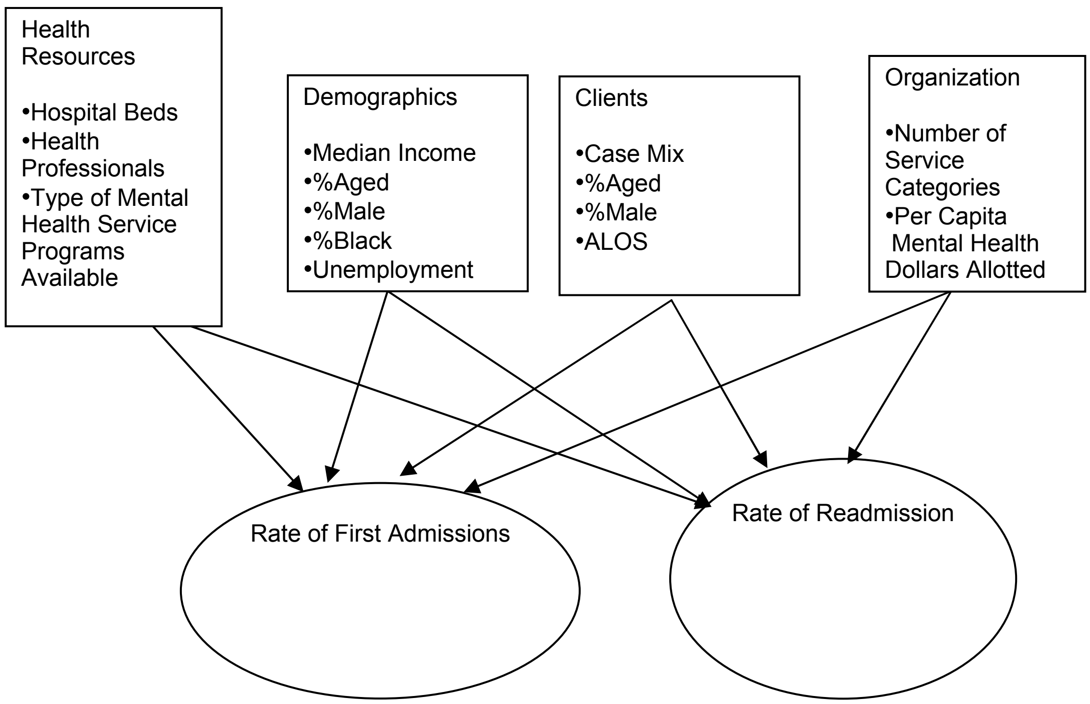
In another recent study, this time with a moderately-sized (n=391) Behavioral Health Organization (BHO) sample, Cuffel, Held and Goldman (2002) explored the relative effects of patient characteristics (such as gender) and system characteristics (such as intensity of post-discharge intervention) using logistic regression models. These authors found no decrease in the likelihood of readmission at 30, 60, and 180 days as a result of care received after discharge. Furthermore, they found that the intensity of aftercare delivery had no effect on readmission. Unfortunately, this study was somewhat limited by a moderate sample size and the lack of the use of Cox regression which may have more appropriately exploited the information available to the investigators.

In another recent article, Frederick, Caldwell and Rubio (2002) tested the hypothesis that home-based aftercare would decrease psychiatric re-hospitalization in a Medicaid managed care setting. The authors of this study used logistic regression to model the likelihood of readmission at 90 days. The authors found that home-based treatment, age, gender, and previous hospitalization were not good predictors of inpatient rehospitalization. The authors cite the small sample size (n=270) used as an impediment to finding significant relationships among re-hospitalization and the independent variable. This study also could have been improved by the use of more sophisticated analytic techniques. Many of the hypothesis tests depended upon simple uncontrolled binomial tests which left ample room for confounding by nuisance variables.

System Responsiveness

The only study classified into a system responsiveness framework in this review also is the only study which used individual components of a state community behavioral health system as the unit of analysis. To complete this study, Wan and Ozcan (1991) proposed the model in Figure 5 and used structural equation modeling to estimate the fit of the model using 144 Community Services Boards (CSBs) from Virginia as subjects. As displayed in their model, four dimensions affect hospitalization and rehospitalization: Health resources, defined as the resources within the CSB; Demographics, defined as the demographic makeup of the CSB catchment area; the Clients dimension captures the case mix and other characteristics of CSB clients in psychiatric hospitals; and Organization, which measures the financial health and complexity of the organization. All of these dimensions were postulated by the authors to influence both admission and readmission with rate of first admission also influencing the rate of subsequent admission.

Figure 5: An Analytical Model of Psychiatric Rehospitalization of CSB clients



Source: Adapted from Wan & Ozcan , 1991, p. 7.

The authors estimated two models using LISREL, one for admission rates, and one for readmission rates. The first equation found that the predictors accounted for 51.5 percent of the variance in readmission rates, length of stay and median income were inversely related to admission rates; whereas the other predictors were positively related. The authors' second equation accounted for 62% of the variance in readmissions although the majority of the variance was accounted for by admission rates. The authors found that average length of stay was inversely related to readmission, while the number of service categories offered and dollars per capita were positively related to readmission.

Community Support

In another design using areas instead of individuals as the unit of analysis, Turner and Wan (1993) explored the role community factors play in the rehospitalization of psychiatric patients in Virginia using data from 95 counties and six cities; 590 subjects were aggregated for analysis. A structural equation model was employed to test the effect of the six latent variables: Resource, comprised of the resources available at the psychiatric institution; Social, comprised of the socioeconomic characteristics of an observation’s catchment area; Behavioral Health, comprised of number of psychiatric patients per 1,000 persons in the area; Chronicity, defined as the number of seriously behaviorally ill persons per 1,000; Race, defined as the percentage of African Americans in the area; and Household, defined as the percentage of female-headed households in the area. The authors found that household composition and socioeconomic status (Social) were found to be significantly related to rehospitalization. The percentage of female-headed households was positively related, while socioeconomic status of the area was negatively related to rehospitalization.

Summary of the Rehospitalization Literature

Hendryx et al. (1999) and Hendryx et al. (2001) found that behavioral health quality measures that are derived without controlling for the effects of nuisance variables result in spurious rankings of agencies. Solomon and Doll (1979) and Lyons et al. (1997) both gave excellent reasons why the rehospitalization rate is not a good proxy for quality in the measurement of behavioral health care. With these issues in mind, what are the nuisance variables that need to be controlled for in rehospitalization studies? Klinkenberg and Calsyn (1996) found in their review article that previous hospitalization was the best indicator of future hospitalization of the client vulnerability variables studied by researchers. This finding was replicated by several other authors (Hoffman, 1994; Olfson et al., 1999; Song et al., 1998). Several authors using disparate samples found that co-morbid substance abuse was a significant predictor of rehospitalization (Haywood et al., 1995; Huguelet, 1997; Olfson et al., 1999). A number of authors pointed out that patients with certain diagnoses were more likely to encounter rehospitalization (Lyons et al., 1997; Olfson et al., 1999; Craig et al., 2000; Hodgson et al., 2001; Song et al., 1998). Some authors found a relationship between female sex and rehospitalization (Polk-Walker et al., 1993; Huguelet, 1997; Hodgson et al., 2001; Turner & Wan, 1993), while others found no relationship between sex and rehospitalization (Frederick et al., 2002; Owen et al., 1997; Foster, 1999). Non-compliance with medication was found to be a significant predictor of readmission by some authors (Haywood et al., 1995; Craig et al., 2000), while others found that the patients receiving pharmacotherapy with conventional antipsychotics were at greater risk for readmission (Olfson et al., 1999). No relationship was found between receipt of post-discharge outpatient services and rehospitalization by some authors (Foster, 1999; Frederick et al., 2002; Solomon et al., 1984; Cuffel et al., 2002), while others found that aftercare either lengthened tenure in the community (Nelson et al., 2000) or reduced the readmission rate (Caton et al., 1985). The length of stay at the index admission was not found to be related to rehospitalization in one article (Lyons et al., 1997) while the authors of another article state that length of stay is inversely related to the proportion of rehospitalization in an area (Wan & Ozcan, 1991). Interestingly, one article pointed out that it may be the variety of aftercare services delivered that reduces readmission (Solomon et al., 1984).

In their 1996 review, Klinkenberg and Calsyn judged many of the studies on rehospitalization as methodologically poor due to the lack of use of multivariate statistics. More recently, it appears that many authors are embracing logistic regression (Lyons et al., 1997; Huguelet, 1997; Moran et al., 2000; Owen et al., 1997; Song et al., 1998; Cuffel et al., 2002). Others used structural equation (LISREL) modeling (Wan & Ozcan, 1991; Turner & Wan, 1993), while others used what this author believes to be the preferred methodology, Cox regression (Foster, 1999; Hodgson & Lewis, 2001). Unfortunately some authors still are embracing the use of often numerous bi-variate chi-square analyses raising the risk of type-1 error (Olfson et al., 1999; Craig et al., 2000; Frederick et al., 2002).

In summary, it appears that the number of previous hospitalizations is the most reliable predictor of rehospitalization, followed by the severity of diagnosis, co-morbid substance abuse, and female gender. The role of aftercare is mixed. Aftercare is likely to have some influence on rehospitalization rates, but its effect may be expressed through the variety of care delivered, instead of simply a binomial indicating if aftercare was delivered or not. In what might be viewed as a fortuitous finding for the financial statements of many Behavioral Health Organizations (BHOs) and HMOs, a short length of stay at index admission does not seem to have much empirical support as being related to rehospitalization.

UNANSWERED EMPIRICAL QUESTIONS

What is lacking from the literature on psychiatric rehospitalization are: 1) a consistent use of a methodology which makes use of the maximum amount of information available while also accounting for censored observations; 2) a useful proxy for psychiatric morbidity, or severity of illness, which can be used as a risk-adjustment covariate in rehospitalization models; 3) some statistical method for using the data available from multiple rehospitalizations simultaneously in one procedure thereby using all the information available; and 4) a comparison of the relative merits of the various statistical techniques available to explore the psychiatric rehospitalization problem.

PROPOSED RESEARCH

Hypotheses to be tested in this dissertation are:

1) A psychiatric hospitalization in a previous year will be positively related with future psychiatric hospitalization and negatively related with community tenure. The rationale for

this hypothesis is derived from the conclusions of the comprehensive review article by Klinkenberg and Calsyn (1996), who found past rehospitalization to be a strong predictor of future rehospitalization. The findings of these authors are also supported by a number of other studies (Hoffman, 1994; Olfson et al., 1999; Song et al., 1998).

2) The demographic variables of sex, age, and race will influence community tenure and rehospitalization rates. Several authors have found relationships between sex and rehospitalization (Polk-Walker et al., 1993; Huguelet, 1997; Hodgson & Lewis, 2001), age and rehospitalization (Hoffman, 1994; Huguelet, 1997) and race and rehospitalization (Song et al., 1998.)

3) A significant negative relationship will be found between the variety of services consumed by an individual in the year prior to the study year, and their rate of rehospitalization. A significant negative relationship will also be found between the number of different services received in the year prior to the study year and the mean hazard of rehospitalization. The rationale for this hypothesis stems from several authors (Wan & Ozcan, 1991; Solomon, Gordon & Davis, 1984) who propose that a more comprehensive outpatient treatment model represented by many available services has a better chance of success.

4) Severity of psychiatric illness will prove to be positively related with the rate of rehospitalization and negatively related to community tenure. Wide support exists in the literature for a relationship between rehospitalization and the existence of schizophrenia, affective psychosis, or major depression (Lyons et al., 1997; Olfson et al., 1999; Craig et al., 2000; Hodgson & Lewis, 2001; Song et al., 1998).

5) The type of new-generation antipsychotic medication the subject is taking will influence community tenure and rates of rehospitalization among patients prescribed a new-generation antipsychotic medication. The number of prescription days filled will be positively related to community tenure and negatively related to rehospitalization. Medication non-compliance was cited by several authors (Haywood et al., 1995; Craig et al., 2000) as increasing rehospitalization rates, while one author (Olfson et al., 1999) stated that conventional antipsychotics were associated with higher rehospitalization rates.

CHAPTER 3: METHODS

PARTICIPANTS

The sample for this study included 558 persons who were enrolled in and received mental health related services under Medicaid managed care in SFY 2000 (9/01/1999 – 8/31/2000) and 2001 (9/01/2000 – 8/31/2001). Persons were selected for the study if they were between the ages of 18 and 65 by the end of fiscal year 2001. Subjects were required to be in a Medicaid-only eligibility category so all their claims would theoretically be captured by the Medicaid claims processing system and no claims would be diverted to Medicare processing. Subjects were drawn from a database comprised entirely of persons who qualified for Medicaid due to chronic illness. Subjects were also required to have an inpatient behavioral health visit in SFY 2001. Subjects were qualified for the study if the first three digits of their primary inpatient diagnosis had an International Classification of Diseases 9th Revision code between 290 and 316. Persons with a diagnosis of dementia (290) were excluded from the sample. Slightly more females (n=312, 55.91%) than males (n=245, 44.09%) were included. About half of the subjects were African American (n=277, 49.64%), 146 of the subjects were white (26.16%); 82 were of unknown or mixed descent (14.7%); 42 (7.53%) were Hispanic; 9 (1.6%) were "other," and two (.4%) were Asian. Almost half of the subjects suffered from schizophrenia. The subjects are displayed in broad diagnostic categories in Table 2. As displayed in Table 3, the mean age of subjects was almost 43 years old.

Table 2: Diagnostic Classification of Subjects at First SFY 2001 Admission

Diagnosis	n	%
Schizophrenia	228	40.86
Bipolar Disorder	89	15.95
Depression	78	13.98
Chemical Dependence	65	11.65
Inorganic Psychosis	52	9.32
Neurotic Disorder	14	2.51
Affective Psychosis	10	1.79
Adjustment Reaction	5	0.9
Organic Psychosis	5	0.9
Special Syndromes	5	0.9
Alcoholic Psychosis	4	0.72
Schizoid Disorder	2	0.36
Conduct Disorder	1	0.18
Total Sample Size =	558	

Table 3: Age of Subjects (End of SFY 2001)

Age	
MIN	18
MAX	65
MEAN	42.88
STD	10.55
Total Sample Size =	558

MEASURES

Chronic Disability Payment System (CDPS)

The Chronic Disability Payment System is a computer program written in the SAS programming language designed to produce diagnosis adjusted risk scores for use in setting payment amounts for Medicaid managed care programs. The program replaces the Disability Payment System (DPS) developed by the CDPS authors and boasts a number of improvements over the DPS such as the use of four million beneficiaries in the development of the program weights, extensive use of physician consultation in the development of the diagnostic categories, as well as better prediction of future expenses than the DPS offered (Kronick, Gilmer, Dreyfus, and Lee, 2000).

The CDPS system works by computing the linear regression of a number of binary indicator variables coding for the presence or absence of fifty-six different diagnostic conditions as well as age and sex (and eleven significant interactions) on the standardized expenses of a patient in a year. The resulting regression weights, calculated on the data supplied from four million beneficiaries, are supplied with the CDPS software. The user has the option to rerun these regressions, thereby calculating weights customized to the population in question. This step is particularly crucial if the user is applying the CDPS to a commercial population. If the regression weights are recalculated for a commercial population, the R-squared statistic of expected vs. predicted expenditures increases from .125 (with no dollar truncation, in prospective mode) to .186 (with recalibration of regression weights) according to a study by

Cumming, Knutson, Cameron, and Derrick (2002). In the present case, the regression weights supplied with the software package were used as these weights were developed on a Medicaid population and thus should be appropriate for the current sample of Medicaid beneficiaries.

In order to determine the CDPS predicted score, the CDPS software simply adds the intercept term and regression weights which are indicated by each subject's diagnostic classifications as well as the weights indicated for the subject's age and sex (Kronick, Gilmer, Dreyfus, and Lee, 2000). The resulting score was then used in this study as the cdpspred variable.

PROCEDURES

Description of the Data

Medicaid managed care claims, rehabilitation and day treatment claims from the Texas Department of Mental Health and Mental Retardation (TDMHMR), Medicaid eligibility data, and drug utilization data were extracted from state databases after the determination was made that this research did not require human subjects' review due to the lack of identifying information on the data. The research director of the author's organization in 2002 gave his permission for the extraction of the data. The base claims table from which the sample was drawn contains 5,404,421 rows, the eligibility table contains 2,606,906 rows, and the drug table contains 558,598 rows. The pharmacy data were obtained from the University of Texas Center for Pharmacoeconomic Studies. The variables available for analysis are displayed in Table 4 for the pharmacy data, Table 5 for the enrollment data, and Table 6 for the claims data.

Table 4: Description of Variables in Pharmacy Data

Variable	Description
amt_paid	Amount state paid for prescription
bn	Name of drug
category	Classification of drug by Pharmacoeconomics staff
date_disp	Date prescription was filled
pcn	Unique id of patient
days	Number of days supply

Table 5: Description of Variables in Enrollment Data

Variable	Description
Plancode	Code for managed care organization
bthdat	Date of birth of client
enrlfrom	Enrollment start date
enrlto	Enrollment end date
membno	Unique patient id
racecod	Race of patient
riskcod	Medicaid eligibility category
sexcod	Sex of patient

Table 6: Description of Variables in Encounter (Claims) Data

Variable	Description
membno	Unique patient id
Admtdate	Date of inpatient admission
Dschdate	Date of inpatient discharge
Admtdiag	Initial diagnosis at admission
Adntsrc	Source of admission
admttype	Type of admission
dfrodos	Start date of service
dtodos	End date of service
diagn1	Primary diagnosis (ICD-9-CM)
diagn2	Second diagnosis (ICD-9-CM)
diagn3	Third diagnosis (ICD-9-CM)
diagn4	Fourth diagnosis (ICD-9-CM)
diagn5	Fifth diagnosis (ICD-9-CM)
dschstat	Status of patient at discharge
plancode	Code for managed care organization
poscod	Code identifying where service was rendered
formcod	Code identifying if claim is from UB92 or CMS1500 form
svccod	Code identifying the service performed (CPT-4)
toscod	Type of service code

Table 7: Description of the Derived Variables

Variable	Description
diff1	Time in days from index discharge to rehospitalization or censoring
status	Dichotomous variable indicating the presence of a psychiatric rehospitalization after the index discharge during the study year
prev_hosp	Dichotomous variable indicating a hospitalization in the year prior to the study year for a psychiatric diagnosis. 1 = yes / 0 = no
sex	Sex of patient 1 = male / 0 = female

Variable	Description
age	Age in years as of 31st of December 2001
white	Indicator of white race 1 = yes / 2 = no
variety00	The number of types of psychiatric services received by the patient in the year prior to the study year
numvis00	The number of visits received by the patient in the year prior to the study year
cdpspred	The CDPS predicted relative risk score calculated only on claims which have a psychiatric primary diagnosis
maxdose	The maximum number of days supply of psychiatric medication during the study year
days	The time-dependent covariate capturing the cumulative days supply in the week prior to readmission or censoring. Note: this variable is calculated by the PHREG procedure and is not available in the dataset.
schiz	Dichotomous variable indicating the presence or absence of schizophrenia as a primary diagnosis (1 = schizophrenia / 0 = no schizophrenia)
psyh	Indicator capturing the presence of a high-cost psychiatric diagnosis (1 = yes / 0 = no)
psym	Indicator capturing the presence of a medium-cost psychiatric diagnosis (1 = yes / 0 = no)
psyl	Indicator capturing the presence of a low-cost psychiatric diagnosis (1 = yes / 0 = no)
subl	Indicator capturing the presence of a low-cost substance abuse diagnosis (1 = yes / 0 = no)
cdrug1	Indicator variable indicating Risperdal use and contrasting with Zyprexa
cdrug2	Indicator variable indicating Seroquel use and contrasting with Zyprexa
los1	The inpatient length of stay of the subject at index admission
numcats	The number of CDPS categories into which the subject fell
plancode	The managed care organization of the client at index admission

All data were processed and stored in SAS® v8.2 (SAS, 2001) using a 1.9 GHZ processor computer with 512MB of RAM, an 80GB raid 0 array, and a 40 GB primary drive.

Data Preparation Steps

As the analysis files all contain many records per person, and most of the statistical techniques used in this analysis require a file with one record per person, a number of programming statements had to be submitted in order to merge and condense the information so that it would be usable in the study. These steps are outlined in the following section.

1) Selection of subjects based on diagnosis: Subjects were selected who had an inpatient behavioral health claim with a primary diagnosis between 291 and 316 of the International Classification of Diseases 9th Revision during SFY2001. Selections were made on the first three digits of the primary diagnosis code (diagn1). Subjects who had a length of stay of less than one day were deleted from the analysis set.
2) Transposition of data: The data set was then transposed to derive the following variables: Length of stay (los) is a series of variables capturing the length of stay of each subject at a particular inpatient event. Variables are indexed i = 1 to number of inpatient stays; admit which is a series of variables indexed i = 1 to number of inpatient stays – 1. This variable captures the number of days between each inpatient event. Admit1 is null if no readmission occurred in the particular patient.

3) Condensing Enrollment: The enrollment segments (the date of enrollment with the subject's managed care organization, as well as the end of the subject's enrollment) were reformatted so contiguous enrollment segments would be in one record if a break in enrollment did not exceed two months. Subjects from step one were merged with these data and only subjects who had an inpatient visit while they were eligible for services at the time of inpatient admission were retained. Subjects were excluded if they had hospitalizations reported under multiple HMOs during this step.

4) Application of Chronic Disability Payment System to SFY2000 encounter file: The Chronic Disability Payment System (CDPS) (2002) is a risk-adjustment tool which is designed to predict the cost of services to be delivered to chronically ill Medicaid beneficiaries. The most common use of a tool like the CDPS is in setting the payment per person or capitation rate for managed care plans. The CDPS can also be used to group patients into classes of illness based upon ICD-9 diagnoses. The CDPS was applied to SFY2000 claims data in this study to obtain variables capturing subjects' predicted level of morbidity in SFY2001 as well as their diagnostic classification(s) in SFY2000.

5) Merging of subjects with CDPS data: Subjects qualified for the study were merged with the CDPS prepared data in order to have variables representing their diagnostic history in the past year. A number of

subjects were lost during this step due to the fact that not all subjects had a history spanning two fiscal years. It is worth noting that all available subjects (not just those with a behavioral health diagnosis) were in the CDPS prepared SFY2000 data.

6) Merging of the data from step four with outpatient visits: The observations resulting from the merging with the CDPS data were then merged with the file containing the outpatient visits. In this data set, four variables were created which measure outpatient utilization in SFY2000 and SFY2001. The first variable num_vis00 captures the number of outpatient visits the patient had in SFY2000, the second variable num_vis01 captures the visits in SFY2001, variety00 captures the number of distinct types of services (e.g., counseling, rehabilitation, etc.) the individual received in SFY2000, and variety01 captures the distinct types of services received in SFY2001.

7) Merging of the data with pharmacy data file: In this step the data from step 6 were merged with the pharmacy data. Four variables were created, all with binary values: ng00 captures if the client had filled a prescription for a new-generation antipsychotic drug (e.g., Risperdal) in SFY2000; ng01 captures this for SFY2001; numdays00 captures the number of days supply of psychotropic prescriptions filled by the individual in SFY2000; numdays01 captures this quantity for SFY2001. It is assumed that the numdays00 and numdays01 variables are a proxy for medication compliance. The

medication file was unduplicated by subject id number, date the prescription was filled, and medication name. In order not to inflate the number of days supply attributed to a subject who received two different doses of the same medication on the same day, the number of days supply was averaged for that medication on the prescription day.

8) Merging multiple inpatient enrollments with later admission dates which occur before the previous admissions discharge date: Discharge dates which occur one day previous to a subsequent admission were also merged as these discharges may be due to administrative purposes where the subject is discharged and then immediately readmitted in order to continue funding for the subject's care.

9) Creation of longitudinal data set for Cox regression analysis: In this step, some of the data were reformatted from a multiple observation per subject format to a multiple variable per subject format in order to facilitate analysis via Cox regression using time-dependent covariates.

Data Analysis

The proposed hypotheses were tested using Cox regression and logistic regression techniques. Cox regression is a technique that is frequently used in biometry and medical research applications. Cox regression has also found applications in sociology often under the moniker of "event history analysis" and in engineering under “failure-time analysis” or “reliability analysis” (Allison, 1984). The essential features of Cox regression that are attractive to many researchers is that the technique allows for the unbiased analysis of time to event data controlling for covariates, it allows for time-dependent covariates, and it does not require the researcher to specify the shape of the underlying distribution. By analyzing time to event data instead of simply the probability of an event occurring vs. not-occurring (like logistic regression), a researcher is able to use all of the information available instead of discarding the information as would be the case in a logistic regression analysis. Cox regression also proves superior to ordinary least square regression (OLS) in that the Cox regression algorithm allows for censoring of persons who discontinued or did not experience the event during the study period. Not taking censoring or time-dependent covariates into account can bias the results of the analysis. For a more complete explanation of the difficulties of using ordinary regression techniques with longitudinal data, see Allison (1984).

One of the most frequent uses for Cox regression is to analyze the survival time for patients with a serious disease such as cancer with one of the covariates representing different treatments. To implement a Cox regression model on this type of problem, the researcher collects the dates that the treatment was applied and the date of death or the end of the study period when observation of the patient concluded. The difference is then taken between the start of observation and the end of observation. Another variable “death” is also constructed which captures if the event occurred for that individual; death in this case would be scored 1 while patients who dropped out of the trial early or completed the observation period without dying would be scored 0. Cox regression can handle both nominal data and ratio or interval level covariates. In the present example, two covariates are constructed: age, which is the age of the person at the conclusion of the study; and treatment, which is scored 1 if the patient was treated with a new therapy and 0 if the patient was treated with the standard therapeutic approach. The model when implemented using SAS® PHREG looks like this:

```
proc phreg data = cancer;
Model survtime*death(0) = age treatment/ ties = efron;
Run;
```

If the model includes a dummy coded covariate with more than two classifications, a test statement must be used to determine if all of the covariates are zero. In this case, there are three treatment categories which can be described by creating two dummy variables, treat1 and treat2.

```
Proc phreg data = cancer_mult;
Model survtime*death(0) = age treat1 treat2  / ties = efron;
Run;
```

The hypothetical cancer data set has the following structure, as displayed in Table 8:

Table 8: Example of Survival Time Dataset

obs	survtime	death	age	treatment
1	110	0	44	0
2	185	1	35	1
3	188	0	33	0
4	220	1	29	1
5	345	1	45	1
45	620	1	39	1

The statement proc PHREG data = cancer asks the SAS® system to activate the phreg procedure using the cancer dataset. The model statement asks the procedure to estimate the survival (survtime) with the censoring value of death set to zero, meaning the subject did not die while being observed for this study and either quit the study early or lived longer than the observation period. The survival time is estimated controlling for the covariates age and treatment. The ties = efron statement asks the procedure to resolve tied data by assuming that ties among the subjects in the value of the survtime variable are due to imprecise measurement. The efron method performs better than the default Breslow method, but not as well as the extremely resource intensive exact methodology (Allison, 1995).

The proportional hazards model is extremely robust as it does not assume an underlying distribution for the hazard function as many other survival analysis techniques do (Allison, 1995). The principal assumption of the model, that of proportionality of hazards, is also robust and can be violated if the results of the analysis are interpreted somewhat differently than the results for a model where the proportionality of hazards assumption holds. The proportional hazards assumption posits that the effect of each covariate does not vary with time. Thus, if two groups are being compared, males and females for instance, the graphs of their survival function will be parallel. If the proportional hazards function is violated, the model then yields an average effect over all time points, which is considered to be a minor problem by some authors (Allison, 1995).

A simple proportional hazards model can be written as: $h_i(t) = \gamma(t)\exp(\beta_1 x_{i1} + \ldots + \beta_k x_{ik})$ The component γ (t) is where the distribution of the hazard function can be specified but is left unspecified to estimate the partial likelihood model. The components $\beta_1 x_{i1} - \beta_k x_{ik}$ represent a set of covariates which are exponentiated. Using the survival time dataset exhibited earlier, the first death is at 185 days. To estimate the probability of death at 185 days, we take the hazard of the individual who died at day 185 over the hazards of all persons who are still at risk for the event. The first observation is not in the risk-set as the individual dropped out of the study (was censored) at day 110. In this case, the hazard of death at day 185 is described by $L_1 = h_2(185)/h_2(185) + h_3(188) + h_4(220) + h_5(345) + .. + h_{45}(620)$, whereas the hazard for patient 45 can be expressed as $L45 = h_{45}(620)/h_{45}(620)$. The first expression can then be substituted into the partial likelihood equation shown earlier; this yields: $\gamma_0(185)e^{\beta x2}/ \gamma_0(185)e^{\beta x2} + \gamma_0(188)e^{\beta x3} + \gamma_0(220)e^{\beta x4} + \gamma_0(345)e^{\beta x5} + .. + \gamma_0(620)e^{\beta x45}$. Since there is not a distribution specified, the γs can be cancelled from the equation yielding: $L1 = e^{\beta x2} /e^{\beta x2} + e^{\beta x3} + e^{\beta x4} + e^{\beta x5} + .. + e^{\beta x45}$. This simplification demonstrates two important properties of partial likelihood estimation: 1) no underlying distribution is assumed; and 2) the estimates only depend upon the rank-order of the event times, not upon their exact value. This simplification also leads to the general partial likelihood case where ties and time-dependent covariates are not considered:

$$PL = \prod_{I=1}^{N} \left[\frac{e_{\beta xi}}{\sum_{j=1}^{n} Yij e_{bxj}} \right] \delta i$$

This can then be rewritten and Newton-Raphson estimation can be used to maximize this quantity with respect to β (Allison, 1995).

As mentioned previously, one of the major trappings of the Cox regression methodology is the ability to use time-dependent covariates. In the current study, the author used the time-dependent covariates to explain the effect of aftercare. Obviously this variable will change with time after release from treatment, and the estimation of a constant effect when the aftercare event has not happened to the individual in question is probably inappropriate. Time-dependent covariates are extremely resource intensive to estimate, so it will be an interesting exercise to see if it is practical to estimate the effect of these covariates on a large (approximately 500 subjects) sample.

Another recently emerging and extremely desirable feature of the Cox regression model is the ability to analyze multiple events. This enables the researcher to extend the information gathering capabilities of Cox regression even further. In essence, when a researcher is using the classical Cox regression model, he/she is choosing to discard the information associated with all but the first time to event values. With repeated events Cox regression, all events are used. Several approaches are available for this technique but none are without assumptions which may or may not be met.

After estimation with PHREG, Cox regression SAS® output yields an overall chi-square test that all the Beta parameters are zero, as well as a section displaying the covariates with their parameter estimates, standard errors, chi-squares and odds-ratios. The odds ratios are interpreted much like those from a logistic regression analysis output. Categorical covariates are interpreted with respect to a reference classification, thus if the risk ratio is 1.84 then the interpretation is that the risk of dying for this classification is, on average, 84 % greater than the reference classification controlling for the other covariates in the model. Quantitative covariates of interval scale or greater are interpreted with respect to a one unit increase in the variable of interest. For example, if age is the covariate of interest, and the risk ratio of age is .33, a helpful strategy is to subtract 1 from the risk ratio and multiply by 100. This yields the percent increase or decrease for each one unit change in the covariate. For this example then 100 * .33 – 1 = - 67%. Therefore, for each year of age increase, the risk of dying decreases by 67 %.

Proc PHREG has a set of regression diagnostic procedures that produce influence diagnostics and measures of residuals. For the present study, the author has selected deviance residuals (resdev) which measure the accuracy of the model and likelihood displacement statistic, which measures how much the log-likelihood will change if an individual is removed from the sample.

Logistic regression is another technique that was employed to test the stated hypotheses. One of the major reasons for the selection of this technique is to compare the logistic regression results to those of the Cox regression models in order to determine how much more sensitive the Cox regression models are than the logistic regression models, in essence investigating how much information is wasted by the dichotomization of the dependent variable.

Logistic regression models the following quantity, yielding estimated probabilities ranging between 0 and 1:

$$\text{Logit}(p_i) = \ln(p_i/1 - p_i) = \beta_0 + \beta_i X_{1i} + \beta_k X_{ki}$$

This model takes into account the non-linearity which is a result of constraining the outcome between 0 and 1. If ordinary least squares regression were used, the estimate would be biased as the OLS model assumes linearity. Inspection of an OLS regression residual plot with a dummy coded dependent variable would probably yield large value residuals at between 0 and .2 and between .8 and 1. Logistic regression bypasses this problem by using an S-shaped logistic function as the base probability distribution (SAS, 2000).

Model fit in logistic regression can be assessed using a graphical technique called Receiver Operator Characteristic curves (ROC). This technique plots a quantity called sensitivity or the accuracy at predicting true events (true positives/total actual positives) against 1 – specificity, which is the proportion of actual non-events (true negatives / total actual negatives). This ROC technique gives a nice graphical representation of the quality of the logistic regression model. A model with high predictability will rise quickly on the y – axis and encompass a maximum amount of area under the curve, whereas a poor model will be essentially a line rendered at a 45 degree angle.

Regression diagnostics were produced to find and eliminate influential observations in order to improve the fit of the models in both the Cox regression and logistic regression cases. Residual plots were used in both the Cox regression and logistic regression cases, to examine the fit of the data vis-à-vis the distributional assumptions of the models.

A complete analysis of interactions was undertaken in order to attempt to understand whether any improvement to the fit of the models could be realized by including interaction terms. Any interactions found were cross-validated by randomly splitting the dataset into two sections and applying the models with the interaction terms to both sections. Only interaction terms that proved significant in both sections were incorporated in the final models.

Proposed Hypothesis Tests

Hypotheses were tested first in a univariate form for each hypothesis. The first equation describes the Cox regression model. This model uses as its dependent variable the number of days or weeks which elapsed from the client's first psychiatric discharge date (index discharge) to the date the client is either: 1) rehospitalized for psychiatric reasons; or 2) censored by way of being disenrolled from Medicaid, or dying. The second equation describes the logistic regression model. This model uses as its dependent variable a dichotomous (0/1) variable which describes if a subject was rehospitalized at any time during the study year after the index discharge date.

Four multivariate tests were also conducted, two for the Cox regression model and two for the logistic regression model. Appropriate regression diagnostic procedures were employed to detect influential outliers for both the Cox regression and the logistic regression models. Collinearity diagnostics were employed to determine the degree of multicollinarity present in the multivariate models. Table 7 (pp. 66-67) includes a description of the derived variables used in the following models.

1) Hypothesis one was tested using logistic regression and Cox regression procedures. The Cox regression model took this form: log(diff1) = $\alpha(t) + \beta_1$(prev_hosp); and the logistic regression model took this form: Logit(status) = $\beta_0 + \beta_1$(prev_hosp)
2) Hypothesis two was tested using logistic regression and Cox regression procedures. The Cox regression model for sex took this form: log(diff1) = $\alpha(t) + \beta_1$(sex); and the logistic regression model took this form: Logit(status) = $\beta_0 + \beta_1$(sex). The Cox regression model for age took this form: log(diff1) = $\alpha(t) + \beta_1$(age); and the logistic regression model took this form: Logit(status) = $\beta_0 + \beta_1$(age). Finally, the Cox regression model for white race took this form: log(diff1) = $\alpha(t) + \beta_1$(white); and the logistic regression model took this form: Logit(status) = $\beta_0 + \beta_1$(white).
3) Hypothesis three was tested using logistic regression and Cox regression procedures. The Cox regression model for number of types of visits took this form: log(diff1) = $\alpha(t) + \beta_1$(variety00); and the

logistic regression model took this form: Logit(status) = β_0 + β_1(variety00). The Cox regression model for number of visits took this form: log(diff1) = α(t) + β_1(numvis00); and the logistic regression model took this form: Logit(status) = = β_0 + β_1(numvis00).

4) Hypothesis four was tested using logistic regression and Cox regression procedures. The Cox regression model for predicted psychiatric morbidity in the study year took this form: log(diff1) = α(t) + β_1(cdpspred); and the logistic regression model took this form: Logit(status) = β_0 + β_1(cdpspred).
5) Hypothesis five was tested using logistic regression models and Cox regression models. The Cox regression model for accumulated prescribed days in the study year took this form: log(diff1) = α(t) + β_1(days)(t); and the logistic regression model took this form: Logit(status) = β_0 + β_1(maxdose). Note the use of the time-dependent covariate represented by the "t" in the Cox regression model. The Cox regression model testing Risperdal vs. Seroquel vs. Zyprexa took this form: log(diff1) = α(t) + β_1(cdrug1) + β_2(cdrug2); and the logistic regression model took this form: Logit(status) = β_0 + β_1(cdrug1) + β_2(cdrug2).
6) The multivariate forms of the hypotheses were explored by four equations: 1) the first equation used Cox regression to determine the effect of each variable except those representing new-generation antipsychotic use. The form of this equation was: log(diff1) = α(t) +

β_1(prev_hosp) + β_2(sex) + β_3(age) + β_4(white) + β_5(variety00) + β_6(numvis00) + β_7(cdpspred) + β_8(days)(t). The second equation explored the relationship of all the variables in the previous model and added two variables that represented which new-generation antipsychotic medication subjects were prescribed. The equation also used only those subjects who were prescribed one of the new-generation antipsychotic medications under study. The form of this equation was as follows: log(diff1) = α(t) + β_1(prev_hosp) + β_2(sex) + β_3(age) + β_4(white) + β_5(variety00) + β_6(numvis00) + β_7(cdpspred) + β_8(days)(t) + β_9(cdrug1) + β_{10}(cdrug2). The third equation used logistic regression to study the relationship of the variables; like the first Cox regression equation, this equation also did not include variables coding for new-generation antipsychotic usage. This equation took this form: Logit(status) = β_0 + β_1(prev_hosp) + β_2(sex) + β_3(age) + β_4(white) + β_5(variety00) + β_6(numvis00) + β_7(cdpspred) + β_8(maxdose). Note that this equation did not include the time-dependent covariate "days," but used the "maxdose" covariate instead. The fourth equation was the logistic regression form of the second equation. This equation also did not include subjects who did not take a new generation antipsychotic medication. This equation can be written as: log(diff1) = α(t) + β_1(prev_hosp) + β_2(sex) + β_3(age) + β_4(white) + β_5(variety00) + β_6(numvis00) + β_7(cdpspred) + β_8(maxdose) + β_9(cdrug1) + β_{10}(cdrug2).

7) An exploratory analysis was also performed which examined the relationship of all of the variables mentioned in the previous section plus a number of variables which were available to the author but were not central to the hypotheses. The additional variables to be entered were psyh, psym, and psyl, which are CDPS-generated dichotomous indicator variables which partition the sample into predicted psychiatric morbidity classifications based upon diagnosis, and subl which was a CDPS-generated dichotomous indicator variable indicating a substance abuse diagnosis. The variable los1, which captured the length of stay of the client at index admission, was also entered. Another variable added was numcats, which was also a CDPS-generated variable that captured the number of CDPS categories into which the subject fell. The variable comm00 was a binary variable capturing the presence or absence of outpatient care delivered to the subject. Finally, the variable plancode was entered. This variable captured the membership of a client in a particular managed care organization. Table 11 (p. 97) describes the variables used in the exploratory analyses in tabular format. Cox regression and logistic regression were used to perform these exploratory analyses.
8) An analysis of the most frequent outpatient service codes used in SFY2000 vs. SFY2001 were conducted by analyzing the twenty frequently used codes in each year using a Wilcoxon signed-ranks test.

Chapter 4: Results

Hypotheses were tested by univariate equations, and then the predictor variables were entered simultaneously to test the multivariate forms of the hypotheses. The Cox regression results were presented first, followed by the logistic regression results for both the univariate and multivariate tests. Variables not significant at $p < .01$ were dropped from the model unless a compelling reason was found to do otherwise. Tables of descriptive statistics are presented in Table 9 while Table 10 displays a matrix of intercorrelations between the variables. Table 11 (p. 97) shows the results of the univariate Cox regression models, while Table 12 (p.109) displays the results of the univariate logistic regression models. The results of the multivariate Cox regression models are displayed in Table 13 (p. 110), while the results of the multivariate logistic regression models are displayed in Table 14 (p. 111).

Table 9: Mean Values, Standard Deviations or Frequencies for Predictor Variables by Rehospitalization Status

Variable	Statistics displayed	Rehospitalized during study year (n = 217)		Not rehospitalized during study year (n = 341)		Chi-square(1) or F(556)
		Mean or n	SD or Percent	Mean or n	SD or Percent	
Weeks from index hospitalization to rehospitalization or censure	Mean; SD	9.93	10.44	27.60	14.31	247.25****
Hospitalization in year prior	n; percent	102	47	90	26.39	24.96**
Sex (M)	n; percent	103	47.47	143	41.94	1.68
Age	Mean; SD	41.02	10.30	44.07	10.55	11.28***
White race	n; percent	65	29.95	81	23.75	2.63
Number of types of visits received in year prior	Mean; SD	5.21	6.64	3.16	4.60	18.74****
Number of visits received in year prior	Mean; SD	26.81	47.04	21.41	52.99	1.5
Predicted psychiatric morbidity	Mean; SD	1.91	0.81	1.56	0.71	29.40****
Prescribed days	Mean; SD	544.04	656.38	391.74	507.62	9.47**
Type of antipsychotic drug (2)						1.52
Risperadal	n; percent	59	37.34	77	43.26	
Seroquel	n; percent	42	26.58	39	21.91	
Zyprexa	n; percent	57	36.08	62	34.83	

Notes: 1) Chi-square test used for categorical variables; ANOVA for continuous variables

2) Type of antipsychotic drug tested using Chi-square test with 2 d.f. and 336 observations

*p < .05 **p < .01 ***p <.001 ****p <.0001

Table 10: Intercorrelations for Community Tenure and Predictor Variables

Variable	1	2	3	4	5	6	7	8	9	10	11
1) Weeks from index hospitalization to rehospitalization or censure											
2) Hospitalization in year prior	-.03										
3) Sex	.03	-.04									
4) Age	.11*	-.01	-.18****								
5) White race	-.05	-.05	.05	-.05							
6) Number of types of visits received in year prior	-.08	.54****	-.05	-.06	.00						
7) Number of visits received in year prior	.03	.20****	.06	.00	.02	.48****					
8) Predicted psychiatric morbidity	-.14***	.16***	-.04****	.03	-.00	.23****	.09*				
9) Prescribed days	.20****	.15****	-.11**	-.01	.16****	.22****	.15***	.11**			
Type of antipsychotic drug											
10) Risperadal	.02	.09*	-.06	.07	-.05	.06	.08	.06	.06		
11) Seroquel	-.06	.04	-.05	-.03	.04	.08	.06	.02	.28****	-.23****	
12) Zyprexa	-.05	-.04	.08*	-.08*	.14**	.04	.04	.05	.10*	-.30****	-.21****

*p < .05 **p < .01 ***p <.001 ****p <.0001

HYPOTHESIS ONE

The univariate form of hypothesis one was confirmed as both univariate proportional hazards and logistic regression models were found to be statistically significant. The Cox regression model found that individuals with psychiatric hospitalizations in the fiscal year prior to study were significantly more likely to be rehospitalized at any time in the study year: χ^2 (1, N = 558) = 17.85, p < .0001. The hazard ratio was 1.77, 95% CI = 1.36 to 2.32, indicating that the hazard of rehospitalization for those with previous hospitalizations was 77.7 percent greater on average during the year of study.

The logistic regression model showed a similarly strong effect when compared to the Cox regression model: χ^2 (1, N = 558) = 24.41, p < .0001. The odds ratio was 2.47, 95% CI = 1.72 to 3.54 indicating that persons with previous hospitalizations were almost two and a half times as likely to be rehospitalized as those that were not hospitalized in the previous year. The overall fit of the model was poor, with a c-statistic of .60 and a pseudo r-square of .06 indicating other variables need to be included in the model.

HYPOTHESIS TWO

The univariate forms of hypothesis two only found support in the age variable with both the proportional hazard and logistic regression models indicating significant differences on this variable. The sex, and white vs. non-white variables did not show differences in any model.

The Cox regression model indicated that increasing age was negatively related to the hazard of being rehospitalized in the study period: χ^2 (1, N = 558) = 12.08, $p < .001$. The hazard ratio of .98, 95% CI = .97 to .99, indicates that for each one year increase in age, the hazard of being rehospitalized at any point in time decreased by approximately two percent.

The logistic regression model results were very similar: χ^2 (1, N = 558) = 10.92, $p < .01$. The odds ratio of .98, 95% CI = .96 to .99 indicates that for each one-year increase in age, the odds of rehospitalization decrease by approximately two percent. The overall fit of this model was poor with a c-statistic of .581 and a pseudo r-squared of .027.

HYPOTHESIS THREE

The univariate form of hypothesis three was not supported in either the Cox regression model or the logistic regression model. While both models produced significant results, subjects who received a greater variety of outpatient services in the prior year had shorter community tenures and were more likely to be rehospitalized; this is exactly the opposite of what was expected.

The Cox regression model indicated that variety of services was positively related to the hazard of being rehospitalized in the study period: χ^2 (1, N = 558) = 18.84, $p < .0001$. The hazard ratio of 1.05, 95% CI = 1.03 to 1.07 indicates that each distinct type of service received was related to an increase in the hazard of being rehospitalized at any point in time by approximately five percent.

The logistic regression model results showed a similarly negative trend: χ^2 (1, N = 558) = 16.63, $p < .0001$. The odds ratio of 1.07, 95% CI = 1.04 to 1.10

indicates that for each distinct service received, the odds of being rehospitalized increased by about seven percent. The overall fit of the model was fair, with a c-statistic of .60 and a pseudo r-square of .04, indicating other variables need to be included in the model.

HYPOTHESIS FOUR

The univariate form of hypothesis four was supported in the Cox regression model and in the logistic regression model. The Cox regression model showed that subjects with greater predicted psychiatric morbidity had an increased hazard of being rehospitalized at any point in time. The logistic regression model also showed that subjects with greater morbidity scores were more likely to be rehospitalized in the study year.

The Cox regression model indicated that the greater predicted psychiatric morbidity a subject is estimated to have, the less community tenure the subject is likely to have: χ^2 (1, N = 558) = 29.53, $p < .0001$. The hazard ratio of 1.40, 95% CI = 1.24 to 1.58 indicates that with each one-unit increase in psychiatric morbidity, the subject's chance of being rehospitalized at any point in time increases about forty percent.

The logistic regression model results showed a similarly strong effect: χ^2 (1, N = 558) = 28.02, $p < .0001$. The odds ratio of 2.1, 95% CI = 1.60 to 2.75 indicates that for each unit increase in psychiatric morbidity, the odds of rehospitalization in the year of study approximately double. The overall fit of the model was fair, with a c-statistic of .65 and a pseudo r-square of .07 indicating other variables need to be included in the model.

HYPOTHESIS FIVE

Hypothesis five was not confirmed for the Cox regression model or the logistic regression model for the type of new generation antipsychotic medication prescribed, but the portion of the hypothesis that predicted that increasing medication compliance would increase community tenure and reduce the chance of rehospitalization found mixed results. The Cox regression model using prescribed days as a time-dependent covariate was not significant; however, the logistic regression model which used the cumulative maximum number of prescribed days was, indicating that patients having more prescribed days have a greater risk of rehospitalization. The logistic regression model results showed a moderate effect: χ^2 (1, N = 558) = 8.76, p < .01. Custom odds ratios were calculated based on a sixty-day increment. The odds ratio of 1.03, 95% CI = 1.01 to 1.05 indicates that for each 60 prescribed day increase, the risk of rehospitalization increased by approximately three percent.

MULTIVARIATE ANALYSES

The multivariate phase of this project was completed by entering all of the variables which were used in the univariate analyses in multivariate equations with other predictor variables. This allows the researcher to control for the other variables in the equation and judge the unique contribution of each variable to the prediction of the dependent variable values above and beyond the contribution of the other predictor variables in the model.

For these analyses, four multivariate models were run: 1) a multivariate Cox regression model excluding the new-generation antipsychotic medications, thereby including subjects with less serious diagnoses; 2) a multivariate Cox regression model including the new-generation antipsychotic variables, thereby limiting the analysis to those subjects who had new-generation antipsychotic medications prescribed to them; 3) a logistic regression equation including the variables capturing new-generation antipsychotic usage; and 4) a logistic regression equation which excludes the variable capturing new-generation antipsychotic use. Please refer to Table 14 (p. 111) for complete results of multivariate Cox regression and Table 15 (p.111) for complete results of multivariate logistic regression. The presence of interactions was explored for all four multivariate models and none were found.

The first multivariate analysis revealed that only previous hospitalization, predicted psychiatric morbidity, and age were significant at .01 or greater. The overall fit of the model was acceptable: χ^2 (3, N = 558) = 53.64, $p < .0001$. The previous hospitalization and predicted morbidity variables were equally influential in this model. The previous hospitalization variable showed a strong effect: χ^2 (1, N = 558) = 12.36, $p < .001$, as did the predicted psychiatric morbidity variable effect: χ^2 (1, N = 558) = 23.92, $p < .0001$. The hazard ratio of the previously hospitalized variable was slightly higher, 1.62 vs. the 1.37 of the predicted morbidity variable, indicating that persons who were previously hospitalized had a sixty percent greater hazard of being rehospitalized at any point in the study period; whereas each point of increase in the predicted morbidity of a subject indicated a thirty-seven percent increase in the hazard of rehospitalization. The age variable also indicated a strong effect: χ^2 (1, N = 558) = 14.21, $p < .001$

and was negatively associated with the hazard of rehospitalization. The hazard ratio of age was .98, indicating that each year of age reduced the hazard of rehospitalization by about two percent. A plot of the estimated community tenure of the subjects is displayed in Figure 6 (p. 103). Of note is a rather large drop in the probability of remaining in the community from discharge to five weeks. After this point, a leveling effect seems to take place until the end of the study period. Figure 7 (p. 104) displays a comparison between previously hospitalized subjects and those who were not found to be hospitalized in the year prior to the study year. Clearly subjects who had been hospitalized prior to the study year have a much greater chance of being rehospitalized. At the end of the study period, only about 40% remained in the community vs. about 60% of the not previously hospitalized subjects. This analysis also controls for psychiatric morbidity and age.

The second multivariate analysis included only those patients who had received Resperdal, Seroquel, or Zyprexa during the study period. In this analysis, the only significant variables were the predicted psychiatric morbidity of the patient: χ^2 (1, N = 336) = 15.79, $p < .0001$, and age: χ^2 (1, N = 336) = 7.24, $p < .01$. The hazard ratio for the morbidity effect was 1.36, indicating a thirty-six percent increase in the hazard of rehospitalization for each one-point increase in the predicted morbidity score. Age again showed a prophylactic effect with each year of age decreasing the hazard by approximately two percent.

The third multivariate model used logistic regression, and the results were similar to the first multivariate model; previous hospitalization, age, and predicted psychiatric morbidity were also the only significant predictors. The overall fit of the model was acceptable: χ^2 (1, N = 558) = 50.75, p < .0001, c-statistic .70 and pseudo r-square of .14. The predicted morbidity variable appeared to be slightly more influential than the previous hospitalization variable, with the age variable providing a slight prophylactic effect against rehospitalization, much like the Cox regression model. The predicted morbidity variable exhibited a strong effect: χ^2 (1, N = 558) = 20.21, p < .0001, with an odds ratio of 1.83, as did the previous hospitalization variable (1, N = 558) = 17.46, p < .0001, and an odds ratio of 2.21. The effect of age was somewhat less pronounced than that of the other two variables but differed in directionality (1, N = 558) = 12.58, p < .001, and an odds ratio of .97. In sum, it appears that a previous hospitalization increases the risk of rehospitalization by a factor of two, as does each additional point of predicted morbidity; on the other hand, each year of age reduces the chance of being rehospitalized by about three percent.

The fourth multivariate model found only the predicted psychiatric morbidity variable (1, N = 335) = 9.78 p < .01 to be significant at the alpha level set for this study. The odds ratio for this variable was 1.77.

Figure 11 (p. 108) displays the receiver operating characteristic curve for the significant multivariate and univariate logistic regression models that lent themselves to this type of plot. A roc curve which rises rapidly and encompasses a great deal of area under the curve represents a good model, and a model which rises diagonally at a 45-degree angle represents a poor model with a c-statistic of

.5. Note the area under the curve for the multivariate model vs. the other models displayed.

INVESTIGATION OF STATISTICAL ASSUMPTIONS

Both Cox regression and logistic regression have somewhat looser requirements concerning the distribution of predictor variables than traditional linear models. Although normally distributed predictors are not required for either model, multivariate normality or a linear association among predictors can enhance the usefulness of the model (Tabachnick & Fidell, 1996). However, both Cox regression and logistic regression are sensitive to influential outliers and multicollinearity (Tabachnick & Fidell, 1996; Allison, 1995). In order to detect these problems, analyses were conducted on both the Cox regression and logistic regression models to detect each problem. In addition to these analyses, the proportional hazards assumption was examined for the Cox regression model.

In order to detect influential observation in the logistic regression models, a bubble plot was created (Figure 8, p. 105) which plotted the predicted probabilities by the change in the chi-square statistic if the observation is removed. The plot also includes a feature which expands the size of the observation in proportion to the observation's c-statistic value. The c-statistic captures the change in each parameter estimate when the observation is removed. Observations which exceeded a DIFCHISQ of four were examined and removed from the model. The value of four is often used as a threshold for this technique as it represents the 95% percentile of the chi-square distribution (Homer & Lemeshow, 1989). All three observations that were removed were subsequently returned to the model, as their removal did not change the results, only the overall fit of the model. With the residuals included, the overall logistic multivariate model 3 possessed a Wald chi-square of 50.75; with the residual excluded, the Wald value increased to 59.50. The pseudo R-square value increased from .137 to .165.

The presence of multicollinearity was investigated by using the variance inflation factor statistic available with SAS PROC REG (SAS, 2002). No collinearity diagnostics are available within the LOGISTIC or PHREG procedures. The variance inflation factor (VIF) values of all of the variables entered in the multivariate models were below ten, which is considered a "rule of thumb" for the use of the statistic (SAS, 2002).

The proportionality of hazards assumption was not checked with a formal method, as Allison (1995) states that the most serious deviation from the assumption is the ability to evaluate the model from an average effects perspective only, which Allison believes is still desirable and yields useful information. A visual test of the proportionality assumption can be viewed in Figure 7 (p. 104). This figure displays the estimated community tenure times of previously hospitalized vs. not previously hospitalized patients. Clearly the lines are not parallel; this is a violation of the proportionality assumption, but, as Allison states, it is a violation most authors can live with.

STATISTICAL SIGNIFICANCE

Readers should keep in mind that multiple statistical tests were performed in the course of conducting this research which attenuates the possibility that some significant findings were due to capitalization on chance. The Bonferoni criterion or other technique to control for type-one error was not formerly employed; therefore, readers should exercise caution when interpreting findings with $p < .05$ or $p < .01$ findings. If the Bonferoni criterion were employed, many of the findings in this paper which were significant at the $p < .0001$ level would remain significant at the .01 level even if 100 statistical tests had been performed.

EXPLORATORY ANALYSES

The addition of more variables in the exploratory analysis phase did not improve prediction in the logistic regression model; however, the introduction of variables capturing the number of co-morbid conditions and a dichotomous variable flagging the presence of a potentially expensive psychiatric condition did improve prediction over the baseline multivariate Cox analysis used in model one. For the exploratory analyses, interaction effects were analyzed at the 2-way maximum level, but no significant interactions were found. Forward selection was used with a significance level of .1 used to set the threshold for entry into the model.

Table 11: Predictor Variables Entered into Exploratory Analyses

Variable	Description
Plancode	Designation of managed care plan
Sex	Sex of subject
Psyh	Indicator of high-cost psychiatric diagnosis
Psym	Indicator of medium-cost psychiatric diagnosis
Psyl	Indicator of low-cost psychiatric diagnosis
Cdpspred	CDPS prospective predicted value
Numcats	Number of CDPS categories
Los1	Length of stay at index admission
Age	Age of the subject in years
Prev_hosp	Indicator of psychiatric hospitalization in previous year
White	Indicator of white race
Numvis00	Number of aftercare visits subject received in previous year
Variety00	Variety of aftercare visits subject received in previous year
Numdays00	Number of days of psychiatric medications subjects received
Ng00	Indicator of new-generation antipsychotic use in previous year
Variety01	Number of different types of outpatient visits received after discharge and before any rehospitalization
Numvis01	Number of outpatient visits received after discharge and before any rehospitalization
after	Time-dependent variable indicating the receipt of outpatient care by that week
comm00	Indicator of outpatient psychiatric services received in previous year
Days	Number of days of psychiatric medications subjects received in study year. Time-dependent variable used in Cox regression only.

In the Cox regression analysis, the continuous variable which captures psychiatric morbidity dropped from the equation and was replaced by two other CDPS-derived variables: psyh, which is an dichotomous indicator variable denoting the presence of a high-cost psychiatric condition; and numcats, which is a variable ranging from zero to eight in the case of this study which captures the number of CDPS categories into which the individual fell. This variable can be considered a good proxy for the number of co-morbid conditions with which an individual is suffering.

The Wald statistic of the model was significant (4, N = 558) = 73.74 $p < .0001$. This represented a significant improvement over the hypothesized model which yielded an overall Wald statistic of χ^2 (3, N = 558) = 53.64, $p < .0001$. The previous hospitalization variable was only significant at the $p < .05$ level and thus is only included here for reference as it possesses such a strong effect in other models and in many other studies. The effect of age again shows a prophylactic effect toward rehospitalization: χ^2 (1, N = 558) = 16.25, $p < .0001$. The hazard ratio was .974, which is in line with the age effect found in many of the other analyses in this study. The psychiatric high-cost variable exhibited one of the strongest effects: χ^2 (1, N = 558) = 20.19, $p < .0001$ and had a very high hazard ratio of 2.22. The number of CDPS categories variable also exhibited a strong effect: χ^2 (1, N = 558) = 27.60, $p < .0001$, and a large hazard ratio of 1.33. Thus it appears that hazard of rehospitalization more than doubles with a high-cost diagnosis, and increases about thirty-three percent with each subsequent different type of CDPS diagnostic classification. Figure 10 (p. 107) displays the community tenure by number of CDPS categories, while Table 16 (p. 112) displays the results of the analysis.

A logistic regression analysis as well as a Cox regression analysis were performed which used only the data from the two largest plans in the analyses in order to determine if subjects' membership in a particular health plan was predictive of rehospitalization. The results of these models are presented in tables 17 and 18 respectively.

The Cox regression model using only data from the two largest plans generated similar results to the logistic regression analysis, with the exception that the numcats variable also was entered into the model. The numcats variable was the strongest effect in this model: χ^2 (1, N = 546) = 23.67, p < .001 and a hazard ratio of 1.31. The overall fit of this model was good: χ^2 (5, N = 546) = 81.62, p < .0001; the fit of this model was a slight improvement over the fit of the previous Cox model presented which was run on all of the data not just the data from the largest health plans.

In the logistic regression model the plancode variable was significant: χ^2 (1, N = 546) = 6.85, p < .01, with an odds ratio of .61. This indicates that the subjects of one plan only had 61% of the chance of getting rehospitalized compared to the subjects belonging to the other health plan. The psyh and age variable were more predictive than plan membership, with the psyh variable being the strongest predictor in the model: χ^2 (1, N = 546) = 25.90, p < .0001, with an odds ratio of 2.96; the age variable was less important that the psyh variable at χ^2 (1, N = 546) = 11.01, p < .001 and an odds ratio of .97. The overall fit of the model was marginally acceptable: χ^2 (3, N = 546) = 40.44, p < .0001. The pseudo-r square was .10 and the c-statistic value was .66.

In order to explore the effect of aftercare delivered during the study year, several variables were added to the Cox regression models which were designed to capture the effects of aftercare on rehospitalization behavior. These variables were: aftercare, a time-dependent covariate which was indexed 1 to 52 and changed from zero to one in the week the patient received the first aftercare visit; numvis01, a variable capturing the number of outpatient visits the subject received after discharge from the index hospitalization and before rehospitalization; variety01, a variable capturing the number of types of outpatient services the subject received. The use of the time-dependent variable aftercare, which signals aftercare has occurred for a patient at a particular point in time, is not central to this dissertation as this dissertation seeks to test and discover predictors of rehospitalization which occur before the event. Nonetheless, the inclusion of this variable is probably important as many other studies use predictor variables which co-occur in the same time frame as the rehospitalization episode the researcher is attempting to predict.

When the aftercare variable was entered into the Cox regression models, it was the most predictive of community tenure in the model using all of the data as well as the model using only data contributed by the two largest plans. The overall fit of the Cox model using all data with the time-dependent aftercare variable increased to χ^2 (5, N = 558) = 103.58, p < .0001, from the (4, N = 558) = 73.74 p < .0001, of the exploratory Cox model which did not include the aftercare variable. The difference between these two models was significant at 1 degree of freedom: χ^2 = (-2 log likelihood of larger model = 1712.62) - (-2 log likelihood of smaller model 1803) = 90.38 p <.0001. The aftercare variable exhibited a negative relationship with community tenure. This relationship was so strong that those who had aftercare had 3.5 times the hazard of being rehospitalized than those who did not receive aftercare. The complete results of this analysis can be seen in Table 19 (p. 113).

A similar story emerged when the aftercare variable was added to the model which only employed data from the two largest plans. In this model, the aftercare variable also became most predictive of community tenure. The fit of this model was also good: χ^2 (6, N = 546) = 107.24, p < .0001. This model also exhibited a substantial increase in predictive power from the Cox regression exploratory model that did not include aftercare as a variable. The likelihood ratio chi-square test for difference in models was significant with one degree of freedom: χ^2 = (-2 log likelihood of larger model = 1672.69) - (-2 log likelihood of smaller model 1757.67) = 84.98 p <.0001. The aftercare effect in this model showed a strong relationship between aftercare and decreasing community tenure. The full results of this analysis can be seen in Table 20 (p.114).

STABILITY OF OUTPATIENT PROCEDURE CODING

In order to determine how stable outpatient billing remained from SFY2000 to SFY2001, a frequency analysis was performed as well as a Wilcoxon signed-rank test. Overall, the top billing code in both years was 8010X, Community Support by Professional, 30 minutes. This code had a large increase in billing percentage from 18.8% in SFY2000 to 28.21% in SFY2001. Total outpatient services delivered to the sample decreased from 18,007 in SFY2000 to 13,032 in SFY2001. The Wilcoxon signed-rank test performed on the top twenty most frequently billed outpatient procedure codes statistically confirmed that services had decreased from SFY2000 to SFY2001 (Signed-rank $S = 53$, $p < .05$).

Figure 6: Estimated Time in Community Controlling for Psychiatric Morbidity, Age, and Previous Hospitalization

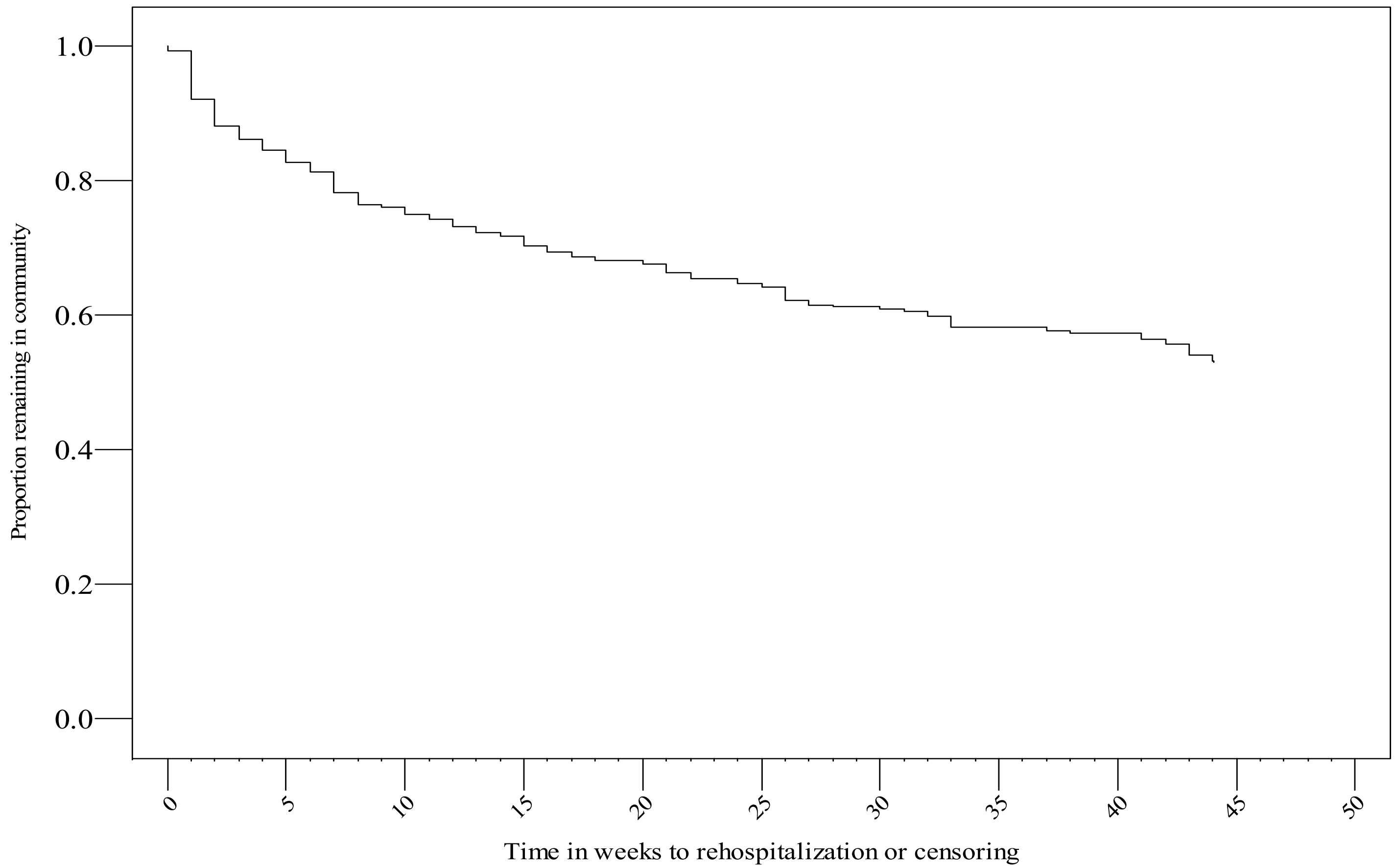


Figure 7: Comparison of Community Tenure of Previously Hospitalized vs. Not Previously Hospitalized Subjects

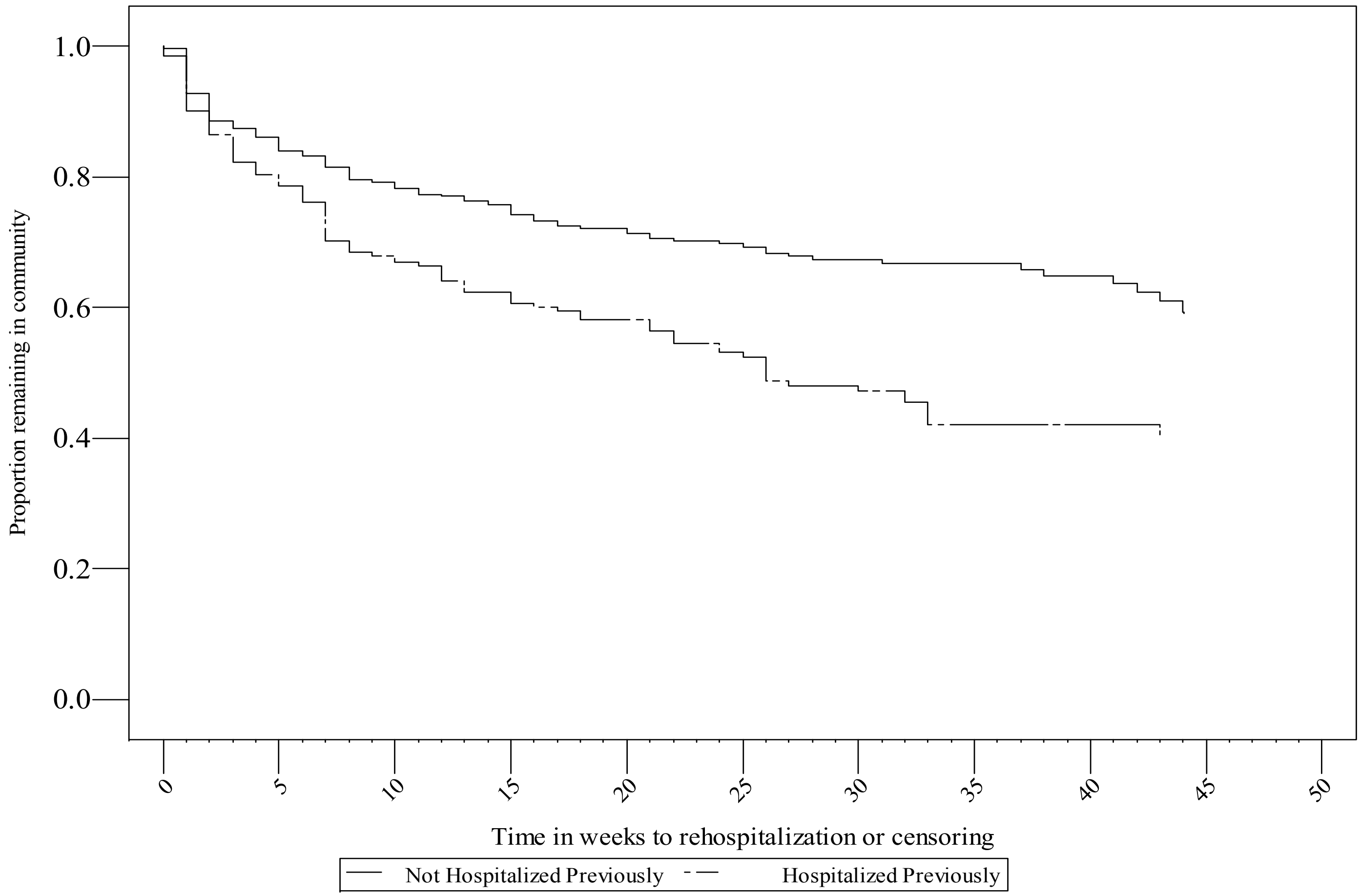


Figure 8: Deviance Residual Plot for Multivariate Logistic Regression Analysis

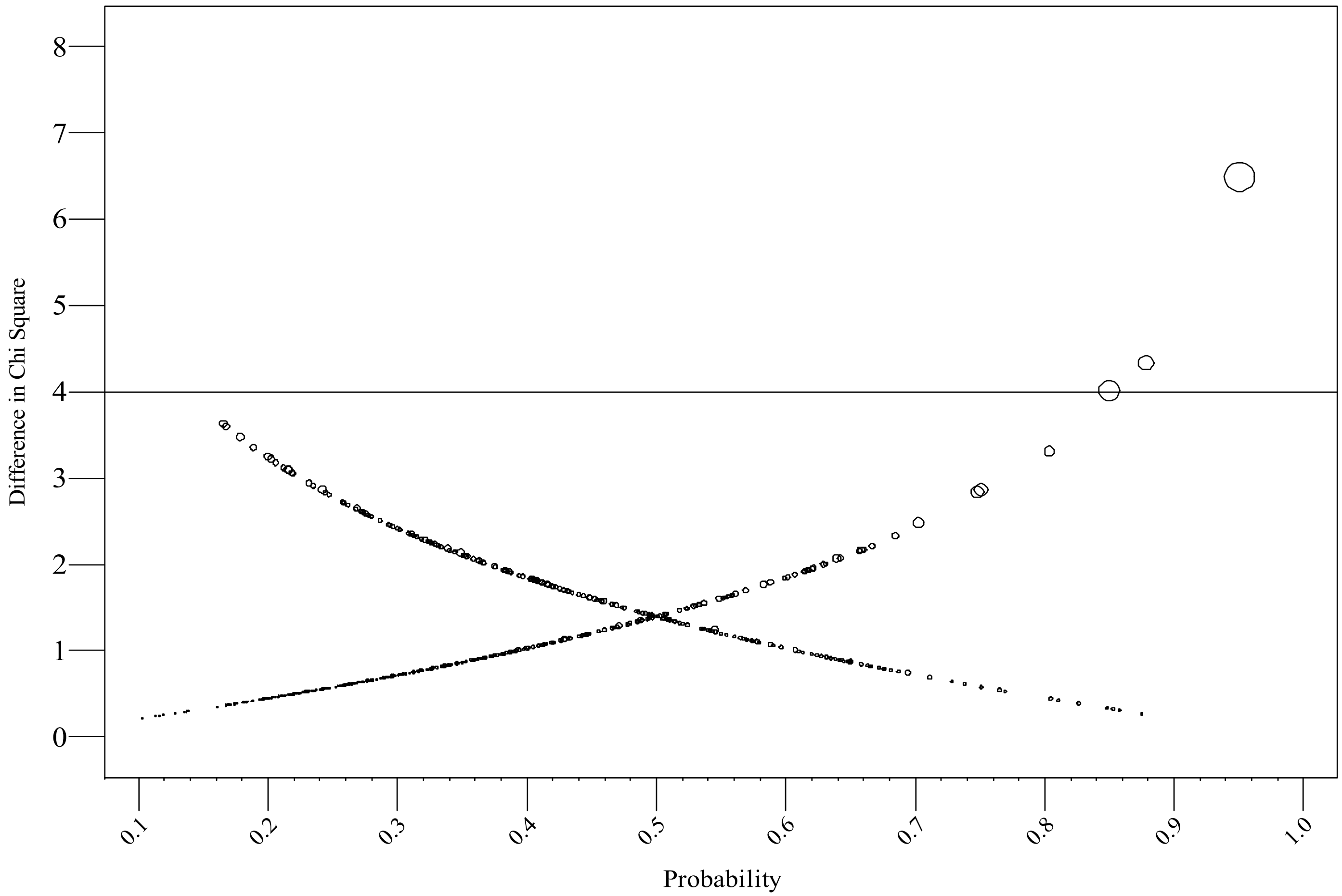


Figure 9: Community Tenure for High-Cost vs. Low-Cost Psychiatric Diagnoses

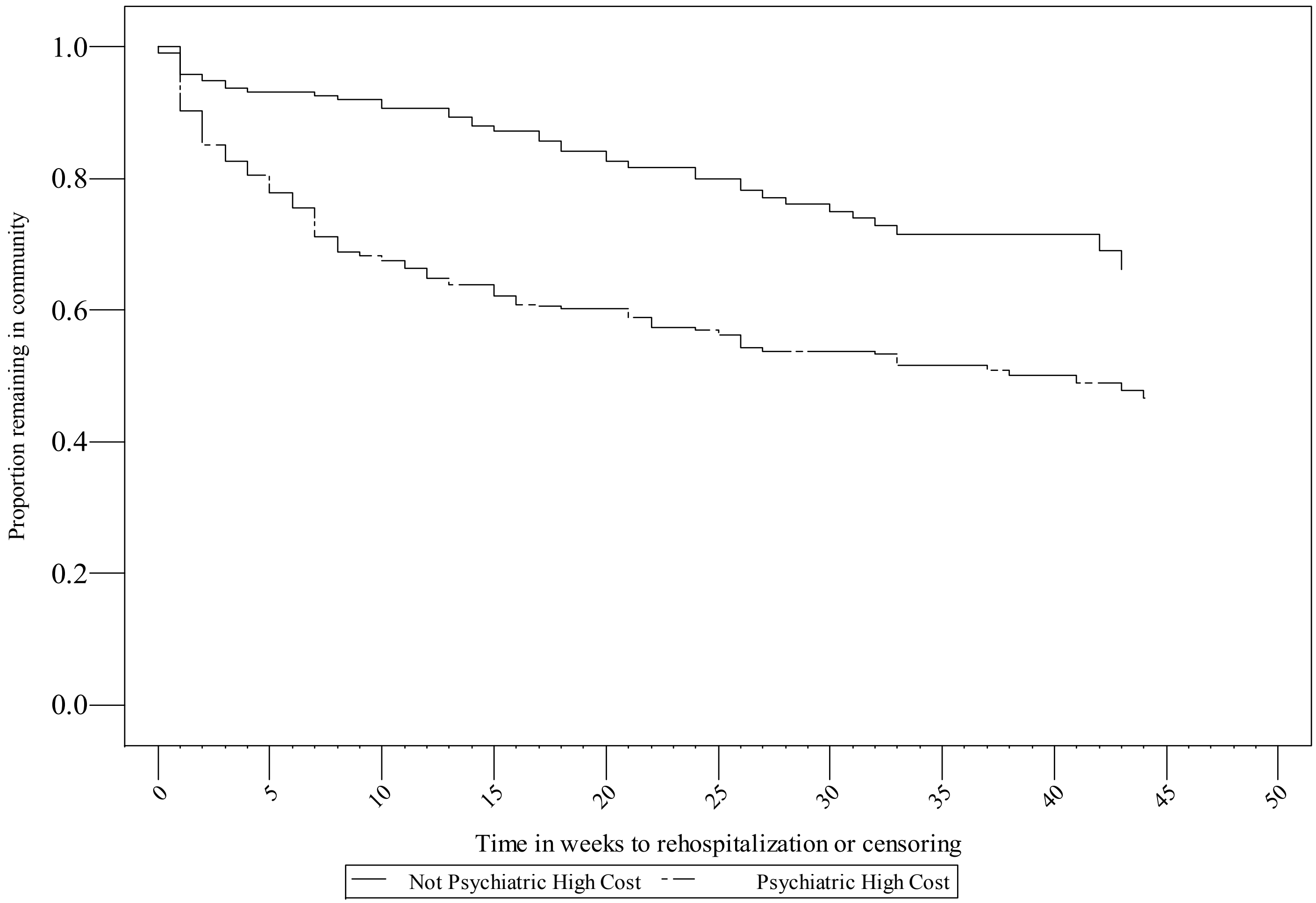


Figure 10: Community Tenure by Number of CDPS Categories

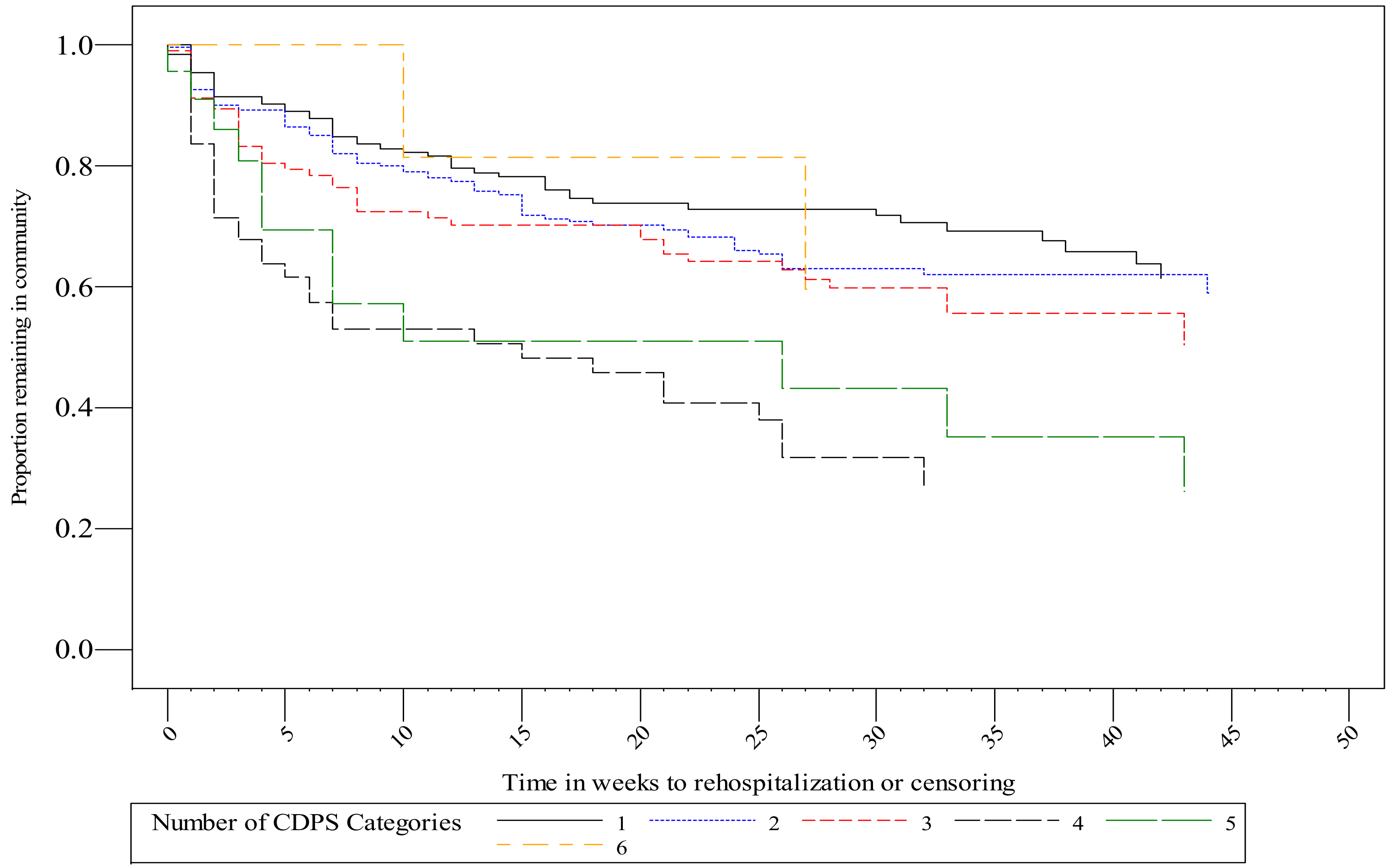


Figure 11: Receiver Operating Characteristic Curve of Significant Logistic Regression Models

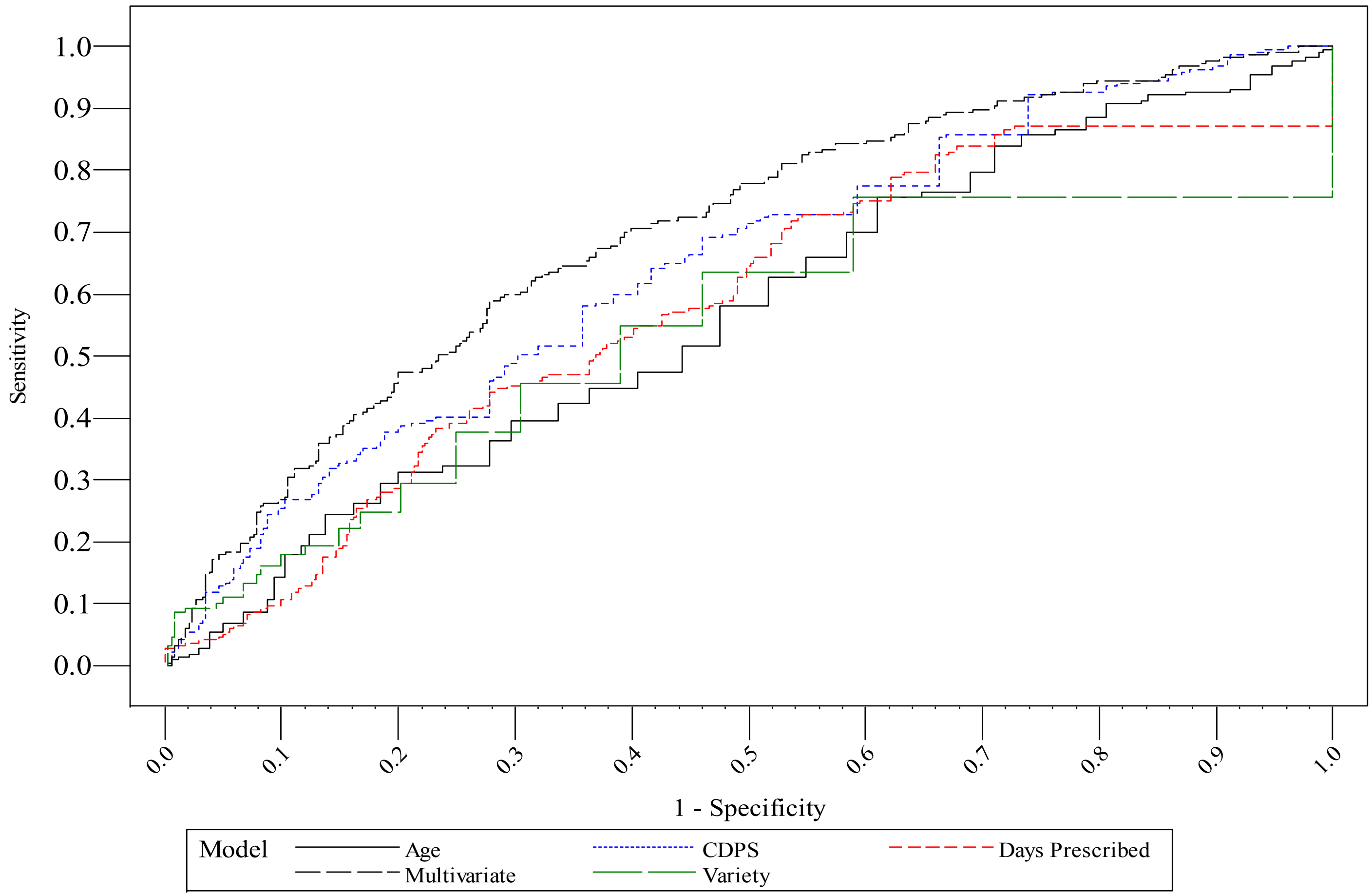


Table 12: Results of Univariate Cox Regression Models

Variable	β	SE	Hazard Ratio	Lower 95% CI	Upper 95% CI	Wald statistic	N
Previous Hospitalization	0.57	0.14	1.77	1.36	2.32	17.85****	558
Sex	0.16	0.14	1.18	0.90	1.54	1.15	558
Age	-0.02	0.00	0.98	0.97	.99	12.08***	558
Race	0.25	0.15	1.28	.96	1.71	2.77	558
Variety of Services Received	0.04	0.01	1.05	1.03	1.07	18.84****	558
Number of Services Received	0.00	0.00	1.00	1.00	1.00	0.63	558
Predicted Psychiatric Morbidity	0.34	0.06	1.40	1.24	1.58	29.53****	558
Type of Antipsychotic Drug							
Risperdal vs. Zyprexa	-0.14	0.11	0.85	0.69	1.07	1.66	336
Seroquel vs. Zyprexa	0.12	0.12	1.14	.90	1.44	1.14	336
Prescribed Days	0.00	0.00	1.00	1.00	1.00	2.60	558

*p < .05 **p < .01 ***p <.001 ****p <.0001

Table 13: Results of Univariate Logistic Regression Models

Variable	β	SE	Odds Ratio	Lower 95% CI	Upper 95% CI	Wald statistic	N
Previous Hospitalization	0.90	0.18	2.47	1.72	3.54	24.41****	558
Sex	0.22	0.17	1.25	0.89	1.76	1.64	558
Age	-0.02	0.00	0.97	0.96	0.99	10.92**	558
Race	0.31	0.20	1.37	.95	2.0	2.63	558
Variety of Services Received	0.06	0.02	1.07	1.04	1.10	16.63****	558
Number of Services Received	0.00	0.00	1.00	0.99	1.00	1.45	558
Predicted Psychiatric Morbidity	.74	0.14	2.10	1.60	2.75	28.02****	558
Type of Antipsychotic Drug							
Risperdal vs. Zyprexa	-0.17	0.15	0.83	0.51	1.36	1.35	336
Seroquel vs. Zyprexa	0.16	0.17	1.17	0.66	2.06	.95	336
*Prescribed Days	0.00	0.00	1.03	1.01	1.05	8.76***	558

*p < .05 **p < .01 ***p <.001 ****p <.0001
*Prescribed Days is the cumulative maximum number of days of medication prescribed and is not time dependent. Odds ratios are based upon 60 day increments.

Table 14: Results of Multivariate Cox Regression Models

Variable	β	SE	Hazard Ratio	Lower 95% CI	Upper 95% CI	Wald statistic	N
Model One							
Previous Hospitalization	0.48	0.14	1.62	1.23	2.12	12.36***	558
Age	-0.02	0.00	0.98	0.96	0.99	14.21***	558
Predicted Psychiatric Morbidity	0.31	0.06	1.37	1.20	1.55	23.92****	558
Model Two							
Age	-.02	.00	.98	.96	.99	7.25**	335
Predicted Psychiatric Morbidity	0.30	0.08	1.35	1.17	1.58	15.80****	335

*p < .05 **p < .01 ***p <.001 ****p <.0001

Table 15: Results of Multivariate Logistic Regression Models

Variable	β	SE	Odds Ratio	Lower 95% CI	Upper 95% CI	Wald statistic	N
Model Three							
Previous Hospitalization	0.80	0.19	2.21	1.52	3.21	17.46****	558
Age	-0.03	0.01	0.97	0.95	0.99	12.58***	558
Predicted Psychiatric Morbidity	0.61	0.14	1.83	1.4	2.24	20.21****	558
Model Four							
Predicted Psychiatric Morbidity	.57	.18	1.77	1.24	2.52	9.78**	335

*p < .05 **p < .01 ***p <.001 ****p <.0001

*Prescribed Days is the cumulative maximum number of days of medication prescribed and is not time dependent

Table 16: Results of Exploratory Cox Regression Model

Variable	β	SE	Hazard Ratio	Lower 95% CI	Upper 95% CI	Wald statistic	N
Previous Hospitalization	0.33	0.14	1.39	1.05	1.83	5.48*	558
age	-0.03	0.00	0.97	0.96	0.99	16.25****	558
psyh	0.80	0.18	2.22	1.57	3.15	20.19****	558
numcats	0.28	0.05	1.33	1.20	1.48	27.60****	558

*p < .05 **p < .01 ***p <.001 ****p <.0001

Table 17: Results of Exploratory Cox Regression Run Only on Large Plan Data

Variable	β	SE	Hazard Ratio	Lower 95% CI	Upper 95% CI	Wald statistic	N
Plan	-0.24	0.07	0.79	0.67	0.90	10.95***	546
psyh	0.79	0.18	2.20	1.56	3.17	19.28****	546
age	-0.03	0.00	0.97	0.96	0.99	17.77****	546
prev_hosp	0.36	0.14	1.43	1.08	1.89	6.3*	546
numcats	0.27	0.06	1.31	1.18	1.47	23.67***	546

*p < .05 **p < .01 ***p <.001 ****p <.0001

Table 18: Results of Exploratory Logistic Regression Run Only on Large Plan Data

Variable	β	SE	Hazard Ratio	Lower 95% CI	Upper 95% CI	Wald statistic	N
Plan	-0.24	0.09	0.61	0.43	0.89	6.85**	546
age	-0.03	0.00	0.97	0.95	0.99	11.01***	546
psyh	1.08	0.21	2.96	1.95	4.50	25.90****	546

*p < .05 **p < .01 ***p <.001 ****p <.0001

Table 19: Results of Exploratory Cox Regression Analysis Including Aftercare Variable

Variable	β	SE	Hazard Ratio	Lower 95% CI	Upper 95% CI	Wald statistic	N
Previous Hospitalization	0.30	0.14	1.35	1.01	1.78	4.36*	558
age	-0.02	0.00	0.98	0.96	0.99	15.88****	558
psyh	0.52	0.18	1.68	1.80	2.40	8.25**	558
numcats	0.22	0.05	1.24	1.12	1.39	15.88****	558
aftercare	1.27	0.19	3.55	2.43	5.20	42.32****	558

*p < .05 **p < .01 ***p <.001 ****p <.0001

Table 20: Results of Exploratory Cox Regression Analysis Including Aftercare Variable and Only Large Plan Data

Variable	β	SE	Hazard Ratio	Lower 95% CI	Upper 95% CI	Wald statistic	N
Previous Hospitalization	0.33	0.14	1.39	1.05	1.86	5.38*	558
age	-0.03	0.00	0.97	0.96	0.99	14.50****	558
psyh	0.53	0.18	1.71	1.18	2.45	8.36**	558
numcats	0.20	0.06	1.23	1.10	1.38	13.47***	558
plancode	-0.21	0.07	0.80	0.70	0.93	8.72**	558
aftercare	1.22	0.20	3.41	2.31	5.02	38.58****	558

*p < .05 **p < .01 ***p <.001 ****p <.0001

Chapter 5: Discussion

Summary and Integration of Results

The purpose of this dissertation was to explore the primary antecedents of psychiatric rehospitalization using modern multivariate statistical techniques. Five questions were proposed for study. Univariate and multivariate techniques were used to explore these questions in both the logistic and Cox regression cases. The dependent variable for the logistic regression case was a dichotomous variable indicating if the subject was rehospitalized during the study period, while the dependent variable in the Cox regression case was the number of weeks from discharge from a psychiatric inpatient facility to rehospitalization.

The first question posed sought to replicate a consistent finding in the literature that suggests that prior psychiatric hospitalization experiences predict future psychiatric hospitalizations. The second question studied the effect of the demographic variables age, sex, and race, while the third question explored the role of the variety and the number of outpatient services delivered to a client in reducing rehospitalization. The fourth question, which is central to this dissertation, asked if a variable could be found which could effectively control for psychiatric morbidity. The fifth and final question asked if the type of new-generation antipsychotic medication prescribed mattered in rehospitalization experience. This final question also asked if the number of prescribed days of medication taken influenced a subject's risk of rehospitalization.

When the totality of the results are considered, it appears that age, previous hospitalization, and psychiatric morbidity are the only significant predictors of either community tenure between psychiatric hospitalizations or the presence of a rehospitalization episode during the year under study. This picture is muddied somewhat by the exploratory analyses offered in this research which show that the effect of previous hospitalization upon community tenure becomes non-significant if the psychiatric morbidity of a patient is controlled for in the model. Furthermore, when exploratory analyses were run just on the data available from the two largest managed care plans in the sample, plan membership became significant in the models. Yet another interpretation develops when the time-dependent variable aftercare is included and becomes the most significant variable in the Cox regression exploratory models. This aftercare variable, while interesting, does not have the promise of the other variables as it is developed contemporaneously with the outcome measures which limits its use for predicting year to year service utilization trends.

Hypothesis One

Psychiatric hospitalizations were predictive of subsequent rehospitalization in both the univariate and multivariate Cox and logistic regression models. In the multivariate exploratory analysis, this effect did diminish to the point of non-significance when a variable indicating the presence or absence of a high-cost psychiatric condition was entered along with a variable capturing the number of co-morbid psychiatric conditions a subject had experienced in the year prior to study.

Hypothesis Two

The age variable was the only variable of the age, race, and sex variables to be significantly related to rehospitalization experience in the univariate or multivariate models. Older subjects were found to be at significantly less risk for psychiatric rehospitalization in all the models tested. The sex variable possessed a very weak non-significant effect in both the univariate and multivariate models, while the race effect also did not contribute significantly to prediction of psychiatric rehospitalization.

Hypothesis Three

The number of outpatient services received was not related to the rehospitalization experiences of the patients in either the univariate or multivariate case; however, the variety of services received was significant in the univariate models, but became non-significant when psychiatric morbidity and age were controlled. Although ultimately these factors were not significant predictors, it is of interest to note that both the number of services received and variety of services received predicted a greater chance of rehospitalization and shorter community tenure during the study period. This was exactly the opposite of what was expected.

Hypothesis Four

The predicted psychiatric morbidity of patients in the study year was predictive of both rehospitalization and shorter community tenure in both the univariate and multivariate logistic and Cox regression models. As predicted, rising predicted psychiatric morbidity scores were associated with increased risk of rehospitalization during the study year, as well as shorter community tenure between hospitalizations.

Hypothesis Five

Contrary to predictions, the psychiatric medication a subject was prescribed did not exhibit a relationship with the subject's rehospitalization experiences. However, the sum of the number of days of psychoactive medication prescribed did exhibit a moderate effect in the univariate logistic regression case, although no effect was found for the univariate Cox regression when the number of days was entered as a time-dependent covariate. Both of these findings disappeared when entered into multivariate analyses where psychiatric morbidity and age were controlled.

Exploratory Findings

The exploratory Cox regression model found that age, number of CDPS categories and psyh were predictive of community tenure after discharge. The variable prev_hosp or previous hospitalization was no longer significant with a p value above $p < .01$. The number of CDPS categories had the strongest effect in the model followed by the psyh variable. The exploratory logistic regression model was not more effective at modeling the data than was the multivariate logistic regression model.

When the data used were restricted to those subjects who belonged to the two largest plans, the managed care organization the subjects were members of was significantly related to rehospitalization behavior. In the Cox regression model, the variable plan had a moderate effect, although the psyh effect and numcats effect were still the strongest in the model. The prev_hosp or previous hospitalization variable was rendered non-significant in this model. Five variables: plan, psyh, age, prev_hosp, and numcats were used in the model. The previous hospitalization variable was retained in the model only because of its significance in many other studies in the past literature. This variable did not attain the $p < .01$ standard used in this research.

The exploratory logistic regression analysis using only the large plan data contained only three variables, plan, age, and psyh. The psyh variable was the largest contributor to the predictive power of the model. The plan effect was the smallest contributor of the three variables to the power of the model.

In two additional exploratory analyses, one using all of the data, and one using just the large-plan data, a time-dependent variable capturing a subject's aftercare experiences was added. This variable was the most influential variable in both of these models. The addition of this variable did not lead to the deletion of any of the variables in the previous exploratory Cox regression models. In both the Cox regression model using all of the data, as well as the Cox model which used only large-plan data, the hazard for those individuals who received aftercare services after discharge and before readmission was approximately three and one-half times greater than those who did not receive these services.

Additional exploratory analyses using both Cox and logistic regression were performed to study the effect of new-generation anti-psychotic drugs on only those subjects who had a diagnosis of schizophrenia. No differences were found among the patients prescribed different types of new-generation anti-psychotic medications probably due to inadequate power.

EXPLANATION FOR FINDINGS

Hypothesis One

As stated previously, patients who experienced a psychiatric hospitalization episode in the year prior to the study are at higher risk for subsequent rehospitalization and have shorter community tenure after discharge from a psychiatric facility. The reasons for this are complex. Even after controlling for psychiatric morbidity and age, the effect remains significant in most models. This variable was also found to be one of the only consistent predictors of rehospitalization in an influential review of the rehospitalization literature by Klinkenberg and Calsyn (1996). Several authors after this review also found the same effect to be in evidence in their studies (Olfson et al., 1999; Song et al., 1998). This study found that the previous hospitalization variable may in some part be acting as a proxy for psychiatric morbidity.

It is only after the number of CDPS categories variables (numcats) is entered as well as the psychiatric high-cost variable (psyh) that the effect of previous hospitalization diminishes. It stands to reason then that the previous hospitalization variable in this study and past studies was capturing at least some portion of the variability which was subsequently removed by the combination of one or more of these additional variables. From the results of the Cox regression exploratory model, it appears as if the numcats variable has the most explanatory power, followed by the psyh variable. The numcats variable may be functioning to remove the variability in community tenure caused by comorbidity while the psyh variable is removing the variation in community tenure associated with a serious psychiatric diagnosis. The effect of the numcats is illustrated in Figure 10. This figure illustrates the steady progression of hazard going from one CDPS diagnostic category to five CDPS categories. The six CDPS category group had very few subjects and would probably follow the same trend toward decreasing community tenure being associated with a rising number of CDPS categories. The progression as seen in Figure 10 is of note. Over 60 percent of those diagnosed with one CDPS category were still predicted to be in the community after forty weeks while only about 25 percent of those with five categories were still predicted to be in the community at the same interval.

A similar story is told by Figure 9, which displays the estimated time in community of subjects who fell into the CDPS psyh (schizophrenia) category vs. those who did not. In this figure, it can be seen that of those subjects with a CDPS designation of psyh, only about 45 percent can be expected to still be in the community after forty weeks, while fully 65 percent of those subjects without this designation can be expected to be in the community at the same interval.

Hypothesis Two

Age was again the only significant variable of the triumvirate of age, sex, and race demographic variables which were considered in this analysis. Perhaps age has the same palliative effect on rehospitalization as this variable does on studies of prison recidivism. Increasing age was also found to have a prophylactic effect on rehospitalization behavior by other authors (Hoffman, 1994; Huguelet, 1997).

The lack of a significant sex effect was not surprising, as many of the studies in the literature appear split on the significance of this variable to predicting rehospitalization, although several authors did find a significant effect (Polk-Walker et al., 1993; Huguelet, 1997; Hodgson & Lewis, 2001). The relatively large sample size of this study, as well as the control of psychiatric morbidity, may have forced this effect into the background. Perhaps sex was acting as a proxy for psychiatric morbidity in some studies where sex and psychiatric morbidity tended to co-vary.

The lack of a race effect was also not surprising as it was not found to be a consistently strong predictor of rehospitalization in many studies (Klinkenberg and Calsyn, 1996), and only one author in the literature reviewed found race as a robust predicator of rehospitalization behavior (Song et al., 1998). It may be that if more subjects had been introduced to this study, and the race variable were broken out into distinct race categories instead of simply white and non-white, then race may have become a significant variable in this study.

Hypothesis Three

In both the univariate Cox regression model and the univariate logistic regression model, the variety of services delivered to a client in the prior year was significant, while the number of services delivered in the prior year was not. The directionality of the variety of services finding was not as anticipated. It was predicted that a richer variety of services delivered to a patient would reduce the hazard of the patient returning to psychiatric inpatient treatment as several authors have found (Wan & Ozcan, 1991; Solomon, Gordon & Davis, 1984). However, the opposite occurred.

The multivariate case of both Cox and logistic regression models told another story from the univariate models. In all of the multivariate models, the variety of services received was no longer significant. Perhaps this variable was acting as a proxy for psychiatric morbidity in the univariate models. Patients who have more acute psychiatric problems would tend to have more services delivered to them by the Medicaid system. Using the CDPS measures for psychiatric morbidity controlled for the illness burden of the patients, thereby forcing the outpatient services variables from the models.

Hypothesis Four

The univariate and multivariate forms of hypothesis four were supported and had effects that exhibited the directionality anticipated by the author. This finding was not surprising, as many authors had attempted to capture a psychiatric morbidity variable in the past by scoring the presence or absence of various disease categories (Lyons et al., 1997; Olfson et al., 1999; Craig et al., 2000; Hodgson & Lewis, 2001; Song et al., 1998), but none have attempted to score all manifestations of psychiatric disease and distill them to a single score. While the CDPS prospective morbidity score was a strong predictor of both community tenure and the likelihood of a subject being rehospitalized, other variables derived from the CDPS proved ultimately more effective in both the Cox regression and logistic regression models.

As described in the hypothesis one portion of this section, the numcats variable and the psyh variable proved more effective in tandem, than using just the CDPS predicted score. When these variables were entered singly or in tandem to the Cox regression exploratory equation, they rendered the CDPS score no longer significant. The psyh variable is an indicator variable that codes for all forms of schizophrenia and schizoaffective disorder; therefore, the psyh variable can also be considered a schizophrenia indicator variable as well as an indicator of high future psychiatric expenditures. Table 21 in the appendix lists the diagnoses associated with the psyh variable.

Hypothesis Five

Much like hypothesis three, the directionality of the findings of hypothesis five were contrary to expectations. The univariate Cox regression form of hypothesis five did not find a significant relationship between the time-dependent variable days of prescribed medication and the outcome variable; however, the univariate logistic regression found a significant relationship. This may be due to the Cox regression time-dependent variable being a truer approximation of the actual effect of medication compliance than the coarser equivalent of maximum sum of days prescribed used in the logistic regression equation. Several other authors have found a significant relationship between medication compliance and rehospitalization behavior (Haywood et al., 1995; Craig et al., 2000) but these studies did not use the number of days prescribed as a time-dependent covariate in the Cox regression methodology.

Exploratory Analysis

The most interesting contribution of the exploratory analyses to this dissertation was the fact that in both the Cox and logistic regression models the psyh variable either singly, or in combination with the numcats variable, became the proxy for psychiatric morbidity. The CDPS prospective value was forced from the models.

Another interesting point is that when the largest plans were analyzed, the effect of plan membership became significant. The plan effect was approximately the same strength in both the Cox and logistic models with the subjects in one plan being only at 67% of the hazard of being rehospitalized at any one time when compared to the subjects in the other large plan. The logistic regression found that the subjects in one plan were only 61% as likely to be rehospitalized over the course of the study year.

The effect of the aftercare variable also cannot be ignored as it represented a dramatic increase in the hazard of being readmitted for those who received aftercare over those who did not. This is indeed a puzzling finding, and may represents the fact that a portion of the community at risk is more likely to avail themselves of psychiatric and psychological services both in the community as well as in inpatient facilities. This variable may also capture another facet of psychiatric morbidity which is not adequately tapped by the psyh or numcats variables, in that it measures the behavioral health systems responsiveness to the sickest individuals. Perhaps the confluence of willingness to seek services on the part of the client, as well as the added psychiatric morbidity variability captured by the aftercare variable, explains why those receiving aftercare were about three and one half times more likely to be rehospitalized than those who did not.

INTEGRATION OF FINDINGS WITH PAST LITERATURE

Convergent Findings

The age, high-predicted psychiatric cost or schizophrenia effect, as well as the previous hospitalization effect were all found to be in evidence in previous studies. The numcats, or number or degree of psychiatric comorbidity variable did not appear to surface in previous studies. Many of the factors found in this dissertation research that are related to rehospitalization risk appear to support a client vulnerability theory of rehospitalization behavior. Earlier in this paper the Klinkenberg and Calsyn (1996) model of rehospitalization was discussed. In their model, these authors propose a function to predict rehospitalization behavior: Recidvism = f(AC,CV,CS,SR), where AC is receipt of aftercare, CV is client vulnerability, CS is community support, and SR is system responsiveness. According to the present research, this equation does hold some utility as a heuristic for researchers. Admittedly the present study did not endeavor to study the CS or community support part of this equation due to the lack of data capturing subjects' interactions with members of their community in the administrative databases available to the present author.

The SR or system responsiveness part of the equation is most directly addressed by the plancode variable that codes the subject's membership in a particular managed care organization, and the aftercare variable which captures the behavioral health systems' attempts to treat the individual in an outpatient setting. The plancode variable is relevant because managed care organizations by their very nature attempt to manage the healthcare delivery to their members promising better outcomes. Clearly the management approach of one of the plans resulted in their members having a far greater rehospitalization hazard during the study year. Some exploratory ANOVAS were run on length of stay, variety of outpatient visits delivered in the prior year, and the number of outpatient visits delivered in the study year. No differences were found except in the case of variety of visits delivered. In this analysis, the plan that had the lower hazard of rehospitalization actually delivered more variety of services to its members. Thus this effect does not, on the surface, appear to be due to restricted access to care in the plan with the lower hazard of rehospitalization. It should also be noted that the variety of services effect was not present in multivariate analyses so this effect should only be taken as evidence weighing against the restriction of services in the managed care plan with the lower risk of rehospitalization and not evidence of a global effect capturing the protective effect of greater variety of service delivery. ANOVAS run using the variables capturing the variety and number of outpatient services delivered to patients in the study year showed the plans retained the same relative placement to each other in the number and variety of services delivered. Therefore, it can be argued that the managed care plan effect is not due to a sudden reversal in service delivery policy by the plan with the greater hazard or rehospitalization which destabilized its seriously mentally ill patient base forcing many into psychiatric inpatient care.

The AC or aftercare component of the equation is addressed by the aftercare variable used in the exploratory analysis. This variable is somewhat less elegant than the rest of the variables used as it occurs in a contemporaneous fashion with the outcome measure and would therefore be not as useful as predictor variables that are measured the year before the outcome variable is measured. However, this finding still cannot be ignored due to the strength and directionality of the finding. This variable was also constructed using both rehabilitative outpatient care visits (generally the most expensive and reserved for individuals with the highest illness burden), as well as standard office therapy session. Perhaps if these visits were broken into two categories the results would have been different. Dichotomizing these services would also offer a richer portrait of the aftercare environment.

Interpreted using the Polk-Walker framework (Polk-Walker et al., 1993), the findings of this paper can be explained by using a pathway and gatekeeper variables framework. In the Polk-Walker theory, the pathway variables are variables that are demographic in nature, while gatekeeper variables are variables such as hospital admission policy and the diagnosis of the patient that are more procedural in nature (Polk-Walker et al., 1993). It appears clear that age is a pathway variable in this model, but the present author disagrees with Polk-Walker about the placement of diagnostic category in the gatekeeper category. It can be argued that a psychiatric diagnosis can be as much a part of the self as sex and age; however, the diagnosis is assigned by the medical community so it can be seen as a gatekeeper to services, especially if a certain diagnostic threshold must be met to gain services in a particular setting. The plancode effect is clearly a gatekeeper variable according to this heuristic. In the final interpretation of the convergent findings, the Klinkenberg and Calsyn (1996) model of psychiatric rehospitalization appears more complete and serviceable as a model for this particular research.

Some possible explanations for why this study found convergence with many studies in the literature finding the previous hospitalization effect, is that the effect is so very strong that it can be detected with rather small sample sizes. This effect also appears to be a very good proxy for psychiatric morbidity; thus when many studies find this effect, they may be finding the effect for schizophrenia instead. As discussed in the exploratory results section of this paper, the previous hospitalization effect all but vanished when psychiatric morbidity was well controlled by the psyh and the numcats variables. Perhaps if more studies had employed controls for psychiatric morbidity, then the previous hospitalization effect would be found much less often in the literature.

The reason more studies did not find a significant age effect is somewhat mystifying since this effect was so strong in the present study and remained unaffected by the introduction of variables selected to control for psychiatric morbidity. The reason for this dearth of support for the age effect in the literature may be due to sample size. In their 1996 review article, Klinkenberg & Calsyn report that five out of six studies reviewed with at least 350 subjects found that younger patients were more likely to be rehospitalized than older patients.

The effect of a serious psychiatric diagnosis being associated with rehospitalization, expressed in this study as the psyh or schizophrenia effect, found general convergence with several studies (Hodgson et al., 2001; Craig et al., 2000; Olfson et al., 1999; Song et al., 1998; Lyons et al., 1997; Huguelet, 1997; Harwood, 1995); however, of these authors, only Song (Song et al., 1998) specifically mentioned schizophrenia as a reliable predictor of rehospitalization behavior. Hodgson (Hodgson et al., 2001) did mention psychosis as a reliable predictor. The population of Hodgson's psychosis group was in all probability predominantly composed of persons with a diagnosis of schizophrenia. However, some of these studies did cite affective psychosis (Craig et al, 2000) or substance abuse (Olfson et al., 1999; Huguelet, 1997; Harwood, 1995) as a discriminator of rehospitalization behavior, although the present study did not find this effect during the exploratory phase of the analysis.

The lack of more findings in the past literature centering on the question of diagnosis is somewhat puzzling as the introduction of variables describing the patient's condition appear at first examination to be excellent candidates for entry into multivariate models. Also surprising is the lack of more convergent studies that capture co-morbidity. The three studies that found co-morbid effects all found them with co-morbid substance abuse (Huguelet, 1997; Song et al, 1998; Olfson et al, 1999). Again, this co-morbid substance abuse effect was not found in the exploratory analysis that was conducted.

Divergent Findings

Probably the most significant findings of this dissertation research which are at variance with some of the literature is the effect of aftercare as well as the effect of medication compliance. The inability of researchers to find a consistent link between the delivery of aftercare services and the delivery of medication services to the reduction of rehospitalization has profound policy and fiscal implications.

Numerous authors have found that either preadmission care (Dausey, et al, 2002) or aftercare in the community was effective at reducing the rehospitalization experiences of the subjects in their studies (Nelson et al., 2000; Wan & Ozcan, 1991; Caton et al., 1985; Solomon et al., 1984). This present study found only a significant finding using uncontrolled univariate techniques and using aftercare as a time-dependent covariate in Cox regression models, and these findings were opposite in directionality from the authors cited above. Both the univariate Cox regression and the univariate logistic regression models found that subjects who consumed a greater variety of aftercare services in the year prior to study had a significantly increased risk of rehospitalization in the study year. There was no effect found for the number of services delivered to the client. When multivariate techniques were employed, the effect of variety of services vanished, suggesting that the variety of services received was serving as a proxy for psychiatric morbidity. To further complicate matters, when the time-dependent aftercare variable was employed in the Cox regression analysis, a huge effect was discovered that suggests that the use of aftercare is highly related to rehospitalization. The lack of researchers to consistently find a positive aftercare effect using multivariate techniques is somewhat distressing given the amount of time and resources that are undoubtedly spent on delivering aftercare services to clients. The questions remain: 1) what are researchers doing wrong in the process of measuring aftercare to include this effect in statistical models; and 2) if a positive aftercare effect does not exist, then would these aftercare dollars be more appropriately spent on another facet of the Klinkenberg & Calsyn (1996) equation like improving community support for example?

Another divergent finding was that the present researcher did not find a link between the type of antipsychotic prescription drug use and rehospitalization. This finding is not strictly divergent with the literature, as no study the author has found to date takes the position that one of the most frequently prescribed new-generation antipsychotic drugs (Zyprexa, Resperdal, Seroquel) is superior to the others for the purpose of reducing psychiatric rehospitalization. However, there are studies in the literature, which found that the use of a new-generation antipsychotic reduced rehospitalization risk compared to traditional antipsychotic medications (Olfson et al, 1999). The discovery of a link between a particular new-generation antipsychotic drug and increased time in community would be a boon both to states and to the pharmaceutical companies who would love the opportunity to supply the magic fiscal bullet for state Medicaid psychiatric services budgets.

The second pharmaceutical related variable which was explored in this study but had a finding which diverged from some of the past literature was medication compliance as measured by days prescribed. The univariate Cox regression analysis failed to find an effect for medication compliance using medication compliance as a time-dependent covariate. A univariate logistic regression analysis found that the number of days of medication prescribed was positively related to the probability of being rehospitalized, exactly the opposite of what was predicted. This significant effect vanished when a multivariate logistic regression model was employed to control for the effect of psychiatric morbidity, leading the present author to believe that days of medication prescribed was serving as a proxy for psychiatric morbidity.

IMPLICATIONS OF FINDINGS

By far the most important implication of this research is the contribution to the literature of a robust and economical system of risk adjustment for use in studies with psychiatric populations. By "porting" the CDPS risk-adjustment system for use in psychological scholarship from a purely actuarial role, this research has proven that systems such as the CDPS produce meaningful psychiatric morbidity and co-morbidity measures which can be used as effective risk-adjustment tools in large administrative data studies. This research has also shown that the sometimes-present aftercare effect is still elusive to researchers. The most consistent effect in the literature -- that of previous hospitalization -- also has had doubt cast upon it by the inclusion of effective psychiatric morbidity covariates. Perhaps the inclusion of more accurate morbidity measures would see this effect vanish completely from the concern of researchers in the field.

Theoretical Implications

Perhaps the most important theoretical implications of this research are best discussed using the model of Klinkenberg and Calsyn (1996). As stated earlier, the model these authors propose to explain rehospitalization is as follows Recidvism = f(AC,CV,CS,SR). This present research did find an aftercare (AC) component but this effect was measured contemporaneously with the outcome variable, so it is unclear how useful this aftercare variable would be in predicting rehospitalization in an applied setting. The CV or client vulnerability portion of this equation has found strong support by this research and other research (Hodgson et al., 2001; Craig et al., 2000; Olfson et al., 1999; Song et al., 1998; Lyons et al., 1997; Huguelet, 1997; Hoffman, 1994; Polk-Walker et al., 1993; Turner & Wan, 1993). It appears from the current research as well as the convergence of similar findings from previous research, that the diagnosis of schizophrenia or schizoaffective disorder and age of an individual may be critical in the prediction of how well that individual will fare in the community after release from inpatient psychiatric care. These client vulnerability variables should be included in future models of rehospitalization. Another variable which should be considered by future researchers is the numcats variable. This variable captures the number of co-morbid conditions from which a client suffers. This variable exhibited a strong effect in both the Cox and logistic regression models.

The managed care organization a subject belongs to, in the case of studies which look at subjects in managed care, should also be considered. This system responsiveness variable would be of key interest to many state agencies and their contractors when engaging in studies seeking to explore the quality and access to services available in managed care.

What the present research adds to this body of knowledge is a dimension that captures how seriously ill the individual is though the addition of the CDPS prospective measure as well as the number of co-morbid conditions (numcats) measure. These crucial measures appear to have been mostly absent in previous research, and their absence has probably resulted in numerous spurious findings in past studies. In their 1996 study, Klinkenberg and Calsyn admonished readers that the findings in the literature should be considered preliminary due to the failure of many researchers to use multivariate techniques. The present author agrees with this assessment.

By using a system for risk adjustment such as the CDPS, future researchers will be able to more accurately control for psychiatric morbidity. This should enable researchers to test theories of rehospitalization while accurately controlling for the psychiatric morbidity of the subjects in the study. This risk adjustment may allow the effect of system responsiveness, aftercare, and community support to be detected if they exist.

Practical Implications

The results of this study could have important implications for applied researchers using administrative data systems to support behavioral health policy decisions. The availability of a free risk adjustment package such as the CDPS that can be used to support both actuarial research as well as research designed to measure the quality and access to care of Medicaid beneficiaries should be viewed by Medicaid researchers as a favorable development. Although implementation of the CDPS package does require some expertise in SAS programming, and the documentation of the package is somewhat sparse, these are obstacles that can be easily overcome with persistence. The payoff to implementing risk-adjustment software is huge as the face-validity of the findings of risk-adjusted studies instantly increases as the researcher can confidently manage the assertion of providers that "my patients are sicker than his/her patients."

Another practical implication of this research is that if future research into the predictors of rehospitalization uses modern multivariate techniques such as Cox regression, with suitably derived risk-adjustment variables, the real etiology of psychiatric rehospitalization is much more likely to be found. It is very clear from the present author's perspective that research using univariate techniques is just muddying the water, choking the literature with spurious findings of false effects which may be just proxies for psychiatric morbidity or co-morbidity.

LIMITATIONS

Predictor Variables

Perhaps the greatest limitation of this study was the lack of available data that could adequately address the community support aspect of the Klinkenberg & Calsyn (1996) model of psychiatric rehospitalization. Without adequate access to data that capture the community support aspect of the rehospitalization process, there is no sure way to tell how important this facet of the equation is to the overall power of rehospitalization models.

This study used a time-dependent covariate to measure the effect of aftercare and found a significant effect for this variable; however, Foster (1999) also used Cox regression with a time-dependent covariate to represent aftercare and did not find a significant aftercare effect. Besides the ambiguity of the effect as measured by time-dependent variables, the point can be made that there may be little utility in introducing variables into rehospitalization models especially for use in large administrative data systems which use as predictor variables values which occur in the same study year as the rehospitalization episode the model seeks to predict. Although models which use values which occur contemporaneously with the rehospitalization episode will probably possess more explanatory value, these models will also be less useful to administrators who must make decisions from fiscal year to fiscal year based only upon historical data.

The construction of the variables capturing psychiatric medication utilization in days may have been constructed crudely vs. some methods that undoubtedly exist for better capturing the essence of medication compliance. Perhaps this research would have found significant results on the days variable if the psychiatric medications chosen would have been restricted to only new-generation antipsychotic drugs, and if the sample would have been restricted to schizophrenic persons only. Of course, this would have limited the applicability of these results to other more general populations.

Not all data regarding hospitalizations and rehospitalizations were available for this research. Although the author was able to access data which captures the putative principal modality of psychiatric hospitalization and rehospitalization, another avenue of psychiatric inpatient care exists. This avenue exists within the TDMHMR and exists in the form of state hospitals. These data were unavailable to the author. In the author's estimation, this presents a small but not inconsequential threat to the validity of this research. The reason the threat to the validity is small is that the chronic care system in this portion of Texas was designed to divert psychiatric patients from the state hospitals to the inpatient facilities from which the author had access to data. If this system was working as intended, then the vast majority of the inpatient visits should be captured within the data files available to the author.

Data Quality

The findings of this study are built entirely on the assumption that the administrative data collected from the managed care organizations' management information systems apparatus are reliable and valid. In order to study the question of administrative data reliability, several unpublished data validation studies have been prepared by the former External Quality Review Organization (EQRO) contracted to the state of Texas. These studies compared data found in the administrative data of the health plans to data abstracted from medical records by the EQRO staff. In these studies, the conclusion drawn was that administrative data systems mature and produce increasingly reliable data. In state fiscal year

2000, 98% of services found in the medical record were found to be represented in the administrative data; this rate was 84% in 1998 and only 58% in 1997 (THQA, 2001). These data were found to be reliable enough to support research that would be used to make statewide Medicaid policy decisions. It is primarily on the basis of these studies that the foundation of this research can be viewed as sound.

RECOMMENDATIONS

This research has shown that risk-adjustment can be integrated into administrative data systems by employing actuarial systems to health services research ends. In order to promote this cause among health services researchers the following recommendations are offered:

- Designers of administrative data systems used for decision support should employ a risk-adjustment system such as the CDPS and publicize the system's existence and advantages to the users' community in order to enhance the validity of studies completed using these data systems.
- Unadjusted rates should not be accepted as accurate measures of access or quality by researchers or state agencies responsible for measuring quality and access to care.
- Cox regression should be employed by researchers when the dependent variable in question is an interval of time.
- Logistic regression should be employed as a tool to derive risk-adjusted rates of service.

- Studies using univariate techniques where multivariate techniques could be employed should be viewed with suspicion by researchers and administrators.
- Rehospitalization rates and/or community tenure after discharge should not be used as a measure of quality. Rehospitalization, even when psychiatric morbidity and diagnosis are properly controlled, is still a little understood phenomenon.
- Researchers interested in predicting rehospitalization behavior in a subsequent year should employ risk-adjustment systems which have an option to produce prospective estimates of morbidity such as the CDPS. Researchers who are interested in controlling for psychiatric morbidity in a study where the outcome variable and the predictor or risk-adjustment variables are concurrent in nature should employ a risk-adjustment system which produces concurrent estimates of psychiatric morbidity.

SUMMARY, CONCLUSIONS AND FUTURE DIRECTIONS FOR RESEARCH

This dissertation research demonstrated that the Chronic Disability Payment System can be successfully adapted from its primary use as an actuarial tool for use as a risk-adjustment system. This research also showed that increasing age, greater number of co-morbid conditions, the receipt of aftercare, and schizophrenia are all related to increased risk of rehospitalization. The age, number of co-morbid conditions and schizophrenia effects are considered most critical by the author as these variables are measured the year before psychiatric rehospitalization is measured. No significant interaction effects were discovered even though an extensive search for the effects was conducted. The prospective nature of these variables makes their use in predictive models more valuable than the aftercare variable which is measured concurrently with psychiatric rehospitalization. The development of these models that predict psychiatric rehospitalization and community tenure after discharge using administrative claims data should enable administrators of psychiatric systems to do a better job of predicting service utilization in coming years. This should enable more accurate budgeting of resources and staff.

The author hopes that future studies can build upon these findings by adding variables which capture the community support aspect of the Klinkenberg & Calsyn (1996) heuristic. Unfortunately, these variables may be difficult to capture in a large administrative data system without committing some additional resources to the collection of these data. Capturing these variables and making them available to health services researchers who can then employ them in multivariate models should allow researchers to further our knowledge of the etiology of psychiatric rehospitalization.

Future research may also want to address the effect of different types of aftercare upon psychiatric rehospitalization behavior. This study did not explore all of the aspects of the aftercare effect; perhaps future research can explore the subtleties of the effects of day treatment vs. that of outpatient psychotherapy. Interaction effects should also be explored; it may be the case that certain types of aftercare work for particular patients with different mechanisms of support in the community. Researchers also may want to explore why several studies, this one included, find that aftercare is associated with increased rehospitalization even when controlling for psychiatric morbidity. It would be interesting to learn if patients who are not rehospitalized have some previously undetected barriers to access to care which are preventing them from being rehospitalized. These barriers may be systemic in nature, e.g., the policy of a particular managed care organization, or they may be personal, e.g., the perceived stigma of a rehospitalization episode. All of these are interesting questions; unfortunately, these questions will entail collecting more than cheap and plentiful administrative data to answer.

Another avenue which should be pursued by future research is the completeness of the administrative data. As mentioned previously, not all administrative data regarding rehospitalization were available to the author because of the unavailability of certain system data. Currently data systems in many parts of the country are fragmentary or data exist in summary form from some systems and in individual record format from others. In order to get a complete picture of the inpatient and outpatient psychiatric care process, it is now necessary to request and coordinate the transfer of data from many agencies. These data are often in different formats (e.g., SAS, Oracle, VSAM, DB2, SPSS), and usually use different variable names to represent the same data element. Presently it is a daunting task to gather all these data and merge them in order to develop a complete picture of service utilization. Future researchers may wish to pursue federal grants (if available) that will facilitate the construction of regional behavioral health data warehouses which will bring together these disparate forms of information, merge them, format them, and make these data available to researchers and policy makers in a single place. This would greatly facilitate the ability of future researchers to produce valid and timely results regarding the inpatient and outpatient behavioral health apparatus.

Researchers may also wish to explore the causes of multiple rehospitalizations. A macro is included in the appendix which may help in this regard. The current study can be criticized as it does not use all the information available. In the future, this information available on multiple rehospitalizations should be employed to further understanding of the rehospitalization phenomenon.

In closing, it is necessary to reiterate how important it is for researchers to employ some form of risk adjustment when measuring access and quality in managed care plans, community mental health centers and other agencies using sound multivariate techniques. As we forge our way through the 21st century, there should be no excuses for measuring quality in behavioral health care without employing proper risk adjustment to make certain the effects of a treatment or a certain managed care plan actually exist and are not merely a proxy for psychiatric morbidity. As health services researchers, we should honor Codman (Iezzoni, 1997) by employing the best tools at our disposal and not placidly settle for the tyranny of unadjusted rates.

Appendix

Table 21: Description of ICD-9 Diagnosis Codes Used to Designate PSYH Variable

Diagnosis	Description
295	SCHIZOPHRENIC DISORDERS*
2950	SIMPLE SCHIZOPHRENIA*
29500	SIMPL SCHIZOPHREN-UNSPEC
29501	SIMPL SCHIZOPHREN-SUBCHR
29502	SIMPLE SCHIZOPHREN-CHR
29503	SIMP SCHIZ-SUBCHR/EXACER
29504	SIMPL SCHIZO-CHR/EXACERB
29505	SIMPL SCHIZOPHREN-REMISS
2951	HEBEPHRENIA*
29510	HEBEPHRENIA-UNSPEC
29511	HEBEPHRENIA-SUBCHRONIC
29512	HEBEPHRENIA-CHRONIC
29513	HEBEPHREN-SUBCHR/EXACERB
29514	HEBEPHRENIA-CHR/EXACERB
29515	HEBEPHRENIA-REMISSION
2952	CATATONIC SCHIZOPHRENIA*
29520	CATATONIA-UNSPEC
29521	CATATONIA-SUBCHRONIC
29522	CATATONIA-CHRONIC
29523	CATATONIA-SUBCHR/EXACERB
29524	CATATONIA-CHR/EXACERB
29525	CATATONIA-REMISSION
2953	PARANOID SCHIZOPHRENIA*
29530	PARANOID SCHIZO-UNSPEC
29531	PARANOID SCHIZO-SUBCHR
29532	PARANOID SCHIZO-CHRONIC
29533	PARAN SCHIZO-SUBCHR/EXAC
29534	PARAN SCHIZO-CHR/EXACERB
29535	PARANOID SCHIZO-REMISS
2954	AC SCHIZOPHRENIC EPISODE*
29540	AC SCHIZOPHRENIA-UNSPEC

Diagnosis	Description
29541	AC SCHIZOPHRENIA-SUBCHR
29542	AC SCHIZOPHRENIA-CHR
29543	AC SCHIZO-SUBCHR/EXACERB
29544	AC SCHIZOPHR-CHR/EXACERB
29545	AC SCHIZOPHRENIA-REMISS
2955	LATENT SCHIZOPHRENIA*
29550	LATENT SCHIZOPHREN-UNSP
29551	LAT SCHIZOPHREN-SUBCHR
29552	LATENT SCHIZOPHREN-CHR
29553	LAT SCHIZO-SUBCHR/EXACER
29554	LATENT SCHIZO-CHR/EXACER
29555	LAT SCHIZOPHREN-REMISS
2956	RESIDUAL SCHIZOPHRENIA*
29560	RESID SCHIZOPHREN-UNSP
29561	RESID SCHIZOPHREN-SUBCHR
29562	RESIDUAL SCHIZOPHREN-CHR
29563	RESID SCHIZO-SUBCHR/EXAC
29564	RESID SCHIZO-CHR/EXACERB
29565	RESID SCHIZOPHREN-REMISS
2957	SCHIZOAFFECTIVE TYPE*
29570	SCHIZOAFFECTIVE-UNSPEC
29571	SCHIZOAFFECTIVE-SUBCHR
29572	SCHIZOAFFECTIVE-CHRONIC
29573	SCHIZOAFF-SUBCHR/EXACER
29574	SCHIZOAFFECT-CHR/EXACER
29575	SCHIZOAFFECTIVE-REMISS
2958	SCHIZOPHRENIA NEC*
29580	SCHIZOPHRENIA NEC-UNSPEC
29581	SCHIZOPHRENIA NEC-SUBCHR
29582	SCHIZOPHRENIA NEC-CHR
29583	SCHIZO NEC-SUBCHR/EXACER
29584	SCHIZO NEC-CHR/EXACERB
29585	SCHIZOPHRENIA NEC-REMISS
2959	SCHIZOPHRENIA NOS*
29590	SCHIZOPHRENIA NOS-UNSPEC
29591	SCHIZOPHRENIA NOS-SUBCHR
29592	SCHIZOPHRENIA NOS-CHR
29593	SCHIZO NOS-SUBCHR/EXACER
29594	SCHIZO NOS-CHR/EXACERB
29595	SCHIZOPHRENIA NOS-REMISS

```
---------
WLW Macro
---------

%macro wlw(dv=,cv=,evcnt=,id=id,max=,data=_last_,ties=efron);

/*****************************************************************
*******
MACRO WLW uses the method of Wei, Lin, and Weissfeld (Journal
of the American Statistical Association, 1989) to produce
tests and partial likelihood estimates for multivariate survival
data, e.g., repeated events.

The macro has the following parameters:

DV    is the two-part dependent variable, specified in the usual
      way for PROC PHREG.

CV    is a list of covariates (the maximum number is 20).

EVCNT is the name of a variable that contains the number of the
      event in the individual's sequence.

ID    is the name of a variable containing a unique ID for each
      individual (either character or numeric).

MAX   is the maximum number of events among all individuals. If
      you don't wish to analyze the higher order events (because
      of small subsample sizes), just enter a smaller number for
      MAX.

DATA  is the name of data set to be analyzed. The default is the
      last created data set.

TIES  is the method for handling ties (BRESLOW, EFRON, EXACT or
      DISCRETE). The default is EFRON.

Example of usage:

%wlw(dv=duration*censor(0), cv=nit hemo age sex, evcnt=number,
                id=idnum, data=_last_, ties=exact)

Author:  Paul D. Allison, University of Pennsylvania
         allison@ssc.upenn.edu
```

```
Adapted from IML program written by Ying So, SAS Institute.
**************************************************************/

/*Preliminary operations on parameter specifications*/
%do i = 1 %to 20;
%let var&i=%scan(&cv,&i);
%if &&var&i = %then %goto out;
%if %length(&&var&i) > 6 %then %let vv&i=%substr(&&var&i,1,6);
%else %let vv&i=&&var&i;
%end;
%out: %let i=%eval(&i-1);
%let tot=%eval(&i*&max);
%do j=1 %to &i;
%let vdef&j=arr&j(_i_) = &&var&j*(evcnt=_i_);
%let cvd&j= &&vv&j..1-&&vv&j..&max;
%let ar&j=array arr&j(*) &&vv&j..1-&&vv&j..&max;
%end;

/*Create data set with covariate by time interactions*/
data expand;
ncv=&i;
set &data;
evcnt=&evcnt;
%do j=1 %to &i;
&&ar&j;
%end;
do _i_ = 1 to &max;
  %do j=1 %to &i;
    &&vdef&j;
  %end;
end;
run;

/*Do PHREG model for all events*/

model &dv=
%do j=1 %to &i;
    &&cvd&j
  %end;
/ties=&ties;
output out=_out1_ dfbeta=dt1-dt&tot/order=data;
id &id;
strata &evcnt;


/*Sum the DFBETA statistics*/

by &id; run;
proc means data=_out1_ noprint;
```

```
by &id;
var dt1-dt&tot;
output out=_out2_(drop=&id _type_ _freq_) sum=dt1-dt&tot;

/*Calculate the robust covariance matrix, perform tests*/
proc iml;
names={&cv};
nvar=ncol(names);
use _a_;
read all var _num_ into b;
b= b`;
use _out2_;
read all into x;

do i=1 to nvar;
stop=i*&max;
start=stop-&max+1;
bi= b[start:stop];
xi=x[,start:stop];
v= xi` * xi;
iv=inv(v);
se=sqrt(vecdiag(v));
reset noname;
cname= {"Estimate", " Std Error"};
cname2={"Event"};
indx= (1:&max)`;
tmpprt=  bi || se;
print / "Results For" (names[i]);
print indx[colname=cname2] tmpprt[colname=cname format=10.5];
print "Estimated Joint Covariance Matrix",, v;

/* H0: All coefficients equal to 0 */
chisq= bi` * iv * bi;
df= nrow(bi);
p= 1-probchi(chisq,df);
print ,,"Testing H0: no treatment effects for" (names[i]), ,
"Wald Chi-Square = " chisq, "DF = " df,
"p-value = "p[format=5.4],;

/* H0: All coefficients equal to each other */
c=j(df-1,df,0);
do k=1 to df-1;
c[k,k]=1;
c[k,k+1]=-1;
end;
chisq= bi`*c`*inv(c*v*c`)*c * bi;
df= df-1;
p= 1-probchi(chisq,df);
print ,,"Testing H0:Equal Coefficients for" (names[i]), ,
```

```
"Wald Chi-Square = " chisq, "DF = " df,
"p-value = "p[format=5.4],;

/* Assume coefficients equal. Estimate the common value */
nparm=nrow(bi);
e= j(nparm,1,1);
h= inv(e` * iv * e) * iv * e;
b1= h` * bi;
se= sqrt(h` * v * h);
zscore= b1 / se;
p= 1- probchi( zscore * zscore, 1);
print ,,"Estimation of the Common Parameter for" (names[i]),,
"optimal weights = "h,
"Estimate = " b1,
"Standard Error = " se,
"z-score =" zscore,
"2-sided p-value = " p[format=5.4];
end;
quit;
run;
```

References

Allison, P.D. (1984) Event history analysis: Regression for longitudinal event data. Beverly Hills, CA: Sage.

Allison, P.D. (1995) Survival analysis using the SAS system: A practical guide. Cary, NC: SAS Institute, Inc.

Associated Press (2001, June 6). Comptroller Warns of Budget Deficit. *Houston Chronicle*. Retrieved January 17, 2003, from http://www.chron.com/cs/CDA/printstory.hts/metropolitan/934310

Boden, P.J. (2000) Medicaid reform in the 1990s. *Healthcare Financing Review. 22*(2), 1-5.

Caton, C.L.M., Koh, S.P., Fleiss, J.L., Barrow, S., & Goldstein, J.M. (1985) Rehospitalization in chronic schizophrenia. *The Journal of Nervous and Behavioral Disease, 173*(3), 139-148.

Chronic Illness and Disability Payment System. (2002) (Version 2.0) [Computer Software]. San Diego, CA: University of California at San Diego.

Cohen, I.B. (1984) Florence Nightingale. *Scientific American, 250*(11), 128-37.

Craig, T.J., Fenning, S., Tanenberg-Karant, M., & Bromet, E.J. (2000) Rapid versus delayed readmission in first-admission psychosis: Quality indicators for managed care? *Annals of Clinical Psychiatry, 12*(4), 233-238.

Cuffel, B. J., Held, M., & Goldman, W. (2002) Predictive models and the effectiveness of strategies for improving outpatient follow-up under managed care. *Psychiatric Services*, *53,* 1438-1443.

Cumming, R.B., Knutson, D, Cameron, B.A., & Derrick, B. (2002) *A comparative analysis of claims-based methods of health risk assessment for commercial populations.* Minneapolis, MN: Milliman USA. Retrieved January 17, 2003, from http://www.symmetry-health.com/PressReleases/MillimanERG.htm

Dausey, D.J., Rosenheck, R.A., & Lehman, A.F. (2002) Preadmission care as a new behavioral health performance indicator. *Psychiatric Services*, *53*(11), 1451-1455.

Dunkelberg. A. (2002) *Medicaid and state budgets: A case study of Texas.* The Kaiser Commission on Medicaid and the Uninsured. Retrieved on January 3, 2003, from http://www.kff.org/content/2002/20020322/4036.pdf

Fagan, P.J., Schmidt, C.W., & Cook, B. (2002) A model for managed behavioral health care in an academic department of psychiatry. *Psychiatric Services, 53*(4), 431-436.

Foster, M.E. (1999) Do aftercare services reduce inpatient psychiatric readmissions? *Health Services Research*, *34*(3)*,* 715-736.

Frank, R. G., & McGuire, T.G. (1996) Introduction to the economics of behavioral health payment systems. In B. L. Levin & J. Petrila (Eds.), *Behavioral health services: A public health perspective* (pp. 22-37). New York: Oxford University Press.

Frederick, S., Caldwell, K, & Rubio, D.M. (2002) Home-based treatment, rates of ambulatory follow-up, and psychiatric rehospitalization in a Medicaid managed care population. *The Journal of Behavioral Health Services & Research, 29*(4), 466-475.

Geller, J.L. (2000) The last half-century of psychiatric services as reflected in psychiatric services. *Psychiatric Services, 51*(1), 41-67.

Geller, J.L., Fisher, W.H., McDermeit, M., & Brown, J M. (2000) The effects of public managed care on patterns of intensive use of inpatient psychiatric services. *Psychiatric Services, 51*(11)*,* 1385-1391.

Gruber, J. E. (1982) Paths and gates: The sources of recidivism and length of stay on a psychiatric ward. *Medical Care*, *20*(12), 1197-1208.

Haywood, T.W., Kravitz, H.M., Grossman, L.S., Cavanaugh, J.L., Davis, J.M., & Lewis, D.A. (1995) Predicting the "Revolving Door" phenomenon among patients with schizophrenic, schizoaffective, and affective disorders. *American Journal of Psychiatry, 152*(6), 856-861.

Homer, D. W., Lemeshow, S. (1989) *Applied Logistic Regression.* New York: John Wiley & Sons.

Henrdryx, M.S., Dyck, D.G., & Srebnik, D. (1999) Risk-adjusted outcome models for public mental health outpatient programs. *Health Services Research*, *34*(1), 171-195.

Hendryx, M.S., Moore, R., Leeper, T., Reynolds, M., & Davis, S. (2001) An examination of methods for risk-adjustment of rehospitalization rates. *Mental Health Services Research, 3*(1), 15-24.

Hodgson, R. E., Lewis, M. & Boardman, A. P. (2001) Prediction of readmission to acute psychiatric units. *Social Psychiatry and Psychiatric Epidemiology, 36,* 304-309.

Hoffman, H. (1994) Age and other factors relevant to the rehospitalization of schizophrenic outpatients. *Acta Psychiatrica Scandinavica, 89*, 205-210.

Huguelet, S.V. (1997) Factors associated with multiple admissions to a public psychiatric hospital. *Acta Psychiatrica Scandinavica, 95*, 244-253.

Iezzoni, L. I. (1997) Risk and outcomes. In Iezzoni, L. I.(Ed.) *Risk adjustment for measuring healthcare outcomes* (2nd ed., pp. 1-41). Chicago, IL: Health Administration Press.

Jordan, J., Adamo, A, & Ehrmann, T. (2000) Innovations in section 1115 demonstrations. *Health Care Financing Review, 22*(2), 49-59.

Kaiser Commission on Medicaid and the Uninsured. (2002, July 30). *State budgets under stress: How are states planning to reduce the growth of Medicaid costs? Preliminary results based on the Kaiser Commission on Medicaid and the Uninsured 50-state budget survey.* Retrieved January 17, 2003, from http://www.kff.org/content/2002/20020730/statbudupdate73002.pdf

Klinkenberg, D.W., & Calsyn, R.J. (1996) Predictors of receipt of aftercare and recidivism among persons with severe behavioral illness: A review. *Psychiatric Services, 47*(5), 487-496.

Kronick,R, Gilmer, T, Dreyfus, T, & Lee, L. (2000) Improving health-based payment for Medicaid beneficiaries: CDPS. *Health Care Financing Review, 21*(3), 29-64.

Ku, L., Ellwood, M., Hoag, S., Ormond, B., & Woodridge, J. (2000). Evolution of Medicaid managed care systems and eligibility expansions. *Health Care Financing Review. 22*(2), 7-27.

Lutterman, T.C. (1994) The state behavioral health agency profile system. In R. W. Manderscheid & M.A. Sonnerschein (Eds.), *Behavioral Health, United States, 1992*

Lyons, J.S., O'Mahoney, M.T., Miller, S.I., Neme, J., Kabat, J., & Miller, F. (1997) Predicting readmission to the psychiatric hospital in a managed care environment: Implications for quality indicators. *American Journal of Psychiatry*, *154*(3), 337-340.

Merril, J. C. (1985) Growth in national expenditures: Additional analyses. *Health Affairs, 4*(4), 91-97.

Moran, P. W., Doerfler, L. A., Scherz, J., & Lish, J.D. (2000). Rehospitalization of psychiatric patients in a managed care environment. *Behavioral Health Services Research, 2*(4)*,* 191-199.

National Committee for Quality Assurance (2000) *HEDIS 2001: Health Plan Employer Data and Information Set.* Washington, DC: National Committee for Quality Assurance.

National Committee for Quality Assurance (2002) *NCQA finds quality despite chaos.* Retrieved January 3, 2003, from http:www.ncqa.org/communications/publications/ncqa%20update/oct02update.htm#2

Nelson, A.E., Maruish, M.E., & Axler, J.L. (2000) Effects of discharge planning and compliance with outpatient appointments on readmission rates. *Psychiatric Services 51*, 885-889

Olfson, M., Mechanic, D., Boyer, C., Hansell, S., Walkup, J., & Weiden, P.J. (1999) Assessing clinical predictors of early rehospitalization. *The Journal of Nervous and Behavioral Disease, 187*(12), 721-729.

Okin, R.L. (1983) The future of state hospitals: Should there be one? *American Journal of Psychiatry, 140*, 577-581.

Owen, C., Rutherford, V., Jones, M., Tennant, C., & Smallman, A. (1997) Psychiatric rehospitalization following hospital discharge. *Community Mental Health Journal, 33* (1), 13-24.

Polk-Walker, G. C., Chan, W., Meltzer, A.A., Goldapp, G, & Williams, B.(1993) Psychiatric recidivism prediction factors. *Western Journal of Nursing Research, 15*(2), 163-176.

Porter, J.M. (1986) *The rise of statistical thinking.* Princeton, NJ: Princeton University Press.

Rockoff, H. (2002) Price Controls. *The concise encyclopedia of economics.* In The library of economics and liberty. Retrieved on November 29, 2002, from http://www.econlib.org/library/Enc/PriceControls.html

SAS (2000). *Categorical Data Analysis Using Logistic Regression Course Notes.* Cary, NC: SAS Institute Inc.

SAS (2001). Statistical Analysis System (Version 8.2) [Computer Software]. Cary, NC: SAS Institute Inc.

Solomon, P., Gordon, B., & Davis, J. (1984) Differentiating psychiatric readmission from nonreadmissions. *American Journal of Orthopsychiatry, 54*(3), 426-435.

Solomon, S., & Doll, W. (1979) The varieties of readmission: The case against the use of recidivism rates as a measure of program effectiveness. *American Journal of Orthopsychiatry*, *49*(2), 230-238.

Song, L-Y, Biegel, D.E., & Johnsen, J.A. (1998) Predictors of psychiatric hospitalization for persons with serious and persistent mental illness. *Psychiatric Rehabilitation Journal, 22*(2), 155-166. Retrieved January 10, 2003, from EBSCO database.

Stigler, S.M. (1986) *The history of statistics.* Cambridge, MA: Harvard University Press.

Stigler, S.M. (1999). *Statistics on the table.* Cambridge, MA: Harvard University Press.

Sullivan, C. B., & Rice, T. (1991) The health insurance picture in 1990. *Health Affairs*, *10* (2).

Tabachnick, B.G., & Fidell, L.S. (1996) *Using multivariate statistics.* (3rd ed.) New York: HarperCollins College Publishers.

Texas Health and Human Services Commission. (2002) Texas Medicaid in perspective. (4th ed.) Austin, TX: Texas Department of Mental Health and Mental Retardation. Retrieved on January 3, 2003, from http://www.hhsc.state.tx.us/Medicaid/reports/PB/2002pinkbook.html

Texas Health Care Information Council. (2002) Guide to Texas HMO quality: 2002. Technical appendix 2. Retrieved on January 3, 2003, from http://www.thcic.state.tx.us/HMOReports02/Methods%20and%20Statistical%20Issues.pdf.

Texas Health Quality Alliance (2001) SHY 2001 Encounter validation executive summary. Retrieved on October 5, 2003, from http://www.hhsc.state.tx.us/medicaid/mc/about/reports/2001annrpts/Encounter%20Validation%20Exec%20_11_26_01_.pdf

Turner, J. T., & Wan, T.T.H. (1993) Recidivism and mental illness: The role of communities. *Community Mental Health Journal, 29*(1), 3-14.

United States Centers for Medicare and Medicaid Services. (2002) *Historical national health expenditures aggregate, per capita, percent distribution, and annual percent change by source of funds: Calendar years 1960-2001.* Retrieved on November 29, 2002, from http://cms.hhs.gov/statistics/nhe (Data filename: nhegdp01.zip.)

United States Department of Health and Human Services. Report No. SMA 90-1942 (pp.163-207). Washington, DC: Government Printing Office.

Walker, H.M. (1929) *Studies in the history of the statistical method.* Baltimore, MD: The Williams & Wilkins Company.

Wan, T.T.H.,& Ozcan, Y.A. (1991) Determinants of psychiatric rehospitalization: A social area analysis. *Community Mental Health Journal, 27*(1), 3-16.

Zelman, W. A., & Berenson, R. A. (1998.) *The managed care blues and how to cure them*. Washington, DC: Georgetown University Press.

Vita

James Frederick. Lukenbill was born to Shirley and Willis B. Lukenbill on May 27, 1969 in West Monroe, LA. James attended high school in Austin, TX, then attended Austin Community College and The University of Texas at Austin. James earned his B.A. and M.A. at The University of Texas at Austin prior to completion of Ph.D. work. James is currently employed as a Senior Manager for Data/Analytics at Affiliated Computer Services in Austin, TX. He lives with his wife Julie and son Aaron in the Austin area.

Permanent address: 1906 N. Cuernavaca Dr., Apt B, Austin, TX. 78733

This dissertation was typed by the author.