Winter 1999

Classification of Participants Into Two Health Resource Utilization Groups By the Health Enrollment Assessment Review (HEAR) Survey

Jody W. Donehoo
Old Dominion University

CLASSIFICATION OF PARTICIPANTS INTO TWO HEALTH RESOURCE UTILIZATION GROUPS BY THE HEALTH ENROLLMENT ASSESSMENT REVIEW (HEAR) SURVEY

By

JODY W. DONEHOO
B.S.N. May 1981, Oral Roberts University
M.A. May 1987, Regent University

A Dissertation
submitted to the Faculty of Old Dominion University
in Partial Fulfillment of the Requirement for the Degree of

DOCTOR OF PHILOSOPHY

URBAN SERVICES

OLD DOMINION UNIVERSITY
December 1999

Approved by

Approved by
Brenda S. Nichols Director

Betty R. Alexy Member

Christine B. Philpot Member

ABSTRACT

CLASSIFICATION OF PARTICIPANTS INTO TWO HEALTH RESOURCE UTILIZATION GROUPS BY THE HEALTH ENROLLMENT ASSESSMENT REVIEW (HEAR) SURVEY

Jody W. Donehoo
Old Dominion University, 1999
Director: Dr. Brenda Nichols

The study examined the ability of the Health Enrollment Assessment Review (HEAR) survey instrument to classify participants into one of two groups according to utilization of health resources anticipated in the following year. Developed by the U. S. Air Force, the HEAR survey is used worldwide by the Department of Defense for enrollees in TRICARE Prime, the military's adaptation of the HMO model of managed care. Individual HEAR reports are prepared for survey participants and their primary care providers in TRICARE Prime. Although it is currently administered worldwide to a majority of the 8.4 million health care beneficiaries of the Department of Defense, the developers expected the health resource utilization (HRU) measure scored from HEAR survey data to be validated in the future when suitable criterion data became available.

This study estimated the reliability and validity of the original HRU model. Further, an alternate HRU model was derived with optimal use of the data available from the HEAR survey. The original HRU model was based on the Pareto principle, which states that "in any population that contributes to a common effect, a relative few of the

contributors account for the bulk of the effect" (Juran, 1992, p. 57). Alternatively, it is sometimes stated as the 80/20 rule: 20% of the contributors account for about 80% of the common effect (Caldwell, 1994).

The target population for the study was adult active duty family members continuously enrolled in TRICARE Prime in the Hampton Roads metropolitan area of Southeastern Virginia in 1997. The survey was mailed to a random sample that yielded 391 usable surveys. A Pareto analysis revealed that 21.2% of participants utilized 50.4% of the primary care visits. Attempting to identify those participants, the sensitivity (true positive rate) of the original HRU model was 25.3% and the specificity (true negative rate) was 90.9%. The reliability coefficient was .619 and the validity coefficient was .200. The sensitivity of the derived HRU model was 34.9% and the specificity rate was 84.1%. It had a reliability coefficient of .816 and a validity coefficient of .195. Neither model was deemed sufficient to classify members into utilization groups.

This dissertation is dedicated
to my wife who has supported me through the years
and to my two sons who have seen their daddy
working on this doctoral degree their entire lives.

ACKNOWLEDGEMENTS

This study has been sponsored and supported by the United States Navy Bureau of Medicine and Surgery Clinical Investigation Program No. P98-L-H00000-035:A.

The TRICARE Mid-Atlantic Region Office provided database information and supplies for mailing under auspices of the United States Navy Bureau of Medicine and Surgery Clinical Investigation Program No. P98-L-H00000-035:A.

The United States Air Force Office of Prevention and Health Services Assessment provided guidance, technical support, and survey forms.

I extend many thanks to my committee members for their patience and hours of guidance on research and editing of this manuscript. The untiring efforts of my Dissertation Chair deserve special recognition.

TABLE OF CONTENTS

Chapter Page

Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

> The American health system is in a state of flux unlike ever before.... Health care costs are escalating at an alarming rate.... Under all scenarios, managing the triad of access to care, cost of care, and quality of care remains the challenge that we must face (Kongstvedt, 1993, p. xvii).

Faced with that challenge, the Health Maintenance Organization (HMO) approach to managed care emerged to tackle all three components. In HMOs, the cost of care is managed by health plans sharing financial risk and reward with primary care providers through various types of arrangements. This partnership results in the management of health services at the point of service delivery by the provider to the patient. When an individual enrolls into an HMO, he or she is assigned to a primary care provider. The primary care provider manages access to care either by personally delivering all necessary health services or by authorizing a referral to another provider for medically necessary services (usually a specialty provider). Providers and health plans have an interest in managing the cost of care and the quality of care, while beneficiaries also want these managed in their own best interest.

As attempts have been made to control the overall growth of health expenditures in the United States, pressures to decrease the overall Department of Defense budget have also been felt in the Military Health System. Serving the seven uniformed services, the Military Health System provides health benefits to approximately 8.4 million people (Office of the Assistant Secretary of Defense for Health Affairs, 1999). The uniformed

services are: the Army, the Air Force, the Navy, the Marines, the Coast Guard, the commissioned corps of the Public Health Service, and the commissioned corps of the National Oceanic and Atmospheric Administration. Military Health System beneficiaries include active duty service members and their families, retirees and their families, and others eligible to receive health services in military treatment facilities (MTF). Military treatment facilities are the clinics and hospitals operated by the Department of Defense.

Top policy makers in the Department of Defense identified that control of healthcare expenditures was essential to keep military treatment facilities a viable option for many of the Military Health System beneficiaries (Statement on the Status of Military Medicine, 1998). As Acting Assistant Secretary of Defense for Health Affairs, Dr. Edward Martin testified that

> the rapid rise in health care costs and the closure of military bases and their medical facilities required the Department to initiate an intensive reengineering effort to design new ways to provide the military health care benefit.... The ever-increasing demand for health care began to exceed our capacity for providing it and precipitated the greatest peacetime management challenges ever faced by the Department. The TRICARE managed health care system was developed as the Department's response to these challenges. It is our military health strategy to provide comprehensive, cost-effective care to active duty members, their families, and other eligible beneficiaries in all the Uniformed Services (p. 7).

The TRICARE program focuses reengineering efforts in the Military Health System on integrating military treatment facilities operated by the government with health services delivery systems from the civilian health care sector.

The military's version of a health maintenance organization (HMO) was named TRICARE Prime and is a central strategy in the TRICARE program. A majority of the 8.4 million Military Health System beneficiaries became eligible to enroll in TRICARE Prime by the middle of 1998. TRICARE Prime contains all the typical features found in most HMOs except the capitation form of provider reimbursement. Additionally, the TRICARE Prime program includes administration of the Health Enrollment Assessment Review (HEAR) survey at the time of enrollment (Josephs, 1996). Developed by the U. S. Air Force, the HEAR survey is used to report health status information to a TRICARE Prime enrollee's assigned primary care manager (PCM: a primary care provider or practice).

Measures scored from HEAR survey data include preventive health service needs, risk factors, primary care complexity level for primary care manager assignment, and Health Resource Utilization (HRU). The HRU was intended to classify survey participants into one of three potential utilization groups: high, medium, or low. The developers of the HEAR specified that the objective for the HRU measure of the HEAR survey was

> identifying TRICARE enrollees [HMO members] likely to be utilizers of high levels of medical resources or PCM [primary care provider] time. The goal is not, as in many of the studies reviewed, to predict future medical expenditures for enrollees; rather we hope to identify which enrollees are

likely to be members of groups which are (on average) high utilizers (Halpern, Murray, Palmer, Reblando & Rust, 1994, p. 51).

Neither of the two technical reports on the development of the HEAR instrument reported the lines of demarcation or thresholds between the three utilization groups (Halpern et al., 1994; Murray & Halpern, 1996). However, it was discovered that the developers were influenced by the Pareto principle when they considered the utilization of health resources (Halpern et al., 1994; Murray & Halpern (1996); J. Frasier, personal communication, Sept. 1998, K. Sotello, personal communication, July, 1998). It was theorized that a relatively few number of people utilize the bulk of the health resources. Assuming that phenomenon was present among TRICARE Prime enrollees, the HRU measure was developed to distinguish the members of the higher utilization groups from the members of the low utilization group. Subsequently, the Pareto principle will be discussed in more detail.

Theoretical Framework – The Pareto Principle

The theoretical framework selected to guide this research was the Pareto principle as articulated by Dr. Joseph M. Juran. A succinct statement of the Pareto principle by Juran was found in a 1992 quote, "in any population that contributes to a common effect, a relative few of the contributors account for the bulk of the effect" (p. 57). Juran (1975) had originally called it the principle of the vital few and the trivial many, but it became widely known as the Pareto principle. Juran later clarified that he had mistakenly renamed his principle as the Pareto principle. None of the writings by Juran or members of the Juran Institute that were reviewed (Juran, 1964, 1975, 1998b, 1992; Juran, & Gyrna,

1988) emphasize a strict adherence to the percentages observed by Vilfredo Pareto in the distribution of wealth among people. Nonetheless, the Pareto principle has become popularized as the 80/20 rule (Swift, Ross, & Omachonu, 1998). More a popularized statement than a true and accurate statement of Juran's principle, this investigator observed that the 80/20 rule could be stated as follows: about 20% of the contributors to a common effect account for about 80% of the effect. The examples given by Gyrna (1988) and Juran (1964) demonstrate that the 80/20 rule was a generality to be used as a guideline rather than a rigid mathematical rule to be applied strictly. In order to understand the Pareto principle fully, it is necessary to review how it developed out of Juran's experience in business and industry.

As a young engineer in the 1920's, Dr. Juran "observed that ... quality defects are unequal in frequency, i.e., when a long list of defects was arranged in order of frequency, a relative few types of the defects accounted for the bulk of the defectiveness" (Juran, 1975, p. 8). Juran described that phenomenon with the phrase, "the vital few and the trivial many" (p. 8). As a corporate industrial engineer for Western Electric Company in the late 1930's, Juran (1975) was responsible for visiting other companies to exchange experiences in industrial engineering practices. While visiting General Motors Headquarters on one such visit, he stumbled onto the work of the French economist, Vilfredo Pareto. A General Motors representative showed him that the executive salaries at General Motors closely followed mathematical models developed by Pareto that demonstrated unequal or maldistribution of wealth. "Pareto observed that 80 percent of the wealth was owned by only 20 percent of the population" (Swift et al., p. 253).

"By the late 1940s, ... I had recognized the principle of the 'vital few and the trivial many' as a true 'universal,' applicable not only in managerial functions, but in the physical and biological worlds generally" (Juran, 1975, p. 9). It was not Pareto's intent that his mathematical models be applied beyond the economics of wealth. Juran's universal principle of vital few and the trivial many was mistakenly attributed to Pareto by a caption identifying the Pareto principle under graphed curves in examples in the first edition of his *Quality Control Handbook* (Juran, 1975). Juran (1988a) applied the Pareto principle to "identify the 'vital few,' whether customers, customer needs, product features, process features, or inputs. Identification of the vital few helps to assure that resources and attention are concentrated where they will do the most good" (p. 6.20). For example, "as applied to the cost of poor quality, the Pareto principle states that a *few* contributors to the cost are responsible for *the bulk* of the cost" (Gyrna, 1988, p. 22.19).

Juran described the Pareto principle as a universal sorter, "the Pareto principle is a universal for sorting any conglomerate mixture into two neat piles, the vital few and the trivial many" (Gyrna, 1988, p. 45). Gyrna continued to explain the principle by describing the analytical process that has become known as Pareto analysis. A Pareto analysis begins by listing the contributors to a particular effect by type under study. The list of contributors is sorted in declining order of frequency to identify the vital few, approximately 20% of the contributors. It is anticipated that this 20% will be responsible for approximately 80% of the effect. It has been recognized that "in a Pareto analysis, there is an endless variety of sources to consider as 'contributors'" (Gyrna, p. 22.19). Consequently, "the ranking of problem areas using a Pareto analysis of course depends on the criterion used" (Gyrna, p. 22.21). When studying widgets for example, an analysis of

the contributors to manufacturing costs may be different from an analysis of the contributors to manufacturing defects.

Later works by Juran began to replace the term the trivial many with the term the *useful many*, particularly in reference to customers (Juran, 1988b, 1992). Using the Pareto principle to identify the vital few customers by sales volumes, Juran noted that marketers avoid openly minimizing the importance of any of the remaining customers. In 1988 he distinguished the vital few, "each of whom is of great importance to us," from the useful many, "each of whom is of modest importance to us" (p. 26). In 1992, Juran recommended that planning for the vital few proceed on an individual basis in contrast to planning for the useful many that should proceed on a group basis.

In describing how to conduct a conceptual analysis, Wilson (1963) suggested that a conceptual analysis should begin by isolating the questions of concept from other questions. Warning against the temptation to analyze everything, it was suggested that only a few concepts be singled out for special attention and the others be left alone. While the literature did not offer a conceptual analysis of the Pareto principle, it became evident that the theory of the Pareto principle was based on two key concepts that can be isolated for special attention: contributors and common effect. The Pareto analysis technique and the graphical representation of its results were developed in the context of quality management in business and industry (see review of literature, chapter II). As the concepts of contributors and common effect are examined, it is important to remember that the overall purpose of Pareto analysis is to sort by priority for quality improvement initiatives.

From the perspective of quality improvement, the concept of common effect is the outcome of a process that is of interest for potential improvement. The common effect could be sales, costs, defects, complaints, etc. A variety of factors may contribute to producing that outcome or common effect and they are considered to be the population of contributors. However, the Pareto principle holds that each of the contributors in the population do not contribute equally to the common effect. The amount of the common effect is disproportionately distributed among the contributors and that enables the contributors to be divided into two groups. Relatively few of the contributors account for the bulk of the effect and they are called the vital few. The remaining contributors account for a relatively small amount of the effect and are called the trivial many or the useful many. A higher priority for quality improvement initiatives is assigned to the vital few.

In the field of economics, Vilfredo Pareto found the maldistribution of wealth to demonstrate a distribution with 20% of the people holding 80% of the wealth. While this 80%/20% distribution is typically applied to the Pareto principle, examples discussed in the literature review (Chapter II) show that quality improvement initiatives can be successful without strictly adhering to the 80%/20% distribution. This investigator observed that as long as there is a disproportionate distribution of contributors to the effect, the Pareto principle has a practical application. In can help focus quality improvement efforts on the few contributors that can have a disproportionately large effect on the outcome. Focusing on a few rather the whole is less daunting when facing an opportunity for improvement. The resulting benefits can far outweigh the costs of the improvement efforts. A relatively small amount of effort can have a disproportionately

large beneficial effect. Finally, it is not that the useful many (formerly the trivial many) are unimportant, they just are not the target of focus for improvement efforts at the outset of a quality improvement project.

Application of the Pareto Principle to Utilization of Health Resources

Having established a general understanding of the Pareto principle, its application as the theoretical framework for the research will now be explored. The common effect in this study was the utilization of health resources. For purposes of this research, utilization was operationally defined as primary care outpatient visits delivered by an assigned primary care provider to members of TRICARE Prime, an HMO-type health plan. The operational definition chosen for health resource utilization is congruent with the intent of the original HEAR developers who conceptualized health resources as the "levels of medical resources or PCM time" (Halpern et. al., 1994, p. 51).

It is reasonable to anticipate that some persons may utilize more primary care visits on an annual basis than others may. That would result in a disproportionate distribution of visits among the people covered by the HMO. If the Pareto principle applies in that situation, the distribution could be theorized to approach a distribution whereby about 20% of the people could be observed to utilize about 80% of the visits.

The contributors in this study were the TRICARE Prime enrollees. In order to decrease overall primary care visit utilization, it would be most advantageous to focus improvement efforts on the vital few persons who will utilize more than their share of outpatient visits. Prerequisite to those efforts would be identifying the vital few. Further, it could be helpful to identify the vital few prospectively so efforts may be focused on

preventing high utilization before it actually occurs. In this case, the vital few would be the high utilization group. Distinguishing the members of the higher utilization groups from the remainder of the population of HMO members is the purpose of the HRU measure in the HEAR survey instrument. The low utilization group would represent the useful many, which would not be a priority for quality improvement initiatives.

Statement of the Problem

Although the HEAR is being utilized system-wide, neither of the two technical reports on the HEAR development (Halpern et al., 1994; Murray & Halpern, 1996) established that the utilization of health resources actually reflected the Pareto principle. It was apparent that the HRU model was developed from the literature and expert opinion rather than by quantitative methods. The HRU classification model was "developed with the expectation that they [it] would be validated sometime after deployment" (U. S. Air Force Office of Prevention and Health Services Assessment, 1999, p. 3). Without any documentation of reliability and validity data offered in the reports, there was no evidence that the sum score from the original HRU model would correctly classify respondents into utilization categories. The current research was designed to estimate initial reliability and validity data for the original HRU model and evaluate its ability to predict health resource utilization. Further, an additional HRU model would be derived systematically and its performance compared to the performance of the original HRU model.

Significance of the Problem

It is in the interest of any health services organization including the government to optimize the utilization of primary care assets and thereby reduce overall costs for health

services. However, if the HRU classification reported from the HEAR survey is to be used to help optimize utilization, it must be a reliable and valid measure. At the time this research was undertaken, a full-scale implementation of the HEAR survey was already underway by the Department of Defense for a majority of its 8.4 million beneficiaries. However, further validation studies of the HRU component of the HEAR were desired (K. Sotello, personal communication, September 27, 1997). A full-scale validation project for the HRU component of the HEAR had been under consideration by the Department of Defense since its creation, but the inability to obtain a national source of reliable utilization data for a criterion-related validation study was delaying such research.

Research Questions

From the research problem and its background, several research questions emerged.

1. Does a Pareto analysis of utilization of primary care manager office visits produce a meaningful criterion against which to evaluate the ability of the HEAR survey to classify participants into utilization groups?
2. What is the reliability and validity of the original Health Resource Utilization (HRU) model to classify participants into utilization groups?
3. Can a measure of health resource utilization be derived systematically from the HEAR survey data? If so, what is the reliability and validity of the derived Health Resource Utilization (HRU) model to classify participants into utilization groups?
4. Which, if either, of the models is best suited for use with the target population?

Summary

The Department of Defense developed the TRICARE program as the managed care plan for the Military Health System. TRICARE Prime, the military's version of an HMO, is central to the TRICARE program. The Health Enrollment Assessment Review (HEAR) is administered to TRICARE Prime enrollees at enrollment. Among other functions, the HEAR is used to classify participants into utilization groups. Named health resource utilization (HRU), that function of the HEAR is based on the Pareto principle developed by Joseph Juran. It was theorized that about 20% of the TRICARE Prime enrollees would use about 80% of the health resources. In this research, health resource utilization (HRU) was operationally defined as the utilization of primary care visits by TRICARE Prime enrollees.

The problem observed was that although the reliability and validity of the HRU classification had not been established, the HEAR was already being implemented across the Military Health System for a majority of its 8.4 million beneficiaries. The first purpose of the research was to establish the reliability and validity data of the original HRU model. The second purpose was to derive an alternative HRU model systematically and to establish reliability and validity of the derived HRU model. Finally, the performance of the original HRU model and the derived HRU model were compared.

CHAPTER II

BACKGROUND OF THE STUDY

The literature review will begin with the utilization of health resources and predictors of utilization. Next, literature applying the Pareto principle to health services will be reviewed. The literature review will conclude with research relevant to or conducted by the developers of the HEAR survey.

Review of the Literature

Prediction of Health Resource Utilization

The review of national ambulatory utilization rates will be reported first. The literature found using surveys directly to predict utilization was minimal and will be discussed next. A significant amount of literature was found that examined the ability of a variety of variables to predict utilization. Many of the studies of predictor variables used data originally collected by a number of means including surveys. Although much of this literature involved the elderly and Medicare, it was included in the review since its coverage of potential predictor variables was extensive. Literature on other groups such as employer groups was included as well since they were more closely related to the target population in this research.

National ambulatory care utilization statistics provided a benchmark for comparison with utilization rates observed in the study sample in the current research. The overall ambulatory care visit utilization rate in the United States in 1996 was 3.37 visits per person, which was not significantly different from the rate of 3.29 visits per

person observed in 1995 (Schappert, 1997, 1998). In 1996, 82.3% of those visits were to physician offices (2.78 visits), 7.5% were to hospital outpatient departments (0.25 visits), and 10.1% were to hospital emergency departments (0.34 visits; Schappert, 1998). Physician office visits included "visits to private, non-hospital based clinics and health maintenance organizations (HMO's)" (p. 2). Females of all ages had 3.9 visits per person and males of all ages had 2.9 visits per person. For females in 1996, 83.0% of those visits were to physician offices (3.21 visits), 7.9% were to hospital outpatient departments (0.31 visits), and 9.1% were to hospital emergency departments (0.35 visits). For males in 1996, 81.4% of those visits were to physician offices (2.32 visits), 7.0% were to hospital outpatient departments (0.20 visits), and 11.5% were to hospital emergency departments (0.33 visits). Females in the 25 – 44 year-old bracket had 3.6 visits per year while males in the same age bracket had 1.9 visits per year. For both genders combined in the 25 – 44 year-old bracket, 2.77 visits per year were observed. Of those visits 79.8% were to physician offices (2.21 visits), 8.0% were visits to hospital outpatient departments (0.22 visits), and 12.1% were visits to hospital emergency departments (0.33 visits). Note that the visits to physician offices included both primary and specialty care.

Aday, Sellers, and Andersen (1981) claimed that some utilization variables can best be obtained by surveys. In particular, determining those individuals in a local area who have not seen a physician in a certain time period can only be assessed by survey. Total utilization rates, self-care practices, and preventive health care behaviors can also be surveyed across a local population. Underreporting of clinical conditions, particularly less serious conditions, is a limitation of health surveys.

The literature was searched from a variety of angles and two articles were found that researched the use of a survey instrument in predicting utilization. Benjamin-Coleman and Alexy (1999) used the SF-36 survey instrument to predict past hospitalization among low-income rural elderly. The research was a follow-up of an earlier study of participants ($N = 222$) in a mobile health unit project. The survey was administered verbally over the phone in a standard manner to participants from the earlier project who could be reached and who agreed to participate ($N = 80$). Eight scale variables were scored by the SF-36 from the 36 items in the survey. It was found that five variables predicted past hospitalization with structure coefficients greater than 0.30 on a discriminant analysis. "Although not significant, the discriminant function accounted for 17% of the variance for past hospitalization (canonical correlation = 0.41)" (p. 226). Overall, 76% were correctly classified with 94% true positives and 75% true negatives. The research compared with the present research. Both studies had a retrospective research design; a survey was administered to predict a dichotomous variable, past utilization. The literature review revealed a finding that was stated in a form similar to a Pareto analysis, "older individuals represent 12% of the United States population but account for 36% of total personal health care expenditures (AARP, 1997).... 38% of all hospital stays and 48% of all days in hospitals" (p. 223). Benjamin-Coleman and Alexy concluded that the SF-36 could be used to screen members of a variety of populations to identify members at-risk for future hospitalization. No mention was made of other similar research using a survey instrument to predict utilization.

Another example of using surveys to collect data with a potential for utilization studies was Saag et al. (1998). The Aday-Andersen model of health behavior was used as

the conceptual framework to guide a study of utilization. In the model "predisposing, enabling, and need variables, together with health system factors, are hypothesized to explain utilization of health services" (p. 966). A telephone survey was completed with 787 participants out of 4,582 calls (17%) to home-dwelling persons over 65 years old. Participants were mostly women (70%) with a median age of 75 years old. The independent variable was urban or rural place of residence. Dependent health care utilization variables revealed no differences between urban and rural participants with respect to total number of visits in the past year and home care assistance. Rural participants used homemaker services more than urban participants did. No further statistical analyses were reported to explain the reasons for the differences observed. The survey research was used to report the incidence of historical use rather than predict future use.

The Tax Equity and Fiscal Responsibility Act (TEFRA) has authorized at-risk contracts for HMOs for Medicare beneficiaries since 1982 (Ash, Porell, Gruenberg, Sawitz, & Besier, 1989). The policy has been for each at-risk HMO contract to cost 95% of what it would have cost to deliver the required care if the enrollees had chosen to remain in the traditional fee-for-service arrangement of Medicare. The adjusted average per capita cost (AAPCC) pricing methodology was established to operationalize that policy. Under the AAPCC pricing methodology, projected fee-for-service costs are established per capita by county and are adjusted by the individual enrollee variables of age, gender, welfare status, and institutional status of the enrollees.

Studies have been conducted to analyze the adequacy of AAPCC as an actuarial methodology and have discussed its acceptability (Beebe, Lubitz, & Eggers, 1985;

Epstein & Cumella, 1988; Ash et al., 1989; Manton & Stallard, 1992; Gruenberg, Kaganova, & Hornbrook, 1996; and Parente, et al., 1996). These studies can be described as either (a) refinements to the existing AAPCC methodology or (b) proposed alternative pricing methodologies. Studies that have examined the AAPCC methodology have relevance for this research since they attempt to identify variables that could be related to utilization. Utilization in context of the AAPC methodology is operationally defined as total costs.

Newhouse, Manning, Keeler, and Sloss, (1989) studied the variance in health expenditures and determined the amount of the variance that could be predicted. Using the RAND Health Insurance Experiment as the data source, a sample ($N = 3,958$) of individuals in six areas of the nation between 14 years old and 65 years old accounting for a total of 7,690 person years was used. The data source excluded active military, retired military, and veterans with service connected disability. There were 818 person years with any inpatient use. It was concluded that the maximum possible variance in total health care expenditures by an individual that could be explained by any number of variables would be 14.5%. When outpatient expenditures were considered alone, they further concluded that the maximum possible variance that could be explained was 50%. It was explained that inpatient expenditures dominated the variance in total expenditures.

Newhouse et al. (1989) used the maximum explainable variance of 14.5% as a benchmark for comparisons of revisions to the AAPCC formula. In addition to the four AAPCC variables (age, gender, welfare status, and institutional status), four sets of predictor variables were used: dichotomous physiologic health; continuous physiologic health; prior use; and functional status, general health self-perceptions, mental health, and

self-reported chronic diseases. Dichotomous physiologic variables were conditions scored as either absent or present, while continuous physiologic measures were actual continuous measures such as diastolic blood pressure above a specified cutoff point. Results were then calculated as a ratio between the observed R^2 and the maximum possible R^2, (maximum $R^2 = 0.145$) established by Newhouse, et al. The portion of the maximum R^2 represented by the observed R^2 was presented as a percentage and the observed R^2 was listed as well. The AAPCC variables used alone explained 11% of the maximum explainable variance ($R^2 = 0.016$) in total expenditures. Various combinations of the four sets of variables were added to the regression models that included AAPCC variables. Adding the prior use variables resulted in 44% of the maximum ($R^2 = 0.0638$), prior use and dichotomous physiologic resulted in 55% ($R^2 = 0.0798$), prior use and continuous physiologic resulted in 60% ($R^2 = 0.0870$), and all variables resulted in 62% ($R^2 = 0.0900$).

Beebe, Lubitz, and Eggers (1985) sought to refine the AAPCC pricing formula by adding prior-use variables and demographic variables readily available from administrative records. Using Medicare history databases (1975 Health Insurance Master Accretions file and the 1976 Person Summary File), a sample was drawn of Medicare HMO enrollees eligible for Medicare Part A and Part B who were alive on January 1, 1976 ($N = 20,773$). Three of the demographic variables in the original AAPCC formula were used in their first regression model: age, gender, and a proxy measure for welfare status, buy-in of Medicare Part B. The second regression model added the dichotomous predictor variable of hospital use in past year to the three demographic variables in the regression model. The continuous variables, hospital days in the past year and amount

applied toward Part B deductible, were added in the third regression model. The regression coefficients revealed that the demographic model explained 0.6% of the variance and the demographic and hospital use model explained 2.2%. Beebe et al. expressed their concern that a more accurate model might invite HMOs to market more aggressively to potential enrollees who had no prior use thereby "cherry picking" enrollees.

The potential for improving the ability of the AAPCC to predict future costs also received attention from Epstein and Cumella (1988). A Medline search was performed on predictors of utilization for the period from January 1, 1970 to December 31, 1985 to collect articles for a meta-analysis. Selecting articles involving American or Canadian elderly populations (60 years old and older), 42 articles were included in the meta-analysis (34 published articles plus 8 unpublished articles). The articles represented 45 data sets and 49% (22) of the data sets involved self-reported data. They examined each article for measures such as a multiple regression coefficient (R^2) that measured the amount of variance explained by the variables under study.

Six groups of predictor variables emerged from the results reported in the studies: perceived health status, functional health status, prior utilization, clinical descriptors, sociodemographic characteristics, and additional predictors (Epstein, & Cumella, 1988). Clinical descriptors were defined as the presence of clinical conditions and mention was made of 21 different conditions. The additional predictor group included variables used in a small number of the studies such as measures of mental health, social support, and other measures. The investigators took two approaches to compare groups of predictor variables. First, the statistical significance approach revealed that perceived health,

functional health, prior utilization, and clinical descriptors were the best predictors. However, they were wary of that approach and mentioned that with large datasets, predictors can reach statistical significance and yet explain very little between-patient variation. Second, they examined results based on correlations and explained variance (R^2) that was reported in 13 of the 42 studies. Prior utilization was the best predictor of current utilization and explained from 6% to 16% of the total variance. Functional health, perceived health status, and clinical descriptors were also good predictors, but sociodemographic variables and the additional variables were poor predictors. Single questions were almost as predictive as multiple questions for functional health and for perceived health.

Discovering from the literature that "certain types of hospitalizations, irrespective of their current costs, could serve as predictors of future high costs" (p.18), Ash et al. (1989) also sought to further refine the AAPCC. Nine disease classification groups (DCG) were identified using diagnoses classified according to the International Classification of Diseases, Ninth Revision, Clinical Modification. A 5% sample ($N = 38,705$) of all Medicare beneficiaries was drawn the Health Care Financing Administration's Continuous Medicare History Sample database. Six regression models were employed using various combinations of variables. The variables included the four original variables in the AAPCC (age, gender, welfare status, and institutional status) as well as other demographic variables, past hospital utilization variables, past Medicare Part B expenses, and disease classification group variables. While all of the variables were significant, the overall ability to predict individual costs was low with 10% of the

variance explained. The model using only past Medicare Part B expenses had the best predictive power with an R^2 value of 0.085.

Including a dichotomous categorical variable for disability in the AAPCC formula was found to significantly improve the prediction of costs (Manton, & Stallard, 1992). First, the original AAPCC costing methodology was updated for the period from 1974 to 1976 by using the second cycle of the National Long-Term Care Survey data, which was administered in 1984. The sample ($N = 22,674$) included elderly persons on Medicare who were alive on April 1, 1984; there were 8,825 males and 13,849 females. "A person was classified as disabled if he or she currently received either personal or mechanical help in at least one of six ADL, i.e. eating, getting in or out of bed, getting around inside, dressing, bathing, and getting to the bathroom or using the toilet" (p. 122).

Including this simple measure of disability, expenditures were found to be different on simulations of financial losses on a magnitude of approximately three to one between disabled and non-disabled persons (Manton & Stallard, 1992). The investigators discussed their concern about passive selection biases by Medicare HMOs against the disabled; active selection bias is against the law. For instance, Independent Practice Association model HMOs were considered to be unbiased toward attracting HMO enrollees while group model or staff model HMOs tend to be passively biased toward the non-disabled. Disabled persons tended to have an ongoing relationship with a private health care provider and private health care providers were much less likely to be in a group model or staff model. Group model or staff model HMOs tended to attract persons who did not have an ongoing relationship with a health care provider, and those persons tended to be non-disabled.

A social/health maintenance organization (S/HMO) demonstration stimulated the need for another study of the AAPCC Medicare HMO costing methodology (Gruenberg, Kaganova, & Hornbrook, 1996). The S/HMO demonstration was conducted in response to the recognition that the traditional AAPCC methodology paid the HMO more for nursing home enrollees than for "frail individuals who reside in the community and who are classified as 'nursing home certifiable'" (p. 60). In the first phase of the demonstration, the identified costing problem was corrected by simply paying the established nursing home rate for the nursing home certifiable individuals who were able to live in the community. Gruenberg et al. responded to the desire for a more refined costing methodology for the second phase of the demonstration. Data were drawn from the Medicare Current Beneficiary Survey conducted in 1991 and that yielded a sample ($N = 8,592$) of individuals at least 65 years old in the community.

> The second phase of the S/HMO demonstration focuses attention on developing improvements in geriatric care that it is hoped will lead to better management of chronic conditions and prevention of or delay in the onset of disability among aged beneficiaries (p. 61).

As an aside, it was interesting to note that employment status was added to the original four factors in the AAPCC in 1995 to take into account the "working aged" (p. 59).

Six regression models were designed involving the AAPCC demographic variables, self-reported diagnoses, perceived health, activities of daily living (ADL) variables, instrumental activities of daily living (IADL) variables, and prior-use variables (Gruenberg et al., 1996). In the regression model, the demographic-alone model resulted in an R^2 of 0.007 for the variance in individual costs explained by the demographic

variables. Explaining 6% of the variance, the comprehensive model contained demographic, health, ADL, and IADL variables. The prior-use model, which contained all of the variables in the comprehensive model plus the prior-use variables, explained 13.2% of the variance in individual costs. Prior use variables were included for comparison purposes only since they were considered to have drawbacks that rendered them unsuitable for adjusting Medicare HMO costs. The comprehensive model demonstrated good fit between expected and observed costs while the demographic model demonstrated poor fit. Drawing 25 random test samples of 50% of the population and comparing them to the remaining 50% of the population validated the comprehensive model. Overall, Gruenberg et al. recommended the comprehensive model. It was also noted that acceptance by payers of self-reported data for use in payment to HMOs could be a challenge.

Using primary care practice as the unit of observation rather than the beneficiary, Parente et al. (1996) found that higher resource utilization was associated with metropolitan practices, smaller practices, group practices, and with internal medicine specialists compared with general or family practitioners. Claims data were used from the Health Care Financing Administration's National Claims History File for services rendered from July 1, 1990 until June 30, 1991. Excluding Medicare HMO beneficiaries and Medicare beneficiaries who spent more than three months in a nursing home or an institutional setting, the data included 100% of the Medicare beneficiaries 65 years old and older from Alabama, Iowa, and Maryland. Parente et al. considered the sample of beneficiaries to be representative of regions with greater proportions of Medicare beneficiaries.

To approximate the gatekeeper approach employed in HMOs, primary care source (PCS) profiling was used whereby fee-for-service Medicare beneficiaries were assigned to the primary care physician who provided more care than any other primary care physician (Parente et al., 1996). To ensure that the primary care source had experience with elderly patients, "only PCSs [primary care sources] seeing at least 25 Medicare patients and providing at least 2 medical visits during each quarter of the study period were retained in the study population" (p. 26). "Across the three states, 2,973 practices have been linked to 728,181 unique beneficiaries resulting in an overall beneficiary-to-practice ratio of nearly 245 patients per practice" (p. 29). The final sample included beneficiaries from Alabama ($n = 244,479$) in 865 primary care source practices, beneficiaries from Iowa ($n = 244,666$) in 808 primary care source practices, and beneficiaries from Maryland ($n = 239,036$) in 1,300 primary care source practices.

The predictor variables were characteristics of the primary care source practice and the dependent variables involved health care use (Parente et al., 1996). It was found that, compared to rural practices, metropolitan practices were more expensive, generated more referrals, and spent more on laboratory tests. After accounting for case mix, internists (compared to general and family practitioners) provided more resource intensive care, generated more hospital admissions, and had higher per patient utilization. Smaller practices provided more ambulatory care, more inpatient visits, and more intensive treatments. Along with studying primary care source characteristics, Parente et al. also examined a few individual patient characteristics. Higher utilization was associated with increasing age, with male gender, and with increasing number of co-morbidities.

The ability of administrative data and interview data to predict utilization was compared by Roos, Roos, Mossey, and Havens (1988). Claims data were merged with 1971 interview data from the Manitoba (Canada) Longitudinal Study on Aging. The study sample ($N = 3,036$) was obtained by selecting individuals age 65 and older who were interviewed and fully covered in the Canadian national health insurance program from January 1970 through December 1973 (or until death). Administrative and interview predictor variables from the first two years (1970 and 1971) were compared against three dependent variables in the last two years (1972 and 1973) using logistic regression procedures. The dependent variables were death, nursing home admission, and hospital admission.

In the Roos et al. (1988) study, the most relevant dependent variable to the present research was the hospital admission outcome. Two predictor measures, an ambulatory illness scale and a hospital illness scale were developed from claims data and reported in the form of Likert measures. Age, female gender, and both illness scale scores were positively associated with hospital admission while spouse alive was negatively related with hospital admission. Hospitalization was associated with five indicators from the interview data: fair or poor health status, presence of a chest condition (cardiovascular disease), treatment for chest conditions, 1 to 30 hospital days in past year, and more than 30 hospital days in past year. The linear regression demonstrated that the three models were very close in the amount of variance explained by each model: the administrative model ($R^2 = 0.006$), the interview model ($R^2 = 0.005$), and the combination model ($R^2 = 0.007$). However, the explained variance in all of the models was quite low and ranged from 0.5% to 0.7%.

Anticipating an increase in the older veteran population along with more favorable eligibility and access policies, Kosloski, Austin, and Borgatta (1987) studied current service utilization and future service utilization of Veterans Affairs Medical Centers. The data source for the research was a survey conducted by the Louis Harris Company from June 1983 to October 1983 using "a self-weighting, national area probability sample" (p. 834). The Harris survey found 3,886 eligible veterans from a screening of 34,500 households, and the final sample ($N = 3,013$) contained veterans 55 years of age and older, 97% male. About one-third of the participants had received services from the VA system. A factor analysis identified 3 factors from 16 variables involving activities of daily living. Those 3 factors along with 27 other variables were entered into a general linear regression model in a hierarchical manner, first for current use and then for future use. That yielded 13 steps for each model.

The thirty predictors explained 32.2% ($R^2 = 0.322$) of the total variance in current Veteran's Affairs Medical Center use and 17.0% ($R^2 = 0.170$) of total variance of intended future use (Kosloski et al. 1987). In a step-wise regression model for current use, the first step used 13 chronic medical conditions in the model to explain 10.2% ($R^2 = 0.102$). An additive composite measure of 42 other medical conditions provided an additional 1.4% ($R^2 = 0.014$) at the second step of the hierarchical regression model. An additional 2.0% ($R^2 = 0.020$) was explained at the third step by two of the ADL factor variables, (a) personal hygiene and mobility and (b) shopping, cooking, and light housework. After the first step explained 10.2% of the variance, the variables that added the most to the measured variance were having private health insurance ($R^2 = 0.036$ with a negative correlation), eligibility for free VA care ($R^2 = 0.036$), and having applied for a

service-connected disability ($R^2 = 0.066$). "Approximately 56% of the sample indicated that they would be likely to use VA services in the future if the service were half the price charged elsewhere" (p. 839). The variables that added the most to the measured variance in intended future use were past use within one year ($R^2 = 0.053$), black race and lower income ($R^2 = 0.046$), and perceived quality of VA hospitals ($R^2 = 0.046$).

The Kosloski et al. (1987) study was similar to the present research investigating utilization in the Military Health System in that both studied federal health systems where beneficiaries can receive free or minimal cost care. Premiums are either not required or are nominal. Kosloski et al. (1987) found that beneficiaries with other health insurance tend to steer away from hospitals in the VA system. Similarly, it is widely recognized that many Military Health System beneficiaries with other health insurance do not use Military Treatment Facilities.

Cafferata (1987) studied the substitution effect of staying in bed (informal use of health services) instead of seeking professional health services (formal use of health services). The data source used was the 1977 National Medical Care Expenditure Survey, which surveyed 14,000 randomly selected households. The Cafferata study drew persons over 64 years old (n = 4,560) for the study sample. The research used traditional regression analysis procedures, logistic regression analysis procedures, and path analysis procedures. It was found that informal use (restricted activity or bed-disability) was positively related to single marital status, but formal use (physician visit or hospitalization) was not. However, living with others (in contrast to married status alone) was negatively related to formal use of health services. It was concluded that informal use of health services did substitute for formal use of professional health services. That

supported the notion of a substitution effect. In addition, "disability days were largely affected by age (positive), education (negative), and four health status measures: perceived fair/poor health, the presence of a limitation, the presence of a chronic condition, and worry" (p. 616).

Hospital utilization and expenditures was the focus of a study by Buczo (1989). The State Medicaid household sample from the 1980 National Medical Care Utilization and Expenditure Survey was used as the data source. The Medicaid beneficiaries in the study sample (N = 7,643) were continuously enrolled during calendar year 1980, were not institutionalized, and were from one of four states: California, Michigan, New York, and Texas. The average age was 30.22 years old (*S. D.* = 25.82) and 61.7% were female. Hospitalization and expenditures were associated with declining self-reported health status, with increasing age, and with death in the study year. The regression model for probability of hospitalization explained 14% of the variance ($R^2 = 0.140$). For those who had at least one hospitalization during the study year, a second regression model explained 32.8% ($R^2 = 0.328$) of the variance in the number of hospitalizations during the study year. The precipitating condition for the hospitalization was most significant variable in the regression model with cardiovascular conditions being the most important among the precipitants. Death during the study year was also significant. The regression model for hospital expenditures explained 37.3% of the variance ($R^2 = 0.373$) with the number of hospital days as the most important predictor. Further, the regression model for number of hospital days explained 28.4% of the variance ($R^2 = 0.284$) and the presence of cardiovascular diseases was the most significant predictor.

Van Vliet and Van De Ven (1992) studied the capitation methodology developed by the Dutch government and proposed improvements. They compared the Dutch methodology to the AAPCC methodology used in the United States for Medicare HMOs. It was noted that "while it seems generally acknowledged in the U.S. that the adjusters mentioned are insufficient, the Dutch government intends to use just these type of adjusters for capitating all health insurance organizations in the Dutch health care system" (p. 1,035). Also of note, the Van Vliet and Van De Ven study differed from studies on Medicare HMOs in the United States in that people of all ages were covered by the Dutch governmental plan, not just elderly people and certain disabled people as in the U.S. Medicare plan. Different data sources were used for three different research approaches.

The first research approach was described as global capitation adjusted by age, gender, supplementary insurance status, and province (Van Vliet & Van De Ven, 1992). The data source was obtained from the largest private health insurance organization in the Netherlands. The sample included about 35,000 individuals continuously covered for five years from 1976 to 1980. Quoting the Newhouse et al. (1989) research mentioned earlier, Van Vliet and Van De Ven used the theoretical maximum R^2 estimate of 13.8% as the benchmark to compare their R^2 results. The model using age and gender alone explained 2.0% ($R^2 = 0.020$) of the variance in total costs. Adding supplementary insurance increased the variance minimally ($R^2 = 0.023$) and then adding province increased the variance slightly more ($R^2 = 0.024$).

The second research approach included prior costs (Van Vliet & Van De Ven, 1992). About 14,000 individuals who had completed a mailed survey in 1976 were selected from the dataset used in the first research. Variables identified in the first

research provided a baseline ($R^2 = 0.024$) and all were retained in subsequent regression models. Adding prior year total costs improved the explained variance significantly ($R^2 = 0.072$). It was comparable to the model with separate outpatient and inpatient costs ($R^2 = 0.073$) and the model that added four other measures to the outpatient and inpatient costs ($R^2 = 0.074$).

The third research approach added health indicators and used data (1981 and 1982) for about 20,000 respondents to the Health Interview Survey conducted annually by the Dutch Central Bureau of Statistics (Van Vliet & Van De Ven, 1992). Again the variables in the first research were retained in all regression models. Adding employment status and family size improved the explained variance marginally ($R^2 = 0.032$), but adding chronic conditions to that more than doubled the explained variance ($R^2 = 0.071$). Then adding physical impairments and health status raised the explained variance further ($R^2 = 0.109$). The comprehensive model with all variables brought the total explained variance to 11.4% ($R^2 = 0.114$), which is 82.6% of theoretical maximum explainable variance ($R^2 = 0.138$).

Concerned about the amount of time that physicians spend with elderly patients, Radecki, Kane, Solomon, Mendenhall, and Beck, (1988) examined the effect of a number of variables by physician specialty on the amount of face to face time spent with patients. Their data source was from a series of nationwide surveys of physician's professional activities administered by the University of Southern California School of Medicine, Division of Research in Medical Education from 1976 to 1978. The investigators selected data on the four physician specialty areas who had substantial numbers of encounters with elderly patients: general practice ($n = 6,853$ encounters), family practice ($n = 9,181$

encounters), internal medicine ($n = 10,878$ encounters), and cardiology ($n = 2,957$ encounters). Participating physicians used a self-administered log diary survey instrument. Data were analyzed using multiple regression procedures. The predictor variables were classified into two categories: encounter information, and physician and practice characteristics. The encounter information category included gender, severity of primary problem, urgency of primary problem, number of visits for primary problem, number of problems/diagnoses recorded, number of diagnostic tests, and number of therapeutic procedures. The physician and practice characteristics category included number of patients per week, physician age, board certification, group practice, academic practice, region, and physician extender present.

Altogether, the predictor variables explained 29% of variance in physician time for general practice, 21% for family practice, 32% for internal medicine, and 33% for cardiology (Radecki et al., 1988). The encounter variables alone explained 19% of variance for general practice, 14% for family practice, 32% for internal medicine, and 33% for cardiology. Three groupings of patients by age were used to compare time spent with patients: reference group ages 45 to 64, the younger old ages 65 to 74, and the older old 75 and up. Comparing the three age groups by specialty to time spent with patients yielded only one significant difference; general practitioners spent less time with the older old. For all ages taken together, time spent was positively correlated with number of diagnostic tests for all four specialties and to number of problems/diagnoses recorded for all but cardiology. For both general practice and family practice, time spent was positively correlated with severity of primary problem and negatively correlated with number of visits for primary problem. Time spent was positively correlated with urgency

of primary problem for family practice and negatively correlated with female gender for internists.

Borus et al. (1985) studied the relationship between mental health treatment and non-mental health ambulatory care utilization. The study sample was drawn from the Bunker Hill Health Center operated by the Massachusetts General Hospital in Boston a mile away. The Health Center was considered to be a true neighborhood clinic since residents in the immediate neighborhood received the vast majority of the services delivered. The neighborhood was described as "a low-income, working-class community of 16,000 people of primarily Irish-American descent in Boston's Charlestown neighborhood" (p. 575). Of the 8,810 patients enrolled to the clinic before fiscal year 1976, Borus et al. studied "the cohort composed of all 400 patients who received a mental disorder diagnosis in the index year (fiscal year 1978), but not in either of the two preceding years.... to represent a 'new' disorder" (p. 575). (It was not specified in which month the referenced fiscal year started.) The participants ($N = 400$) were grouped by those who had been treated by a mental health specialist ($n = 202$) and those who had not ($n = 198$).

The health resource utilization for each patient was studied for the 24 months preceding an index month and for the 24 months following the index month (Borus et al., 1985). The index month during which the mental disorder was diagnosed was excluded from the analysis to control for peaking of visits during that month. The untreated group started out with significantly more visits in the pre-index months than the treated group. There was a sharp rise in non-psychiatric visits in the three months immediately

preceding the index month with a sharp decline in the three months immediately following the index month. Borus et al. concluded

> (1) specialist mental health treatment has an offset effect on ambulatory utilization and charges for nonpsychiatric health care in patients with a mental disorder diagnosis; and (2) the visits and charges associated with such specialist mental health care overshadow these offset savings in nonpsychiatric care and boost overall (nonpsychiatric plus mental health specialist) utilization and charges of the specialist treated subgroup above that of the subgroup treated solely by the nonpsychiatric physicians....those with less severe mental disorders showed a greater offset effect than the patients with more severe and chronic diagnoses (p. 580).

In short, receiving treatment from a mental health specialist was related to a decrease in non-mental health charges, but the cost of the treatment by a mental health specialist was more than the cost avoided. The end result was higher overall costs.

Diehr, Price, Williams, and Martin (1986) studied predictors of outpatient mental health care utilization. The data source was the Consumer Choice and Cost Containment Study of Washington State employees ($N = 2,304$) followed from July 1979 to December 1980. The employees were enrolled in one of three health plans: a fee-for-service plan (Blue Cross), a consumer-owned prepaid group practice (Group Health Cooperative of Puget Sound), and a prepaid independent practice association (United Healthcare). Any outpatient visit was considered to be a mental health visit if it was associated with a diagnosis from the International Classification of Diseases for psychoses, neuroses,

mental retardation, nervousness and debility, or adverse effects of psychotherapeutics. The predictor variables significantly related to mental health care use included good or fair perceived health status, family size of three or four members, middle income, clerical occupation, and one to two chronic conditions. Gender was not related to mental health care use. From the overall sample, 191 participants received treatment for mental health. Higher mental health use and cost were associated with more education, better perceived health status, and professional or administrative jobs.

Smoking has been implicated in increasing health resource utilization. Vogt and Schweitzer (1985) compared utilization rates of non-smokers, former smokers, and current smokers using the same data used by the McFarland, Freeborn, Mullooly and Pope (1985) study discussed later. It was a large 5% random sample of Kaiser Permanente Northwest members that was selected to maintain detailed records of all medical care contacts. From that larger sample, Vogt and Schweitzer drew a sample ($N = 2,582$) of members who had been continuously enrolled in Kaiser Permanente Northwest from September 1967 through 1974 and who had participated in a household interview in 1971. The study sample contained 41% non-smokers, 25% former smokers, and 34% current smokers. Three age brackets were reported: less than 40 years old (36%), 40 to 64 years old (46%), and 65 years old and older (18%). Inpatient utilization rates were expressed as the number of hospital days and the number of discharges while outpatient utilization rates were expressed as costs that had been adjusted by a revision of the California Relative Value Studies.

Utilization was compared using Z scores with significance levels indicated (Vogt & Schweitzer, 1985). There were no differences between outpatient utilization costs for

current smokers and non-smokers. However, the outpatient utilization for laboratory costs, surgery costs, total costs, and total doctor office visits were higher for former smokers than non-smokers. Former smokers had higher rates of inpatient treatment for ischemic heart disease than non-smokers, but not higher than current smokers. Using step-wise regression procedures, only one model including age and gender indicated a relationship between smoking and utilization. As time since quitting smoking increased for former smokers, hospital discharge rates decreased ($R^2 = 0.065$). It was interesting to note that no differences were found between outpatient utilization by current smokers and outpatient utilization by non-smokers. However, differences were found between outpatient utilization by former smokers and outpatient utilization by non-smokers. Vogt and Schweitzer explained that there were

> indications that smokers are less likely to seek preventive medical care services than are either non-smokers or former smokers. These issues raise questions about the degree to which morbidity differences between smokers and nonsmokers may be counterbalanced by a lower concern for health and a corresponding reduced tendency to seek medical care until it is absolutely necessary" (p. 1065).

Chetwynd and Rayner (1986) studied 978 women in New Zealand who visited one of four general practitioner clinics or one family planning clinic in a three-month period in 1979 and 1980. The sample included women 18 to 60 years old (mean age = 34.7) working twenty hours or less a week. Compared to non-smokers, they found that smokers had significantly more illness episodes, more general practitioner visits, and more hospital admissions. There were no significant differences between smokers and

non-smokers in regards to specialist visits, or emergency admissions. Next, the participants were stratified into three age groups: 18 to 29 year olds, 30 to 44 year olds, and 45 to 60 year olds. There were no significant differences between smokers and non-smokers in the oldest age group in regards to illness episodes, general practitioner visits, and hospital admission, but for the other two age groups, smokers' utilization of health services was significantly higher. "Unfortunately, the sample did not include large enough numbers in the older age group to allow any more detailed examinations.... However, this finding results from the much larger variations in health care experience amongst the older group" (p. 232).

The Body Mass Index (BMI) measurement can be used to assess obesity. Using the BMI, Quesenberry, Caan, and Jacobson (1998) studied the relationship between obesity as measured by the BMI and health resource utilization and costs. The main data source was the membership health survey administered in March 1993 to members of the Kaiser Permanente Medical Care Program in the Northern California Region. An age and gender stratified random sample ($n = 33,888$) was selected with a response rate of 58.2% for the surveys ($n = 19,728$). The final study sample ($N = 17,118$) excluded respondents with missing or illogical data. Ranges of the BMI scores (20 – 24.9, 25 – 29.9, 30 – 34.9, and 35 plus) were used to group the respondents. The 20 – 24.9 BMI group was used as the reference group. The 30 – 34.9 BMI group was considered to be moderately overweight and the 35 plus BMI group was considered to be severely overweight. The average age among the four groups ranged from 48.6 years old for the severely overweight group to 53.9 years old for the 25 – 29.9 BMI group. The annual rate of outpatient visits was directly related to BMI (reason for outpatient visits was not studied).

For all ages, the reference group had 6.09 outpatient visits per year. The moderately overweight group had 7.13 visits per year (17% more) and the severely overweight group had 7.55 visits per year (24% more). Age was found to interact with BMI producing the strongest associations between BMI and outpatient visits for the youngest group (20 – 39 years old) and the middle group (40 – 59 years old) with the relationship disappearing for the oldest group (75 years and older).

Application of the Pareto Principle to Health Care Delivery

Dr. Joseph M. Juran developed the Pareto Principle. Along with Dr. W. Edwards Deming and Philip B. Crosby, Juran, focused attention on quality in business and industry.

> The trio of Deming, Juran, and Crosby are the real leaders – the big three who have achieved guru status and made QC (quality control), TQI (total quality improvement), COQ (cost of quality) and SPC (statistical process control) familiar workplace acronyms. Most other quality improvement programs are generally considered derivatives or combinations of these gentlemen's ideas (Oberle, 1990, p. 47).

Not surprisingly, most of the applications of the Pareto principle in the health services industry were found in literature discussing quality. In particular, Pareto charts or diagrams were mentioned both as an SPC tool (Amsden, Butler, & Amsden, 1991; Carey & Teeters, 1995; Clark, Cushing, & Bredenberg, 1998; Pfadt & Wheeler, 1995) and as a continuous quality improvement (CQI) tool (Fields & Siroky, 1994; Juran, 1994; and Ziegenfuss & McKenna, 1995).

Amsden et al. (1991) developed a simple manual for using SPC tools tailored to service industries. Use of the SPC tools was described in the contexts of quality and continual improvement. "In order to make each service conform to its standards, services are produced through a series of repetitive operations.... These repetitive processes can be monitored and measured using statistical tools" (p. 2). Procedural steps were detailed for each of the SPC tools that required no more than a calculator, graph paper, and/or SPC forms to complete. Several principles underlie SPC. No two services are delivered exactly the same, but it is desirable for the variation between services to be minimal. Variation tends to increase over time, but it can be measured. When measured, a pattern tends to emerge with the distribution of measurements conforming to the normal distribution curve. However, variations due to non-random (assignable) causes tend to distort the normal curve.

SPC aims to identify when assignable causes are operating on a process (Amsden et al. 1991). However, it is impractical to measure everything. The characteristics of the service that are critical to customers need to be determined and the critical operations to create the critical characteristics need to be identified. These processes become the focus of the SPC techniques. In SPC, control charts are used to monitor critical processes over a period of time to identify when they are in control. When assignable causes are operating on a critical process, the process can be examined and the specific assignable causes can be identified. When only random or chance causes are operating on a process, the process is said to be stable and in statistical control. However, when assignable causes are present, the process is unstable and out of control. Although the processes involved in delivering a service may be in control, it does not mean that the service is within

specifications. Specifications reflect the wishes of the designers and ultimately the consumer of the service. In control simply means, "the processes are producing the services as they have been designed to do" (p. 18). Assignable causes typically can be eliminated at the point of detection. In contrast, processes that are in control but are not performing to specification call for a reexamination of the design of the process.

Within the SPC approach to quality management, a Pareto analysis is a problem-solving tool "useful in tracking down the sources of variability that result in special causes of variation" (Pfadt & Wheeler, 1995). It can be used either when service processes are out of control or when they are out of specification (Amsden et al., 1991). It is used to "sort out the really important problems from the more numerous but less important problems" (p. 39). A Pareto diagram (chart) is a graphical representation of the Pareto analysis. Other SPC tools include brainstorming, flow-charts, frequency histograms, cause and effect diagrams (fishbone diagrams), storyboarding, scatter diagrams, and check-sheets.

Several SPC projects used Pareto analyses in conjunction with control charts. Clark et al. (1998) used SPC to monitor trends in trauma mortality at the Maine Medical Center, a 598-bed hospital in Portland. In 1985 the hospital collected data for trauma mortality review retrospectively back to 1975 and began to use a simple method to collect the data concurrently. Cases were excluded that had one or more confounding factors present such as age greater than 80 years or intracranial gunshot wounds since it would be unlikely that the quality of medical care would be implicated as a major determinant of outcome in those cases. Before excluding the cases, it was determined that the overall annual number of trauma deaths was relatively constant averaging 37.5 trauma deaths

annually. After excluding those with confounding factors, 236 patients remained in the sample and represented an average of 10.7 trauma deaths annually. These patients were classified into one of three categories: failure of resuscitation, organ failure, and neurological deterioration. Overall, the number of annual trauma deaths was in control. However, other results were found when the cases were analyzed by general cause of death. Although deaths from resuscitation failure and organ failure were in control, the investigators performed a Pareto analysis on the small number of principal diagnoses for otherwise unassigned deaths. It was evident from the Pareto diagram that head trauma was represented more than twice as frequently as each of the other four causes; 20.0% of the diagnoses for otherwise unassigned deaths were responsible for 52.9% of the trauma deaths.

Fields and Siroky (1994) described a number of the SPC techniques and offered examples to illustrate how health care professionals could use each technique. No detail was given about the source of the data for the examples or whether the data was actual data or hypothetical data created for illustration purposes only. The use of the Pareto chart was illustrated with data about unplanned transfers from inpatient units to the intensive care unit. Of the four units depicted in the chart, it was shown that the telemetry unit alone (25.0% of the units) was responsible for 56.0% of the unplanned transfers. Subsequently, the telemetry unit was identified as the focus of attention. The investigators concluded that a "Pareto chart can target improvement efforts to address issues with the greatest promise for cost-effective, efficient results" (p. 7).

SPC and Pareto analyses were found useful in examining how the entire medication process could be improved at Lutheran Hospital in Park Ridge, IL (Carey &

Teeters, 1995). A cause for concern became apparent when 54 medication errors were reported for the month of October 1991 representing 0.03% of the doses administered during the month. The hospital Pharmacy and Therapeutics committee appointed a medication error subcommittee to study the issue. Having recently completed continuous quality improvement training, the members decided to use SPC techniques. Data were collected for the six-month period from October 1992 through March 1993. It was discovered that most of the errors were associated with intravenously administered medications although intravenous (IV) medications represented only about one-third of the medications administered during the period.

Selecting IV medication errors as the focus of their efforts, the team started exploring the problem with a Pareto analysis (Carey & Teeters, 1995). They discovered 62.1% of all the IV errors were related to two of the seven (28.6%) types of errors studied: wrong IV dose administered and omitting the IV dose. A training module on calculating IV drip rates was developed and initiated in April 1993. Overall improvement observed twelve months later was attributed to reduction in omitted IV doses. However, the training had not been effective in the reducing the number of wrong IV doses. Examining this persisting problem, additional Pareto charts revealed that the main cause of incorrect doses was errors in setting IV pumps and that the main problem drug was Heparin. It was reported that the hospital was planning to purchase new IV pumps and special training programs would be offered with the new pumps. Monitoring of the medication process was ongoing.

Ziegenfuss and McKenna (1995) illustrated the use of ten CQI tools with an example examining bed utilization and timely discharge from a university hospital. The

CQI tools identified were the same as the SPC tools mentioned earlier thereby demonstrating that the terms CQI tools and SPC tools are sometimes used interchangeably. Again, the source of the data was not identified and very little detail of the data was given. A Pareto diagram constructed from a sample ($N = 100$) of patients with delayed discharge releases was able to identify the contributors to the delayed discharges. Two of the five causes (40.0%) displayed were responsible for 65.0% of the delays.

It has been shown in the previous studies that the techniques of CQI or SPC have been used successfully to address problems that span more than one hospital department. D. Juran (1994) described a CQI project undertaken in response to a challenge by the hospital executive vice president. The problem was late inpatient arrival to computerized tomography (CT) scans at Beth Israel Hospital in Boston, a 452-bed teaching and research hospital affiliated with Harvard Medical School. An interdisciplinary team began exploring the problem with Pareto diagrams. It was discovered that 23.1% of the nursing floors were responsible for 59.4% of the delays and that 30.8% of reasons for delays were responsible for 66.0% of the delays. The team considered these results supportive of the 80/20 rule and used them to draw attention to the vital few factors. The team proceeded to flowchart the processes revealed by the Pareto analysis. Several changes had almost immediate success. Remaining delays were examined closely. Within one year of the project, 80% of the patients were arriving within five minutes of their appointment, up from the pre-project rate of 4%.

In summary, the literature reviewed on the use of the Pareto principle in the health care industry demonstrated that the Pareto principle was useful even when the

disproportionate distribution of effects from contributors did not strictly follow an 80%/20% relationship. Table II-1 lists the actual proportions interpreted from the literature reviewed.

Table II-1

Distribution of Contributors and Effects in Literature Describing the Use of the Pareto Principle in the Health Care Industry

Contributors	Effects	Description	Authors
20.0%	52.9%	trauma deaths	Clark et al. (1998)
25.0%	56.0%	unplanned ICU transfers	Fields & Siroky (1994)
28.6%	62.1%	IV medication errors	Carey & Teeters (1995)
40.0%	65.0%	hospital discharge delays	Ziegenfuss & McKenna (1995)
23.1%	59.4%	late CT scan arrival	D. Juran (1994)
30.8%	66.0%	late CT scan arrival	D. Juran (1994)

HEAR Instrument Development and Testing

Much of the work of the HEAR project was performed under government contract by the Battelle Memorial Institute, Centers for Public Health Research & Evaluation, in Arlington, Virginia and was reported in two phases. Phase One reported the literature review, analysis, and recommendations (Halpern et al., 1994). Phase Two reported the development of HEAR instrument and the results of field test (Murray & Halpern, 1996). More recently with the HEAR survey in production use for some time, a high resource utilization validation study was reported (U.S. Air Force Office for Prevention and Health Services Assessment, 1999).

Foundational studies for the HEAR development. According to Murray and Halpern (1996), the HEAR developers relied heavily on expert opinion, Freeborn et al.

(1990), and Yen et al. (1994). Freeborn et al. is discussed first followed by other relevant literature written by Freeborn including some of his earlier works. A review of Yen et al. follows along with other works to which Yen contributed.

The study sample used by Freeborn et al. (1990) consisted of members of the Kaiser Permanente HMO in the northwestern U.S. who were 60 years old or older ($N = 501$) and who had been continuously enrolled in the HMO for the six year period from 1976 to 1981. The sample was drawn from the larger 5% random sample of Kaiser Permanente Northwest members on whom detailed records of all medical care contacts since January 1, 1967 were maintained. HMO members enrolled less than the full six-year period were younger and more likely to be female in comparison to enrollees in the sample. The predictor variables were taken from a mailed survey conducted in 1980 with the sample participants. The sample was categorized into consistently high users ($n = 131$, 26.1%), consistently low users ($n = 120$, 24.0%), and mixed users ($n = 250$, 49.9%). The discussion focused on the high and low users and not the mixed users.

The high users (26.1%) accounted for 51% of ambulatory care contacts and for 47% of the office visits (Freeborn et al., 1990). The high users were more likely to be older, and female. They were more likely to have a regular physician, have more medical and mental health conditions, have arthritis or rheumatism, have high blood pressure, and have a heart condition. Marital status and income were not significant indicators for utilization. For the study sample, "The high users most frequently sought care for chronic diseases with treatable symptoms.... Most contacts (82%) of the high users were for continuing conditions" (p. 534). While few of the participants had any services for conditions with an emotional component, a higher proportion of the high users received

mental health services (5%) than the low users (2%). Poorer perceived health status was found in significantly more of the high users (57%). A number of the indicators in this article were used in the HRU formula. Freeborn et al. concluded that "consistent with past research, we found that a minority of the elderly makes persistently larger demands on the health care system over extended periods" (p. 539). That was consistent with the Pareto principle.

Freeborn also participated in an earlier study that investigated participants grouped by number of doctor outpatient visits annually (McFarland et al., 1985). The sample ($N = 1,401$) consisted of adults continuously enrolled in Kaiser Permanente, Northwest, from January 1, 1967 to September 1973 who participated in a household interview in 1971. The sample was drawn from the same 5% random sample of Kaiser Permanente Northwest members that was used in the Freeborn et al. (1990) study. The sample was divided into eight groups by four age categories and by gender. The age categories were: under 40, 40-49, 50-59, and 60 and older. Visits consumed by each of the eight groups were stratified into three utilization groups by quartiles, low users (bottom quartile), medium users (middle two quartiles), and high users (top quartile). Utilization of doctor office visits tended to remain stable over time. Remaining in the same user group in the following year was observed in 45% of the low users, 57.8% of the medium users, and 53.6% of the high users. Further, 65% of the high users for two consecutive years were high users in the following year, 70% were high users for three years, 76% were high users for four years, and 80% were high users for five years. It was concluded that two or three years of consecutive years of high use were sufficient for consideration as consistently high users.

Taking into account annual use over the almost seven years of the study period mentioned above, participants were finally placed into one of three utilization categories (McFarland et al., 1985). These categories were consistently low users (n =116, 8% of sample, mean age = 52.1), consistently medium users (n = 1,100, 79%, mean age = 50.0), and consistently high users (n = 185, 13%, mean age = 50.3). The consistently high users (13%) utilized 31% of the total doctors office visits with an average of 7.72 visits per year. The consistently medium users (79%) were responsible for 68% of the visits (2.80 visits per year) and the consistently low users (8%) had 1% of the visits (0.33 visits per year). The pattern of utilization by the consistently high users was characterized by visits for chronic treatable diseases with continuing care. Visits by consistently low users were most frequently received for well services and miscellaneous. "Utilization patterns were unrelated to marital status, income, occupation, and perceived social class. Health habits such as smoking history, current drinking practices, and present level of physical activity were also not associated with patterns of utilization" (p. 1,226). Visits by consistently high users were associated with anxiety (44% of visits), depression (22%), and insomnia (21%). Both discriminant function analyses and multiple regression analyses revealed that consistently high users were more likely to be female and older as well as report fair or poor health status and a higher number of physical symptoms (total $R^2 = 0.13$)

The other study that influenced the HEAR developers was Yen et al. (1994). A sample of 7,796 employees who selected the indemnity health plan was drawn from a population of 10,446 employees at a large manufacturing company. The indemnity plan members were selected because claims and encounter data were available while encounter data were not available with the HMO plan members. Participants were

grouped into three approximately equal sized groups by age: under 35 years old, 35 – 45 years old, and 45 years and older. Cost measures were compared across the six single-year periods from 1985 to 1990 inclusively as well as three multiple year periods, 1985 – 1987, 1988 – 1990, and 1985 – 1990. The top 20% of the sample was responsible for 90.9% to 92.6% of the total costs in the individual years from 1985 to 1990. When examining multiple year periods, the top 20% of the research participants was responsible for 82.3% of the costs 1985 to 1987, 84.0% of the costs from 1988 to 1990, and 76.0% of the costs in the six year period from 1985 through 1990. "A small segment of employees dominate the employer-paid medical claims costs" (p. 513). This finding is also consistent with the Pareto principle as stated earlier. Higher costs were associated with single marital status, self-reported medical problems, absenteeism (more than six days), and smoking. Gender was not a significant factor in the costs. The statistical analyses included Spearman's rank correlation and multiple logistic regression models. The maximum variance explained by the variables in the research was 12.8%.

An earlier study by Yen et al. (1992) examined the predictive ability of a number of health related measures on one large manufacturing company's economic costs, both from medical claims and absenteeism costs. The 1,294 employees were covered continuously by the company's traditional health plan from 1986 to 1987 and had completed a health risk appraisal in 1985. Using nine multiple regression models (3 employee groups modeled with 3 outcome cost measures each), selected health measures were able to predict between 12.5% and 22.9% of the dependent economic cost measures. Six health-related predictors that were significantly related to outcome cost measures were age, perception index, personal health problem, self-reported 1984 work absences,

smoking, and drug/medication. The perception index was calculated from selected indicators in the health risk assessment and reflected an overall view on life. Note that physical activity, blood pressure, and cholesterol measures were not significantly related to medical claims costs.

Both Freeborn et al. (1990) and Yen et al. (1994) were used in the formulation of the HEAR survey to identify a number of variables previously used to predict health resource utilization. Additionally, the HEAR developers found support for the concept that a population of high users exists and could be described. The two studies differ in that the Freeborn et al. (1990) research looked at an older population and doctors office visits while the Yen et al. (1994) research investigated a working age population and total costs. Similar to the Freeborn study, the older group in the Yen et al. study was also more likely to be in the high utilization group. These studies found a higher number of medical conditions among the higher utilization groups. In the Freeborn et al. (1990) research, marital status was not significant and female gender was significant while the opposite was true in the Yen et al. (1994) research.

HEAR project phase one. The HEAR was designed to perform three functions listed below:

1. assess preventive care needs,
2. predict high resource and PCM time utilization (HRU), and
3. determine primary care level (PCL) of individual beneficiaries by complexity of their healthcare needs to be considered for PCM assignment (or empanelment) purposes (Halpern et al., 1994; Murray & Halpern, 1996).

The developers described the HEAR as modular in design with three components. Each component represents one of the three functions listed above. With the modular design, the components potentially could be used separately. The three components of the HEAR were reported in separate sections.

The literature on health risk assessments (HRA) was reviewed extensively and interviews with seven experts in the field were reported (Halpern et al., 1994). It was found that "health risk assessments have traditionally been used to analyze the increased risk of morbidity or mortality for individuals based on their sociodemographic, behavioral, and clinical characteristics" (Halpern et al., p. 43). HRAs have been used to describe the health risks of populations as well as individuals. However, Halpern et al. cautioned about the use of existing HRAs with certain subpopulations such as the non-white, young, or elderly since the reliability in those groups has been questioned.

The HEAR project headed in a different direction from the purpose of health risk appraisals. Typically with HRAs, a measure of appraisal age is computed for comparison against the individual's actual age. The HEAR project focused on the need for an instrument for use primarily by primary care managers to manage and coordinate preventive services for their assigned TRICARE Prime enrollees and only secondarily to provide information to the patient (Halpern et al., 1994). The implementation of clinical preventive services in the Military Health System was considered far from satisfactory. Prevention objectives of the U.S. Public Health Service were outlined in the "Healthy People 2000" program. Specific guidance toward implementing those objectives followed in another initiative named "Put Prevention into Practice" (PPIP). Focused on health services providers, PPIP provides several products to facilitate implementation of clinical

preventive services. It was implied that an aspiration of the HEAR survey was to meet the challenges of encouraging providers to implement clinical preventive services with patients in their practices. Halpern et al. concluded,

> Many of the HRAs currently in use in civilian or military settings do not specifically collect information on needs for preventive services. Therefore, none of the instruments in their current forms will be appropriate for use in TRICARE enrollment. In addition, it will be important to use information from the HEAR instrument to collect baseline data related to Healthy People 2000 objectives, HEDIS measures, and TRICARE benefit measures as well as preventive service needs for PCMs (p. 47).

With the HEAR survey, the PCM report was designed in such a manner that indicated actions to take concerning preventive services were readily apparent, particularly those covered as a TRICARE benefit (Halpern et al, 1994). The PCM report could serve as a "tickler list" for preventive services (Appendix A). Similar to laboratory reports, recommended values were presented for preventive service intervals and services that were not current were flagged. The flags would inform the PCM which preventive services were indicated at that time for the beneficiary so that the PCM could encourage the beneficiary to obtain the indicated preventive services.

Another component of the HEAR survey was the Primary Care Level (PCL) measure (Halpern et al, 1994). The PCL was designed to recommend assignments of TRICARE Prime enrollees to health services providers who have the skill level appropriate to the complexity of the enrollees' health services needs. It differs from the

HRU component, which considers the potential demand of enrollees to utilize health resources. Halpern et al. found no literature specifically related to the PCL measure and concluded that nothing similar had been developed previously. The closest related literature was found in two areas: patient classification and comparisons of outcomes according to levels of PCMs such as internal medicine physicians versus family practitioners. The investigators also relied on interviews with selected experts. It was determined that the Ambulatory Care Groups (ACG) might offer some contribution. They were being developed to be the ambulatory equivalent of the Diagnostic Related Groups (DRG) which has been used by the Medicare Prospective Payment System used for inpatient services since the mid-1980's. The investigators recommended that the only way to reach the goal of determining the indicated PCL would be creating a new measure through interviews with a panel of experts and that the literature would be of little help. In effect, HEAR developers would have to create the PCL from nothing.

The final component of the HEAR to be discussed is the Health Resource Utilization (HRU) component (Halpern et al, 1994). The literature review of the HRU component for the HEAR project covered studies that involved identifying both individuals and groups utilizing high amounts of health services. Few studies were found that were related to utilization of PCM time and those studies dealt with length of visits. Halpern et al. also cited Newhouse et al. (1989) who determined that the maximum amount of variance in total resource utilization that could be explained by any methodology was less than 15%. It was concluded that the variability between groups of individuals was dominated by variability within groups of individuals. As part of their subsequent literature review, the investigators specifically looked for amount of explained

variance reported in each source. The literature review conducted by Halpern et al. supported their conclusion that the maximum potential variance that could be explained would be 15%; only two of the studies reported explained variance in excess of 15%.

The literature review Halpern et al. (1994) conducted related to the HRU found studies that could be divided into two categories. One category of studies was related to the general population and the other category was related to special populations. They reported that most studies focused specifically on high resource expenditures. That differed from the goal of the HEAR project which de-emphasized cost measures in favor of classifying participants into utilization groups. Halpern et al. found indication that models existed to predict membership in such high resource utilization groups, but they were considered proprietary by the commercial managed care organizations that developed them. It was intuitive and reasonable to expect that managed care organizations could be the most likely to have the sophisticated information systems and the robust data necessary to develop such models. Further, such models could provide a competitive advantage to the managed care organization in a managed care business environment that is becoming more and more intensely competitive. As a result, they would be quite reluctant to divulge valuable information. That is a potential limitation to any literature review on the subject.

A number of recommendations were offered by Halpern et al. (1994) for developing an instrument to identify members of a high resource utilization group. They identified several types of questions that should be considered: (a) prior utilization, (b) chronic disease checklist, (c) attitudes regarding health and mental health, (d) risk factors, and (e) family status information. The strongest predictors discovered in their literature

review were utilization in the previous year, chronic diseases, and attitudes about health and mental health. They recommended that prior utilization questions be simple. A survey could inquire about hospitalization and ambulatory care visits in the past year. Items regarding restriction of activity such as the number of workdays missed could also be included to provide information on the impact of illness on the beneficiary. Lists of chronic diseases should be limited to a reasonable number and ongoing care for chronic diseases should also be considered.

Halpern et al. also recommended that the survey items should be stated in layman's language, rather than medical jargon. Inquiry regarding attitudes about health and mental health should include health status, satisfaction, stress, and mental health status. While literature they reviewed indicated that the presence of risk factors was of little value in predicting health resource utilization, they considered that risk factors could interact with other information and could provide some additional information since the HEAR survey instrument included them anyway. Finally, the unique lifestyles of military families were considered by many to have a potential impact on health services-seeking behaviors. For example, family separation from the active duty service member as well as separation from the family of origin had been mentioned often according to Halpern et al. Subsequently, collection of some sort of information to capture this uniqueness was recommended.

Contrary to the preventive services assessment portion of the HEAR survey, the implications for the PCM from the HRU measure were not readily apparent. It would be reasonable to assume that the PCM might simply overlook the HRU result while receiving value from another portion of the PCM report concerning preventive services.

The HRU may be more relevant to the PCM as a measure of anticipated workload generated by his or her beneficiary panel at the aggregate level than at the individual beneficiary level.

HEAR project phase two. This phase involved development and field-testing of the HEAR instrument and followed the literature review conducted at Phase One of the HEAR Project. While the report of Phase Two (Murray & Halpern, 1996) detailed the rationale for selection of each group of items for the assessment of preventive services, detail was not provided for selection of items to measure the HRU component. It appeared that the two studies mentioned earlier, Freeborn et al. (1990) and Yen et al. (1994), as well as expert opinion formed the basis of the developed HRU formula.

The field testing of the HEAR survey instrument was performed with convenience sampling ($N = 817$) at four military sites in Texas: Fort Hood (Army), Brooks Air Force Base, Dyess Air Force Base, and Corpus Christi (Navy); (Murray & Halpern, 1996; and J. Frasier, personal communication, Sept. 1998). Participants were described as "largely relatively young, healthy, and active adults" (Murray & Halpern, p.17). However, the HEAR survey was designed for use with all TRICARE Prime enrollees including retirees and their family members. The returned HEAR surveys were scored at a central site. No further details of the sample or the field testing procedures were documented in the report.

Inconsistencies were noted in the report of the field test of the HEAR survey (Murray & Halpern, 1996). It was stated in the narrative of the report that "we used thirteen categories to predict which enrollees were likely to be high resource utilizers" (Murray & Halpern, p. 11). However, the table included in the appendix to the report

revealed that data were collected on seventeen indicators. The indicators in that table were consistent with the indicators in the HRU formula as detailed in the system documentation (Bell et al., 1996). The indicators in the final production HRU formula are listed in Table III-3.

The indicators in the report narrative were compared to the indicators in the actual HRU formula in the system documentation. Age was listed in the narrative, but not included in the HRU formula. The chronic conditions, cancer, neurological diseases, diabetes, liver diseases, kidney diseases, mental disorders, and HIV/AIDS, were mentioned in the narrative, but did not appear in the formula. Gender was not in the narrative, but was found in the HRU formula. Questions of a more general nature regarding mental health were included in the formula than were mentioned in the narrative. The report did not reveal anything that would further explain the apparent discrepancies.

The table in the appendix of the report provided cross-tabulated pilot test results of each HRU indicator with utilization group (Murray & Halpern, 1996). The table reflected three HRU classification groups, rather than the two groups in the initial test HRU formula. It was noted that the totals in the table did not equate to the total number of participants specified in the narrative ($N = 817$ narrative and $N = 796$ table). The three utilization groups which were displayed in the table were low utilization ($n = 716$, 89.9%), medium utilization ($n = 54$, 6.8%), and high utilization ($n = 26$, 3.3%). The table revealed that the indicator for the number of medications was present in neither the high nor the low utilization group and occurred only once in the medium utilization group ($n = 54$). The following indicators were positive only twice among members of the high

utilization group ($n = 26$): work absences, ER visits, and hospital visits. The indicators for heart problems and emphysema were positive only three times for the high utilization group. In contrast, the following indicators were highly represented among the higher utilization groups: mental health (26 in the high utilization group and 49 in medium utilization), stress (25 in high utilization and 42 in medium utilization), family problem (20 in high utilization and 28 in medium utilization), female gender (18 in high utilization and 28 in medium utilization), and single marital status (15 in high utilization and 28 in medium utilization). The following indicators showed up in more than half of the members of the medium utilization group: smoking (33 in medium utilization and 13 in high utilization), and alcohol (30 in medium utilization and 12 in high utilization). Surprisingly, perceived poor health was represented less than expected (11 in high utilization and 10 in medium utilization) while arthritis was quite higher than might be expected in the general population (10 in high utilization and 7 in medium utilization).

As a result of field testing, modifications were made to the HEAR survey form and to the formulas for HRU, PCL, and preventive services (Murray & Halpern, 1996). Two changes were made to the HRU measure. Since all 26 of the participants classified into the high HRU category in the table had positive responses to the mental health indicator, the threshold was raised to two positive responses from the six mental health indicators. The initial test HRU formula apparently classified participants into two utilization groups HRU sum score of six establishing the threshold for high utilization. As a result of the field testing, the formula was modified to classify participants into three utilization groups. The high utilization threshold remained at an HRU sum score of six or more, medium utilization was set equal to an HRU sum score of five, and low utilization

for an HRU sum score of four or less. No explanation for this change in the HRU formula was offered in the report.

The results reported were examined closer by this investigator. The frequencies of positive indicators among the medium and high utilization groups provided information about the contribution of each indicator to the classification ability of the HRU. Positive indicators such as number of medications, work absences, ER visits, hospital visits, heart problems and emphysema that were represented rarely among the two higher utilization groups could be considered for removal from the formula. The following five indicators were highly represented among those two groups and could be considered for retention: mental health, stress, family problem, female gender, and single marital status. The two indicators of perceived poor health and arthritis would warrant additional research along with hypertension and doctor visits.

High resource utilization validation report. This study became available to the investigator about eight months after the last surveys were mailed in the present research. Two of the three functions described in earlier reports (Halpern et al., 1994; Murray & Halpern, 1996) were described as the second objective of the HEAR survey by the U.S. Air Force Office for Prevention and Health Services Assessment (1999). "The second objective involved developing new predictive models for resource utilization and appropriate level of primary care provider. Both models were developed with the expectation that they would be validated sometime after development" (U.S. Air Force Office for Prevention and Health Services Assessment, p. 3). To meet that expectation, the validation study "evaluated the validity of the HRU algorithm in a population of TRICARE Prime enrollees" (U.S. Air Force Office for Prevention and Health Services

Assessment, p. 1). The population from which the study sample was drawn was located in the TRICARE Southwest Region 6 which covers Oklahoma, Arkansas, western Louisiana, and all but the western-most portion of Texas (TRICARE Management Activity, 1999; U.S. Air Force Office for Prevention and Health Services Assessment). The participants in the study sample ($N = 7,596$) were TRICARE Prime enrollees who "completed the HEAR during a four-month period (September 1996 – December 1996) and who maintained a continuous enrollment in TRICARE Region 6 during the succeeding twelve months (October 1996 – December 1997)" (U.S. Air Force Office for Prevention and Health Services Assessment, p. 1). For example, the research period was January 1997 to December 1997 for those who completed the survey in December 1996.

Total cost was used as the measurement of utilization (U.S. Air Force Office for Prevention and Health Services Assessment, 1999). That contrasted with the earlier reports on the HEAR development and testing which did not mention an operational definition for health resource utilization (Halpern et al., 1994; Murray & Halpern, 1996). In the process of designing the validation study, two major challenges encountered were (a) achieving consensus among the investigators involved concerning the operational definition of health resource utilization, and (b) finding a database that contained data on utilization in which the investigators had a reasonable level of confidence in the integrity of the data (K. Sottello, personal communication, July, 1998).

The data source chosen for the cost data in the research was the Department of Defense Corporate Executive Information System (CEIS; U.S. Air Force Office for Prevention and Health Services Assessment, 1999). All but one percent of the treatment purchased in the civilian community had associated costs in CEIS. However, since the

cost data for fifteen percent of the episodes of care treated in military treatment facilities was missing in CEIS, an estimation methodology was used to complete the data. Other estimation models were used to arrive at the cost of care episodes in military treatment facilities.

> The Patient Level Cost Accounting (PLCA) algorithm estimates costs for three types of direct care episodes: 1) inpatient stays, where there are several factors for staffing, physician salary, bed days, and DRG case complexity; 2) same-day surgery, which is based on physician time estimates, work center and MTF; and 3) outpatient visits, which depend only on average pharmacy cost and overall staffing expense for the work center. For example, every patient seen in a family practice clinic would be given the same cost estimate, regardless of diagnosis, level of care, or number of prescriptions. (p. 11)

The estimation models that were detailed support an observation made early in designing the present research that all outpatient visits are typically counted as equal measures in military treatment facilities. The quality of the data was identified as a limitation of the study and a disclaimer was given that no attempts were made to validate the data integrity.

The cost variable was used to classify each participants into one of two cost groups with the high cost group consisting of the 20% of the participants who were responsible for highest costs per person (U.S. Air Force Office for Prevention and Health Services Assessment, 1999). While the HRU measure was designed to classify survey participants into three utilization groups, the investigators combined the medium

utilization group and the high utilization group into a single high utilization group. The cost groups were compared to the HRU groups in a two by two matrix design. Only 41% of the participants classified into the high HRU group were actually members of the high cost group. Rather than presenting the results in the table format used in the study, the results are listed in Table II-2 in the same manner as the classification tables that are used in Chapter IV of this research to facilitate comparison. This investigator calculated the percentages in the resulting table. It was evident from Table II-2 constructed by this investigator that only 13% of the high cost participants were correctly classified by the HRU formula as high HRU.

Table II-2

Classification of HRU Groups Compared to Total Cost in U.S. Air Force Validation Study

Total cost	HRU classification [b]		
of care [a]	low	medium/high	Accuracy (% correct)
low cost	5,788 [c]	284 [d]	95.3% correct [c]
high cost	1,326 [d]	198 [c]	13.0% correct [c]
			78.8% overall correct

Note. $N = 7,596$.
[a] Total cost of care is the dependent variable that was from the Department of Defense Corporate Executive Information System.
[b] HRU classification was determined from the scoring of the original HRU model.
[c] Correctly classified and the last column reflects this as a percentage of the row total
[d] Incorrectly classified.

The U.S. Air Force Office for Prevention and Health Services Assessment investigators (1999) concluded that the "HRU algorithm is not sensitive enough to correctly identify high-cost enrollees. This makes it a poor tool for identifying individuals

for utilization/case management or cost-control interventions targeted at high cost utilizers" (p. 8). The investigators recommended that the HRU measure not be used to identify high cost individuals as planned and the measure should be researched further for use as a risk-adjustment tool and for resource planning uses. Multiple regression techniques should be used to derive a mathematical model for the measure. Further, "changes to the HEAR HRU algorithm should include coding to identify missing and conflicting responses, and produce an 'invalid' HRU outcome" (p. 12) when present.

Summary

Literature was reviewed on the prediction of health service utilization, the use of the Pareto principle in the health services industry, and the development of the HEAR survey. It was discovered that females in 25 – 44 year-old age bracket utilized 3.6 ambulatory visits per year in 1996 (Schappert, 1998). That population most closely reflected the target population in this research. It was also found the SF-36 health survey was useful in predicting past hospital utilization in one study, which incidentally was fairly similar in design to the present research (Benjamin-Coleman & Alexy, 1999).

A number of studies were reviewed that studied the ability of variables to predict utilization. Measures of prior use of health resources were found to be good predictors in a number of studies (Ash et al., 1989; Epstein & Cumella, 1988; Gruenberg et al., 1996; Newhouse et al., 1989; Van Vliet & Van De Ven, 1992). Health status measures were also mentioned as good predictors (Newhouse et al., 1989). Demographic variables were poor predictors used alone (Epstein & Cumella, 1988). Mixed results were found with current smoking as a predictor (Chetwynd & Rayner, 1986; Vogt & Schweitzer, 1985).

Overweight as measured by the body mass index (BMI) was a good predictor at younger ages while its effect disappeared in the oldest age group (Quesenberry et al., 1998). While the literature addressed a number of potential predictor variables in variety of combinations, no consensus was found.

Pareto analyses were found to be useful with a variety of quality improvement initiatives in the delivery of health services. However, none of the studies reported percentages anywhere near a proportion of 20% of the contributors accounting for 80% of the effects (Carey & Teeters, 1995; Clark et al., 1998; Fields & Siroky, 1994; Juran, 1994; Ziegenfuss & McKenna; 1995). One study reported 20.0% of the contributors accounting for 52.9% of the effects (Clark et al., 1998).

The developers examined predictors of health resource utilization addressed in the literature and used expert opinion to formulate a list of predictors for the HRU model (Halpern et al., 1994). When the prototype HEAR survey instrument was field-tested, minor revisions were made that included an expansion of the two utilization groups classified by the HRU model into three groups (Murray & Halpern, 1996). While the final report of the present research was being prepared, the U.S. Air Force Office for Prevention and Health Services Assessment (1999) reported a validation study on the HRU model using total expenditures for the health resources. Only 13% of the members of the high utilization group were classified correctly by the HRU model.

CHAPTER III

METHODS AND PROCEDURES

A descriptive research project was conducted to study the ability of the Health Enrollment Assessment Review (HEAR) to classify participants into two groups, low and high Health Resource Utilization (HRU). There were two purposes for the research. First, the study provided reliability and validity data for the original HRU model with the original scoring procedures as designed by the HEAR developers. Second, all potentially relevant items from the HEAR survey instrument were used along with the PCM visit data to derive a revised HRU model systematically to classify participants into low and high Health Resource Utilization (HRU) groups. The procedures involved administering the HEAR survey to a sample of TRICARE Prime enrollees in the first half of calendar year 1998. The classification results from the original and derived HRU measure were then correlated with the criterion, the actual number of primary care manager (PCM) visits utilized by the participants during the previous calendar year (1997). Using a criterion from the year previous to administration of the survey instrument resulted in a retrospective research design although it is clearly noted by the investigator that the HEAR survey was designed as a prospective instrument. This research was viewed as a pilot study for further larger scale validation studies.

Target Population and Approval to Use Human Subjects

The target population for the research ($N = 15{,}138$) was adult family members (dependents) of active duty service members continuously enrolled in TRICARE Prime in 1997 and assigned to a PCM under the Navy's primary care contract in the Hampton

Roads metropolitan area in Southeastern Virginia. Since the target population was a known size, the sample size required to make inferences about the population could be determined. A suggested sample size table by Krejcie and Morgan (as cited in Issacs and Michaels, 1981) revealed that the minimum number of participants in a random sample required to generalize to a population of 15,000 would be 375 participants, and 377 participants would be required for a population of 20,000. Subsequently, it was determined conservatively by interpolation that a sample size of at least 376 participants would be required for the research. A sample of about five hundred participants was initially targeted. Although the final sample ($N = 391$) was not quite that large, however, it was assumed to be sufficient to generalize to the target population.

Agency Setting of the Target Population

The term Military Health System encompasses all health services delivery systems available to beneficiaries authorized to receive health services from military treatment facilities. TRICARE is the program initiated in the Department of Defense to apply managed care principles commonly found in the civilian health services industry to the Military Health System. Active duty service members are entitled to receive all of their health services from military treatment facilities while all other Military Health System beneficiaries can receive health services from military treatment facilities on a space-available basis. Health services in military treatment facilities are provided with no out-of-pocket expense to the beneficiary except for a nominal per diem flat fee for meals while in an inpatient status. Active duty service members do not have to pay for any health services in any setting including civilian health care settings. For all other Military

Health System beneficiaries, the government shares costs with beneficiaries when they receive care in the civilian health services industry under what was formerly widely known as CHAMPUS, the Civilian Health and Medical Program for the Uniformed Services. The CHAMPUS program was integrated into the TRICARE program that encompasses all health services delivered in the Military Health System, whether provided by providers in a military treatment facility or purchased from TRICARE-authorized providers in the civilian health services industry.

The Department of Defense promotes TRICARE as the triple-option health plan for the entire Military Health System. The first option is the same as the traditional fee-for-service option historically known as CHAMPUS. It remains relatively unchanged and was renamed TRICARE Standard. The second option, the TRICARE Extra option, was built on the CHAMPUS foundation. It was designed to be similar to the Preferred Provider Organization (PPO) approach widely used in the civilian health services industry. TRICARE Extra offers a network of selected TRICARE-authorized healthcare providers who agree to discount their billed charges below the TRICARE Maximum Allowable Charge and accept the discounted rate as payment in full.

The third option of the TRICARE program is called TRICARE Prime. It is the military's adaptation of the HMO approach to managed care. As mentioned earlier, it contains all the typical features of an HMO except the capitation form of provider reimbursement. All active duty service members are considered to be enrolled in TRICARE Prime and are subject to the most of the benefits and rules of the program. However, they might be passively enrolled with no action required on their part. Beneficiaries other than active duty service members who are under 65 years of age may

elect to enroll in TRICARE Prime where it is available by submitting an enrollment application. TRICARE Prime enrollees indicate their preference for a primary care provider to be assigned as their Primary Care Manager (PCM). Serving as their personal primary care provider, a PCM could be an individual provider, a group of providers, or an entire clinic. One of the rules of the program requires TRICARE Prime enrollees to seek all of their non-emergency, non-mental health services from their PCM, unless they are exercising their point of service option. The point of service option allows enrollees to seek care without a referral from their PCM, but then they are responsible for a significantly higher share of the cost. In return for seeking health services from their PCM, primary health services and specialty health services are offered within convenient TRICARE Prime access standards.

Military Health System beneficiaries are grouped into five categories for priority access to services from military treatment facilities (Josephs, 1997). The access priorities specified in the Department of Defense policy memorandum are listed below in declining order of priority:

(a) active duty service members;

(b) active duty family members enrolled in TRICARE Prime;

(c) survivors, retirees, and their family members enrolled in TRICARE Prime;

(d) active duty family members NOT enrolled in TRICARE Prime; and

(e) retirees, their family members, and survivors NOT enrolled in TRICARE Prime (p. 1).

A Department of Defense policy memorandum further details implementation of the public law statutes establishing those priorities for access to military treatment facility services. Beneficiaries who choose to enroll in TRICARE Prime fall into a higher priority group for access to military treatment facilities than those who do not enroll.

The Health Enrollment Assessment Review (HEAR) survey is part of the national TRICARE Prime option. In order to facilitate health assessment of TRICARE Prime enrollees, the Department of Defense policy "establishes the HEAR as *the* [emphasis added] TRICARE health assessment survey instrument.... [with the] expectation that the HEAR be used across the Department of Defense for collecting health assessment data on *all* [emphasis added] our TRICARE enrollees" (Josephs, 1996). The guidance by the Assistant Secretary of Defense for Health Affairs was to implement administration of the HEAR Survey across the entire Department of Defense for all TRICARE Prime enrollees by January 1, 1997. The TRICARE Mid-Atlantic Region was permitted to delay initiation of the HEAR survey until April 1, 1998 when contractor assistance would be available.

Description of the Target Population

The Military Health System serves a variety of people. They include young military recruits, families, and retirees and they may be located almost anywhere in the world. Active duty service members range from the most junior enlisted members to four star admirals and generals. Considering the Military Health System beneficiary population as a whole, it would be reasonable to assume that a considerable amount of random variation could exist among the people in the population. Stated another way, there could be a considerable within-group variation. The HRU measure scored by the

HEAR was designed to classify individuals into one of three HRU categories according to the amount of health resources that they could be anticipated to utilize in the following 12 months. In effect, the HRU was designed to measure the amount of variation between the three utilization categories and that could be referred to as between-group variation. However, when there is a large amount of variation among members within the group, it is more difficult for an instrument to classify a group of individuals into smaller groups. An approach that can be used in this situation is to divide the larger, more heterogeneous group into smaller, more homogeneous groups. By reducing the within-group variation, it would be more likely for the measuring instrument to detect between-group variation.

The priorities for access to military treatment facility care mentioned earlier provided a logical and commonly used method for dividing the military population into several more homogenous sub-populations. The first three beneficiary categories established to specify priorities for access to military treatment facility services were for TRICARE Prime enrollees only:

(a) active duty service members (all are considered enrolled in TRICARE Prime);

(b) active duty family members enrolled in TRICARE Prime; and

(c) survivors, retirees, and their family members enrolled in TRICARE Prime (Josephs, 1997, p. 1).

From personal experience working in the Military Health System, the investigator had observed that the utilization patterns among those three groups of TRICARE Prime enrollees could vary considerably. Due to constraints of time, funding, and available data, only one of the sub-populations was selected for the present research, the active duty

family members enrolled in TRICARE Prime. Although the HEAR survey was designed for use with persons 17 years of age and older, only adults 18 years old and older were included in the target population. As will be mentioned in the discussion about gaining approval to use human subjects, eliminating 17-year-old participants from the research avoided additional measures that would be required to allow minors to participate in the research.

The beneficiary category of adult active duty family members differs from the beneficiary category comprised of retirees, family members of retirees, and survivors of deceased service members. For example, the older group of retirees, family members, and survivors might report more chronic conditions than the younger group of active duty family members. Also, the beneficiary category of active duty family member contains mostly female spouses of male active duty service members. The adult active duty family members are younger and their utilization could be expected to mirror that of the female working-age United States population. Although there have been an increasing number of female active duty service members, the vast majority of active duty service members continue to be male. Subsequently, the vast majority of the active duty family member spouses were female, or wives. The beneficiary category of retirees, their family members, and survivors was more evenly balanced between the genders.

Active duty service members were not included in the target population for the research. Enrollment in TRICARE Prime is mandatory for active duty service members under Department of Defense policy for the TRICARE program. Active duty service members are usually assigned to the PCM that was assigned to provide services for their entire military unit. Serving as a Naval Nurse Corps officer in the reserve component with

former active component service and serving as a civil service program specialist for the TRICARE Mid-Atlantic office, the investigator has observed that utilization among the active duty service member population is unique. Traditionally, service members have not been permitted to stay out of work because of sickness unless they visit their doctor or other health care provider. The provider determines whether any sick days are to be authorized. At the conclusion of the authorized sick period, service members usually have to report to the provider again for permission to return to work. As a result, active duty service members who are in the high utilization group could be individuals with quite minor conditions that would cause very few civilians to seek professional care. Further, service members who become seriously injured or develop chronic conditions are usually given a medical discharge from the military. For active duty service members, the Military Health System could be compared to a very comprehensive occupational health service.

Retirees, their family members, and survivors were not selected for the research because it was found in the government data file that more active duty family members were represented in TRICARE Prime in the Hampton Roads area. When Military Health System beneficiaries turn sixty-five years old, they become eligible for Medicare and generally lose their eligibility to enroll in TRICARE Prime. A similar phenomenon could exist for civilians; HMO enrollees might disenroll from their employer-sponsored HMO when they become eligible for Medicare.

Identification of the target population began with the 79,976 adult active duty family members enrolled in TRICARE Prime as of February 1, 1998 and who were assigned to PCMs operating under a Navy primary care contract in the Hampton Roads

metropolitan area of Virginia. The contract covered the three overlapping catchment areas of the Naval Medical Center Portsmouth, the 1st Medical Group at Langley Air Force Base, and the McDonald Army Community Hospital at Fort Eustis. The Naval Medical Center Portsmouth is the large, tertiary level military treatment facility whose forty-mile catchment area covered the Hampton Roads area of Virginia and a small part of the northeast corner of North Carolina. The Naval Medical Center Portsmouth provides comprehensive specialty services except cardiac surgery. Virtually all of the primary care for the TRICARE Prime enrollees (excluding active duty service members) who were the responsibility of the Naval Medical Center Portsmouth in 1997 was delivered under a Navy primary care contract. The 1st Medical Group is an Air Force military treatment facility with a small inpatient capability north of the Naval Medical Center Portsmouth. McDonald Army Community Hospital, also with a small inpatient capacity, is further north of the 1st Medical Group, but south of Williamsburg. Both the Army facility and the Air Force facility are located in the northern area of Hampton Roads metropolitan area often referred to as the Peninsula.

The Navy primary care contract required PCM services to be offered under the TRICARE Prime program at eight sites located throughout the Hampton Roads area. Three TRICARE Prime PCM sites were located in contractor-procured spaces and five PCM sites were located in government-owned spaces. One site was located in North Hampton Roads inside McDonald Army Community Hospital. The other seven sites were located in South Hampton Roads. The eight PCM sites that were operated under the contract resembled urgent centers or staff model HMO centers and offered convenient evening and weekend hours.

The contractor maintained a database of office visits that was considered to be highly reliable and complete. It included all PCM visits by TRICARE Prime enrollees at those sites. An electronic extract of the PCM visit data was delivered to the Navy monthly. It was imported into a personal computer database and a copy was provided to the TRICARE Mid-Atlantic office. The Navy's copy of the database containing actual PCM visit data was known to lag behind the Navy enrollment database that was used for sampling. The period from January 1, 1997 to December 31, 1997 was chosen for the study because the PCM visit database that included data through December 1997 was available at the time of sampling.

Approval to Use Human Subjects

The study required approval to use human subjects from two organizations. Old Dominion University reviewed the research since it was conducted for the purposes of a doctoral dissertation and the Navy reviewed the research since all participants were beneficiaries of the Military Health System served by the Navy primary care contract. Approval was requested to administer the HEAR survey to the participants and to use a government database file containing data regarding office visits delivered by the primary care contractor. Further, financial support was requested from the Navy for costs associated with paper supplies, duplication, and mailing.

The human subject reviews were conducted in accordance with organizational policies and procedures as well as applicable law. The Director of the TRICARE Mid-Atlantic Regional Office endorsed the Navy application for approval to use human subjects on November 21, 1997 and forwarded it to the Head of the Medical Analysis and

Review Center at the Naval Medical Center Portsmouth. There it was assigned to an Institutional Review Board where it was anticipated to receive an expedited review. An application for human subject review was also submitted to Old Dominion University College Of Health Sciences Institutional Review Board on December 16, 1997. In early January 1998 the Old Dominion University College Of Health Sciences Institutional Review Board referred the study for a university level review. About the same time, the Chair of the Navy Institutional Review Board determined that the study did not qualify for an expedited review and referred the study for a full review by the Navy Institutional Review Board.

In preparation for the Naval Institutional Review Board, Navy personnel recommended several revisions. It was recommended that the age of the participants be raised from seventeen years of age as originally proposed to eighteen years of age. It was reported that Virginia law had a number of additional requirements if minors were used. The marginal benefit of including seventeen-year-old participants was compared against the additional requirements for minors and that recommendation was implemented.

The Naval Institutional Review Board also advised that informed consent could not be obtained as originally proposed. It was explained that informed consent would require the investigator to discuss research participation with participants personally and counsel them on the risks of participation in the research before they completed and signed an informed consent form. As an alternative, it was recommended that a cover letter be used to explain the research and stipulate that return of the survey would be considered consent to participate. The cover letter was designed as recommended and a "Notification of Additional Information – Field Test Research Project" was printed on the

reverse side. The notification included all the information that would have been included on an Informed Consent Form if one was used. The cover letter and the notification are enclosed in Appendix B. Revisions to the Naval application were submitted January 9, 1998 and corresponding revisions were submitted to the Old Dominion University Human Subjects Review Board on January 19, 1998.

The Old Dominion University Human Subjects Review Board recommended that the computer file containing the visit data and the survey instruments be randomly coded. Then, when the surveys were returned, the visit data and the survey data would be matched using the codes and the links destroyed thus ensuring the anonymity of the data used for analysis. The review board applications were revised in accordance with the recommendation that data be linked by a database management system and then, all fields with individually identifying information would be excluded during the importing procedure. The importing routine would in effect segregate the data collection phase from data anaiyses phase and render the data anonymous in the process. Survey forms would then be shredded after ensuring data integrity. Recommended revisions were made to the Old Dominion University Human Subjects Review application and submitted February 10, 1998 then forwarded for consideration with the Navy application.

On February 17, 1998, the Old Dominion University Human Subjects Institutional Review Board found the study exempt from review under 45 Code of Regulations 46.101(b)(2)(i) on the condition that the data would be rendered free of identifiers before analysis. Documentation of the university approval is included in Appendix C. The Commander, Naval Medical Center Portsmouth granted approval on February 27, 1998 and identified the study as the Naval Bureau of Medicine and Surgery Clinical

Investigation Program Number P98-L-H00000-35:A. The memorandum dated March 4, 1998 documenting Navy approval is also enclosed in Appendix C. The Navy memo specified that "presentations and/or publications resulting from your study shall acknowledge and identify the [Navy] Bureau of Medicine and Surgery Clinical Investigation Program as the sponsor of your study" (p. 3). The approval by the Navy indicates that the proposed study was found to be in compliance with Department of Defense requirements as well as in specific compliance with all applicable federal laws including the Code of Federal Regulations (CFR).

When the response to the first mailing of HEAR surveys was lower than anticipated, it became evident that a second mailing was indicated. In the interim since the first mailing, the full implementation of the TRICARE Prime program began in the Hampton Roads area in Virginia on April 1, 1998 and included administration of the HEAR survey to all new TRICARE Prime enrollees. Subsequently, the cover letter required minor revisions to reflect that change. A request to make the change to the cover letter and to conduct a second mailing was sent to the responsible Naval Institutional Review Board on May 28, 1998 and the Commander, Naval Medical Center Portsmouth granted approval in June 1998. The Chair of the Old Dominion University Human Subjects Institutional Review Board gave verbal approval to proceed with the second mailing (personal communication, V. Derlega, May 1998). Follow-up reminder cards were sent to all potential participants selected for the second mailing; no reminder cards were used with the first mailing.

Risks to individual participants were minimal. The study was not experimental research and the participants were not subjected to any experimental procedures. Any

risks or discomforts were related to the personal nature of the questions. The survey included questions regarding the participant's health, medical history, and personal habits. Although the HEAR software could generate letters for the PCM and for beneficiaries reporting the individual survey results (Bell et al., 1996, 1997), it was decided not to include this follow-up in the research as recommended by the Institutional Review Boards.

The HEAR survey was already in full use in several of the military's ten TRICARE regions across the continental United States. The Department of Defense had mandated the HEAR survey be a part of the TRICARE Prime program and be administered to all TRICARE Prime enrollees beginning in January 1997 (Josephs, 1996). It is one of five survey instruments approved by the Department of Defense for administration to beneficiaries across the entire Department of Defense (Martin, 1996).

The cover letters accompanying the mailed survey form in both mailings were signed by the investigator, a doctoral candidate, and endorsed by the Director of the TRICARE Mid-Atlantic Regional Office. The Director was a Navy Captain, a senior military officer equivalent to a Colonel in other branches of the military. The endorsement by the regional Director assured the participants that the research was supported by the TRICARE Mid-Atlantic Region office and that participation would not affect any of their health care benefits. The letter identified the investigator as the point of contact for any questions or assistance as well as the Chair of each of the Institutional Review Boards.

Procedures were employed to protect the identity of the participants. The investigator assigned random numbers to the enrollee database records in order to select

the sample in a random fashion. Each HEAR survey form was pre-printed with another random number that was recorded for all participants when the survey forms were mailed and entered into the enrollee database. The PCM visit data for each participant in the mailing sample from the PCM visit database were merged into the enrollee database after the surveys were mailed. The returned surveys were read by the optical scanner producing a survey results database identified only by the preprinted random number from the survey scan form. The PCM visit data, the raw HEAR survey data, and other non-identifying data were queried using the preprinted random survey number code to match records. The statistical analysis software was used to pull the query results and create a file for the statistical analyses. Identifying information was eliminated during the query and data importing procedures with only the random survey number remaining to uniquely identify each record. That step effectively destroyed the link between survey responses and any identifying information. As a result, the data prepared for analysis by the statistical analysis software were rendered anonymous. The Institutional Review Boards recommended that the original scanned survey forms be maintained for quality control purposes for a period of time before shredding them. After the data were rendered anonymous for the data analysis, the data were never linked to any identifying information again.

Sampling Procedures

The target population of adult active duty family members was identified from the enrollment file maintained on the Naval Medical Center Portsmouth's mainframe database system, the Composite Health Care System (CHCS). The CHCS is the software

installed on mainframe computers in military (medical) treatment facilities worldwide. The CHCS host that serviced the three military treatment facilities in Hampton Roads was located at the Naval Medical Center Portsmouth (Virginia). Government personnel routinely downloaded the TRICARE Prime enrollment file from the CHCS and provided it to the TRICARE Mid-Atlantic regional office monthly. There, other government personnel routinely exported the file into a personal computer database. A copy of the February 1998 enrollment file was provided to the investigator by TRICARE Mid-Atlantic regional office. There were 153,305 active enrollment records in the database file, one for each current TRICARE Prime enrollee in Hampton Roads. The file included all of the active duty service members, active duty family members, retirees, and retiree family members who were enrolled in TRICARE Prime on the Naval Medical Center Portsmouth CHCS platform. That represented 79,976 enrollees served by the primary care contract and about 20,000 enrollees served by military PCMs. The remainder of the enrollment records in the file represented active duty service members (approximately 53,000).

Enrollees in the target population were identified by querying the February 1998 enrollment file. Several selection criteria were used in the query. Individuals with a date of birth before February 24, 1980 were selected to ensure that each participant was at least 18 years old. Only enrollees assigned to one of the eight PCM sites operating under the Navy primary care contract were selected. Since complete PCM visit data were available for only those eight PCMs, the selection was restricted to those sites. By using the February 1998 enrollee file and selecting current enrollees with an enrollment start

date earlier than January 2, 1997, enrollees who were continuously enrolled in TRICARE Prime during calendar year 1997 were selected ($N = 15,136$).

Random numbers were assigned to all records in the enrollment file for the target population and the records were sorted in ascending order by the random sample number. The first 1,000 records were selected from the target population for the first mailing and the next 995 records were selected for the second mailing. Address labels were generated from the enrollment file imported from the CHCS into the personal computer research database file. The random survey number printed on the survey was recorded when the surveys were assembled in the envelope for mailing. Later, that random survey number was entered into the research database. The survey random number was used to match survey responses to the PCM visit data for each participant.

It was known that maintaining correct beneficiary addresses in the CHCS was a challenge for the government. However, addresses for TRICARE Prime enrollees in CHCS were supposedly updated no less than annually when their enrollment was renewed. Subsequently, the addresses could be expected to be current within the past year. Nonetheless, it was anticipated that some of the surveys mailed to participants selected for the sample would not reach a correct address. So, after accounting for undeliverable surveys, it was anticipated that a sufficient number of surveys would be delivered to obtain the desired sample size during the first mailing. However, a second mailing was required as mentioned.

Data Collection

A major purpose of this research was to subject the health resource utilization classification capability of the HEAR survey instrument to the scrutiny of quantitative analyses. Data were obtained from two governmental databases and from returned HEAR surveys. The return rate of the HEAR surveys will be discussed. Finally, the HEAR survey item will be discussed in some detail.

Government Data

Only a few data elements were used from the government data. They were complete for all of the participants since they were continuously enrolled in TRICARE Prime in calendar year 1997. The dependent variable for the research was constructed from the PCM visit government database, which contained PCM visit data at a ratio level of measurement. The database was delivered to the Navy monthly by the primary care contractor serving the Hampton Roads Virginia metropolitan area. The number of annual PCM visits per participant ranged from 0 to 27 visits in 1997. As discussed in Chapter IV in detail, it was determined that 21.2% of the participants had seven or more PCM visits and were responsible for 50.4% of the total number of PCM visits in the study sample. Subsequently, the actual number of PCM visits consumed in 1997 by the participants was used to classify them into two PCM visit groups. Participants in the low utilization group had zero to six visits and participants in the high utilization group had seven or more visits in 1997.

Two data elements were used from the second government database, which contained TRICARE Prime enrollment data for the Hampton Roads, Virginia area. The

enrollment data was extracted monthly by government personnel from the Composite Health Care System (CHCS) housed at the Naval Medical Center Portsmouth and serving all of the Hampton Roads, Virginia area. The date of birth from the government database was easily converted to the date data type by the personal computer database software. This was used because of the difficulty of converting the data for the date of birth from the HEAR survey responses in ASCII text format to a date data type. The two fields were compared through direct observation and found to be identical for all participants in the final sample. Subsequently, the date of birth field was retained from the government data. The date of birth was used to calculate age as of March 1, 1998 and that resulted in a variable at the ratio level of measurement. The gender data in the government data were compared with the gender data from the survey responses. The gender item was retained from the survey responses since it was already coded numerically and the government data were coded as text. The comparison of dates and gender between the government data and the survey data identified inconsistencies in six cases that were subsequently eliminated from the analysis sample.

HEAR Survey Instrument

The HEAR survey instrument was developed by MEDTAP International and Battelle Memorial Institute in collaboration with the Office for Prevention and Health Services Assessment (OPHSA) of the United States Air Force" (Murray and Halpern, 1996, p.1). The survey was designed as a self-administered instrument consisting of 82 items (Bell et al., 1996, 1997). The survey was printed on a form that could be scanned by an optical mark reader device. There were two forms of the surveys. The longer form

provided an opportunity for the participant to indicate identifying information and address. The shorter form was designed for use when the scoring software could link to an existing database containing the identifying information on the participant. Identifying information was required for production use of the survey to generate the beneficiary and PCM reports of the survey results. Survey completion time was estimated at approximately 20 to 30 minutes. The U. S. Air Force Office of Prevention and Health Services Assessment maintained a supply of HEAR scan forms and provided the longer HEAR forms for this research (Appendix D). The written instructions sent with the survey forms explained that the identifying information on the form should not be completed. The random number pre-printed on the form was recorded and used to link data from returned surveys to the PCM visit data.

The first page of the long form (eight pages) is a title page while the second page contains instructions and a Privacy Act Statement. Pages 3 and 4 are the identifying information while pages 5 through 8 contain the 82 survey items numbered alphanumerically. Similar survey items are grouped by letter and then sequentially numbered within each lettered group. The 82 survey items are divided into the 14 groups identified in Table III-1 with 1 to 18 items per group. Instructions for the participant to skip certain items are embedded among some of the survey items. For instance, the cholesterol item reads as follows (Appendix D):

C.1 Blood cholesterol is a fatty substance found in blood.
Have you ever had your blood cholesterol checked?
__ Yes (*go to C2*) __ No (*go to C4*) __ Don't know (*go to C4*)

Table III-1

Description of HEAR Survey Items by Section Grouping

Sect.	Description	No. of Items
A	demographic and general health	8
B	blood pressure	6
C	cholesterol	5
D	physical activity	3
E	women's health	5
F	men's health	1
G	smoking	6
H	alcohol use	8
I	stress	3
J	general satisfaction and family	3
K	mental health symptoms and treatment	6
L	absenteeism and difficulty walking	3
M	inpatient/outpatient services	7
N	medical conditions and family history	18
	Total Items	82

Note. Sect. = lettered section identified on the HEAR survey form

Survey Data Collection and Response Rate

Two mailings were required. After the required human subject approvals were obtained, the Naval Medical Center Portsmouth mailed the surveys on March 4, 1998. The survey forms were mailed in an envelope with the TRICARE Mid-Atlantic logo on it along with a cover letter, a notification of additional information about the research project, and a business reply envelope. Since only 186 surveys were returned and the research required a minimum sample size of 376 participants, it was determined that a second mailing was indicated. The U. S. Air Force Office of Prevention and Health Assessment supplied 995 more survey forms. Selection of additional participants for the second mailing was conducted in the same fashion as the first mailing with the next 995

random numbers queried in sequence from the personal computer database. The second mailing was mailed on June 22, 1998 after the Navy Institutional Review Board and the Old Dominion University Human Subjects Institutional Review Board approved a second mailing. Reminder cards were mailed July 2, 1998 to all 995 addresses included in the second mailing.

For the first mailing conducted on March 4, 1998, 1,000 surveys were mailed and 33 surveys (3.3%) were returned by the Post Office as undeliverable. Of the 967 surveys that were delivered successfully in the first mailing, participants ultimately returned 186 surveys (19.2%). For the second mailing conducted on June 22, 1998, 995 surveys were mailed and 72 surveys (7.2%) were returned as undeliverable. Since the first mailing resulted in a lower than anticipated response rate, reminder cards were mailed July 2, 1998 to all of the addresses included in the second mailing. Of the 923 surveys delivered in the second mailing, participants returned 211 surveys (22.9%). The second mailing had a 3.7% higher return rate than the first mailing.

Overall, 1,995 surveys were mailed with 105 surveys (5.3%) were returned as undeliverable. Of the 1,890 surveys that were delivered, participants returned 397 surveys. The investigator deemed 6 returned surveys to be unusable resulting in a final sample of 391 (20.7%). Reasons for disqualifying returned surveys included mismatches between the gender and the date of birth in the government data and the survey data.

Response Bias Analysis

While the PCM visit data were collected from a government database, the HEAR survey data were dependent upon the response of the potential participants in the

research. The survey response was less than anticipated (20.7%). It should be noted that the potential participants were consumers who were randomly sampled to participate anonymously in a mailed survey. It would be reasonable to anticipate a lower return rate from consumers than one could achieve with a more controlled administration of a survey instrument.

Response rates were reported for several of the TRICARE regions (T. Baker, personal communication, July 1, 1999) where the HEAR survey was administered at the time of enrollment into TRICARE Prime. A response rate of 17% was reported for the TRICARE Heartland region which began administering the HEAR survey April 1, 1998, the same time that it began in the TRICARE Mid-Atlantic region where the research was conducted. The HEAR survey was developed in the TRICARE Southwest region where an overall 37% response rate was reported. Response rates were reported for three other TRICARE regions that have been administering the HEAR survey for several years: TRICARE Central (38%), TRICARE Golden Gate (38%), and TRICARE Southern California (36%).

A response bias analysis was performed on the sample and the population. Excluding members of the sample from the target population, a comparison of four non-identifying variables from the sample ($N = 391$) and the population ($N = 14{,}745$) revealed one significant difference. The study sample had proportionately more members ($n = 83$, 21.2%) in the high utilization group than the target population did ($n = 2{,}525$, 17.1%) according to the chi-square test results ($\chi^2 = 4.497$, $df = 1$, $p = .034$). Since the study sample was not representative of the target population, it was concluded that it would be inappropriate to generalize the research results to the target population.

Another potential source of response bias was a difference in procedure between the first and second mailing. A reminder card was used with the second mailing but not with the first mailing. Further, the public start of the national TRICARE program began for the participants April 1, 1998 which was between the two mailings. Subsequently, participants from the first mailing ($n = 182$) were compared to participants from the second mailing ($n = 209$) using a chi-square analysis. No significant differences were found between the participants from the two mailings in age, number of annual visits, and the classification into utilization groups by the Pareto principle. Participants from the two mailings were also compared by measures from the original HRU model: the 17 HRU indicator variables, the computed HRU sum score, and the final classification into utilization groups. One significant difference ($\chi^2 = 5.86$, $df = 1$, $p = .016$) was found between the participants from the two mailings; more participants ($n = 28$) from the second mailing reported general dissatisfaction in life (as scored by the original HRU model) than participants ($n = 11$) from the first mailing. No differences were found between the participants from the two mailings (chi-square analysis) using measures from the investigator-derived HRU model: variables, subscale scores, HRU sum score, and final classification into utilization groups.

Demographic Description of Study Sample

Demographic information describing the study sample is shown in Table III-2. There were 391 participants in the final sample and they utilized a total of 1,579 primary care manager visits over one year resulting in a mean number of 4.04 annual primary care visits per person. The vast majority of participants were female ($n = 369$) and most of the

participants were married ($n = 366$). Married spouses of active duty service members dominated the study sample. Most of the spouses were dependent wives ($n = 349$, 94.4%), rather than dependent husbands ($n = 14$, 4.6%). Viewed from another perspective, the vast majority of active duty service members who were sponsors for active duty family members were male. While there was a small percentage of males in the sample, a Chi-square analysis detected no significant difference between the sample and the population in regards to gender ($\chi^2 = 2.59$, $df = 1$, $p = .107$). Subsequently, males were not eliminated from the study sample since the same proportion was reflected in the target population. A small number of the participants were never married and probably represented children of active duty sponsors ($n = 17$, 3.6%). Children of active duty service members who are in college continue their military dependent status up until their 23rd birthday and remain eligible for military benefits including health services.

Table III-2

Demographic Description of Study Sample

Variable & level of variable	*n*	%	Cum.%
Gender			
male	22	5.6	5.6
female	369	94.4	100.0
Marital status			
married	366	93.6	93.6
never married	14	3.6	97.2
separated	9	2.3	99.5
divorced	2	0.5	100.0
Race/ethnicity			
White, Non-Hispanic	273	69.8	69.8
Black, Non-Hispanic	53	13.6	83.4
Asian/Oriental	21	5.4	88.8
Other	15	3.8	92.6
White, Hispanic	10	2.6	95.2
Pacific Islander	9	2.3	97.5
Black, Hispanic	8	2.0	99.5
American Indian or Alaska Native	2	0.5	100.0
Age			
18 – 22	28	7.2	7.2
23 – 27	46	11.8	18.9
28 – 32	85	21.7	40.7
33 – 37	104	26.6	67.3
38 – 42	77	19.7	87.0
43 – 47	35	9.0	95.9
48 – 52	12	3.1	99.0
53 – 57	3	0.8	99.7
58 – 62	1	0.3	100.0

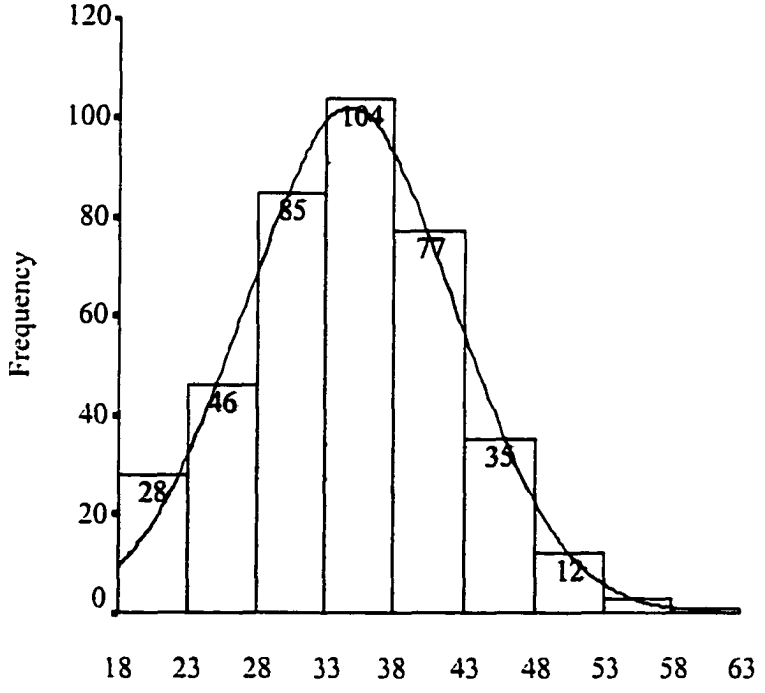
Note. $N = 391$; number of Primary Care Manager visits = 1,579; mean number of visits/ person = 4.04; mean age = 34.7 (*S. D.* = 7.5); minimum age = 18.06; and maximum age = 61.74

The majority of the participants ($n = 273$, 69.8%) indicated their race/ethnicity as white, non-Hispanic. Black, non-Hispanic participants was the second largest group ($n = 53$, 13.6%). The average age of all the participants as of March 1, 1998 was 34.7 years old (*S.D.* = 7.5). The ages ranged from 18.06 to 61.74 years of age. As a continuous

variable, age was normally distributed (skewness = 0.145; kurtosis = 0.056). For description purposes, participants were classified into age groups at intervals of five years. The number and percentage of participants is listed in Table III-2 and a histogram with the normal curve superimposed on it is shown in Figure III-1. The age groups were also normally distributed (skewness = 0.156; kurtosis = -0.106).

Figure III-1

Histogram of Age Intervals With Normal Curve



Scoring Raw Survey Data to Create Variables and Classification Measures

The raw survey responses were scored according to the original HRU model. The HRU outcome was classification into one of three utilization groups, which were subsequently collapsed into two groups. All of the HEAR survey items were closely

examined and variables were constructed for consideration in deriving a HEAR model from the survey items. The approach to handling missing data will be detailed.

Coding of Original HRU Variables and Calculation of HRU Sum Score

Examining the HRU scoring formula revealed more about the original HRU model. The formulas are detailed in the documentation of the software developed under an Air Force contract (Bell, Rosebrough, and Wall, 1996, 1997). Of the 82 items in the HEAR survey instrument, 17 intermediate variables were computed and will be referred to as HRU indicator variables (Table III-3). Each HRU indicator variable was coded with either a value of one if the condition was present or a value of zero if the condition was absent. The formula computed each of the seventeen HRU indicators independently and each item was weighted equally. When multiple survey items were utilized for a particular HRU indicator, a positive response on any one or more of the survey items resulted in a score of one for the indicator variable. The only exception to that scoring approach was mental health; two positive responses were required from the six mental health items to score the HRU mental health indicator as positive (or a value of one). The HRU indicator values were totaled resulting in a range of possible HRU sum scores of 0 to 17.

Table III-3

Original Health Resource Utilization Indicators [a]

Ref.	HRU indicator	Survey item #	Item Description
1[b]	female gender	A.2	gender
2	single marital status	A.3	marital status
3	perceived poor health	A.8	perceived health status
4	hypertension	B.3	told had hypertension 2 times
5	smoking	G.2	smokes how often
6	alcohol	H.5	thoughts to cutdown in past month
		H.6	received complaints about
		H.7	drinking
		H.8	felt guilty about drinking
			had 5 drinks in a day
7	stress	I.1	how often too much stress
		I.2	how much stress in past 2 weeks
		I.3	effect of stress on health
8	family problem	J.1	family satisfaction
		J.2	serious family problems
9	mental health	K.1	little pleasure
		K.2	felt depressed
		K.3	bothered by nerves
		K.4	bothered by worry
		K.5	bothered by anxiety
		K.6	mental health treatment past year
10	work absences	L.1	days spent in bed in past 2 weeks
		L.2	days absent from job past 2 weeks
11	number of medications	M.1	how many different prescriptions
12[c]	doctor visits	M.3	office visits past year
13	ER visits	M.4	ER or urgent care visits past year
14	hospital visits	M.6	hospital nights past year
		M.7	hospital visits past year
15	heart problems	N.3	cc_heart attack
		N.11	angina
16	emphysema	N.4	emphysema
17	arthritis	N-5	arthritis

[a] Information taken from Bell, et al. (1996, 1997) textual material and presented here in tabular form.

[b] Reference numbers were added by the investigator for comparison with Table E-1.

[c] Used as a criterion variable, deleted from computation.

Finally, the HRU sum score was used to classify the participant into one of the three utilization categories: low utilization (sum score from 0 to 4), medium utilization (score equal to 5), and high utilization (score greater than or equal to 6). Since the classification reflected relative standing among the three groups, the HRU measurement was at an ordinal level of measurement. "The fundamental difference between nominal and ordinal measurement is in the latter case information concerning not only equivalence and nonequivalence but also concerning relative standing or ordering among objects is implied" (Polit and Hungler, 1987, p. 341).

The seventeen HRU indicators were selected by the developers because of their utility in predicting utilization and not because they are the most common reasons for adult PCM visits (Murray and Halpern, 1996). The documentation only detailed the developers' literature review and their interviews with a panel of experts, but made no mention of statistical analyses being used in selecting the indicators. Only minor revisions were made to the HRU measure and other computed measures after the field test of the HEAR (Murray and Halpern, 1996).

Although the HEAR software was designed to generate two types of reports in letter format, one for the PCM and another for the beneficiary surveyed, neither type of letter was used in this research (Bell et al., 1996, 1997). A sample of the PCM letter is included in Appendix A (Bell et al.) to show how the HRU result is typically reported on participants to their primary care managers. Both reports address clinical preventive services, counseling services, and chronic conditions/impairments. Additionally, the PCM report also reports HRU level, primary care level, missing or incomplete information, and particular findings to note such as risk factors. The section on clinical preventive services

in the PCM letter also suggests that clinical judgement be utilized along with the survey findings, while the other sections simply report the findings. The beneficiary letter suggested that the beneficiary bring their HEAR report to their next routine visit with the PCM.

The HEAR surveys were scanned by an optical mark reader device. A staff member from the Air Force Office of Prevention and Health Services Assessment who was familiar with the HEAR survey and the government HEAR software programmed the computer file specification required for the optical mark reader to scan the returned survey forms. The output from the optical mark reader was raw survey data in ASCII text format, which were then imported into a personal computer database (Microsoft Access 97). There, the random survey number was used to link it to the PCM visit data. Next, the raw survey responses and the PCM visit data were imported as a flat file to the statistical analysis software. No individually identifying data were included in that final file import procedure. That, in effect, destroyed the link with any identifying information and rendered the analysis data file anonymous.

The HEAR scoring software was not used in the research because the random number preprinted on the survey form was required for matching with the PCM visit data in the government data and the software was not designed to read that number. Routines were written for the statistical analysis software to score the raw survey data in the same manner that the government-developed HEAR software would have scored the data if it had been used. The formulas were obtained from the HEAR software documentation (Bell et al., 1996, 1997).

The computation of the seventeen HRU indicators listed in Table III-3 is detailed next. The survey items used to compute each indicator are identified by the alphanumeric enumeration used on the HEAR survey form (Appendix D). Female gender (HEAR survey item A.2), single marital status (A.3), perceived fair or poor health status (A.8), smoking some days or every day (G.2), and being informed twice of high blood pressure (B.3) were computed as positive indicators. The alcohol indicator was positive for any of the following in the past month: thoughts about cutting down (H.5), complaints about drinking from others (H.6), guilty feelings about drinking (H.7), or at least one day with 5 or more drinks (H.8). Either general dissatisfaction (J.1) or family problems (J.2) was another HRU indicator. The stress indicator was positive for responses of feeling too much stress often (I.1), feeling a lot of stress in the past two weeks (I.2), or feeling a lot of effect on health from stress (I.3). Two or more of the six mental health survey items (K.1 – 6) increased the HRU sum another point. In the past two weeks, five or more bed days (L.1) or five or more missed days from work because of illness or injury (L.2) was a positive HRU indicator. Each contributing one to the HRU sum were the following occurrence in the past 12 months: 21 or more office visits (M.3), five or more emergency department or urgent care center visits (M.4), and either more than one hospital admissions (M.7) or more than six nights in the hospital (M.6). The last four HRU sum indicators were more than five current prescription medications (M.1), heart attack (N.3) or angina (N.11), emphysema (N.4), and arthritis (N.5).

The retrospective design of the research necessitated one modification of the formula used in the original HRU model. The survey was administered in the spring and early summer 1998 and the HRU model was designed to classify participants into

utilization groups for anticipated use in the year immediately following administration of the survey. However, data used in the research for the dependent variable reflected primary care visits actually utilized in 1997, which was prior to administration of the survey. One survey item (M3) inquired about number of outpatient visits during the past twelve months. This variable about past use was designed to contribute to the prediction of future use. However, the time period for the data about past use (1997 to 1998) in this research design overlapped the time period for the data for the dependent variable (1997). That was considered to be unacceptable by a consultant to the Dissertation Committee and the variable was subsequently dropped from the formula used in this research for the original HRU model. That left sixteen indicator variables that were used in the original HRU formula to calculate the HRU sum score for this research. The values of the HRU sum score used to classify participants into utilization groups were unchanged. Those with an HRU sum score of 0 to 4 were classified into the low utilization group, a score of 5 five into the medium utilization group, and a score of 6 to 16 into the high utilization group. When the revised formula was applied to the survey data, the participants were classified into only two groups, low utilization, and high utilization. Only one fewer participant who had been classified into the high utilization group was reclassified by the revised formula into the lower group.

Classification into Utilization Groups Based Upon the HRU Sum Score

The original HRU model classified 343 persons (87.7%) into the low utilization group, 26 persons (6.6%) into the medium utilization group, and 22 persons (5.6%) into the high utilization group (Table III-4). However, only two groups result from a Pareto

analysis. The Pareto principle describes one group as the vital few and the other as the useful many (Juran, 1964, 1975, 1988b, 1992). To approximate the Pareto principle from the results of the original HRU model, the medium and high utilization groups were combined into one high utilization category (n = 48, 12.2%) for the purposes of the remainder of the research. This reclassification is congruent with the validation study of the HRU by investigators from the U.S. Air Force Office of Prevention and Health Services Assessment (1999) who used the same approach of combining the high and medium utilization groups into one high utilization group.

Table III-4

Classification of Participants into Utilization Categories by the Original HRU Model

		Persons			Visits		
Utilization	HRU Sum	No.	%	Cum. %	No.	%	Cum. %
low	0	5	1.3	1.3	12	0.8	0.8
low	1	115	29.4	30.7	397	25.1	25.9
low	2	118	30.2	60.9	489	31.0	56.9
low	3	67	17.1	78.0	233	14.8	71.7
low	4	37	9.5	87.5	150	9.5	81.2
Subtotal		342	87.5		1281	81.2	
medium	5	25	6.4	93.9	124	7.9	89.0
Subtotal		25	6.4		124	7.9	
high	6	11	2.8	96.7	79	5.0	94.0
high	7	6	1.5	98.2	34	2.2	96.2
high	8	5	1.3	99.5	49	3.1	99.3
high	9	0	0.0	99.5	0	0.0	99.3
high	10	2	0.5	100.0	12	0.8	100.0
Subtotal		24	6.1		174	10.1	
Total		391			1,579		

Note. HRU Sum = score range of 0 and 16.

Construction of Variables for the Derivation of an HRU Model and Handling of Missing Data

Since the HEAR survey instrument contained 82 items with multiple possible responses, it was decided to examine all of the survey items to construct variables in a derived HRU model. There were four primary considerations in constructing the variables:

1. taking advantage of as much of the measurement contained in the raw survey data as possible and reasonable,
2. increasing nominal measures used in the original HRU model to ordinal measures where permitted by the raw data,
3. handling missing data, and
4. standardizing the coding direction of the variables.

The direction of the coding for variables at both the nominal and ordinal level of measurement was standardized with zero representing the absence of a condition. For nominal level data, a code of one represented the presence of a condition. For ordinal level data, the values of the codes increased along with the worsening levels of the particular condition. Essentially, lower code values along the ordinal continuum represented better states of the particular condition and higher code values represented worse states of the particular condition. For ordinal data, the coding scale was anchored at zero and increased by consecutive integers up to the maximum value of the particular variable. This standard approach to coding enhanced consistency across all of the data for the purposes of scale construction and interpretation. No effort was made to standardize the number of levels of the ordinal level variables. Rather, effort was applied to use all of

the levels available from the survey data. There were two types of missing data. The software coded blank items as system-missing data by default, while it was programmed to code responses such as *don't know* as user-missing data. The items for gender, and marital status were left untouched except for recoding missing data.

It was discovered that the variables could be grouped into six groups for the derived HRU model according to the content of the variables. The groups were demographic information, general health-related information, mental and emotional health, medical services, disease conditions, and family health. The data dictionary detailing variables, the levels of each variable and coding of each variable can be found in Appendix E (Table E-1) and in the fourth column a reference to the comparable indicator variable from the original HRU model (Table III-3). The survey items used for each variable are identified by the alphanumeric enumeration used on the HEAR survey form (Appendix D).

Demographic variables. The first group of variables constructed for consideration in the derived HRU model contained demographic information. Age was calculated as of March 1, 1998 from the date of birth (A.1) field retained from the government database. The HEAR survey item for gender (A.2) had no missing data and was coded with zero for females and one for males. Zero was chosen for females since they were much more heavily represented in the survey sample. Each of the five choices in the marital status survey item (A.3) was coded. The content of a particular item and its measures of central tendency in the study sample were taken into consideration when coding missing data. Marital status (A.3) was left blank in two cases (0.5% of the sample), and they were coded as married since the married response was the mode. All

eight choices in the race/ethnicity item (A.4) were coded. Race was left blank in six cases (1.5%) and they were recoded to the other racial or ethnic background category since it was the most non-specific category. The recoding of missing demographic data could have introduced a source of error into the derived HRU model, however, ultimately the demographic variables were not used in the model as discussed later.

General health variables. The next group of eight variables was named general health and they described general health-related information. The body mass index was computed from the responses for height (A.6) and weight (A.7). For the 29 cases (7.4%) in which data were missing for either height or weight, the mean body mass index value in the sample was used, 26.0381. The formula from HEAR scoring software for the body mass index and overweight risk factor was used to compute the variable for overweight (Bell et al., 1996, 1997). After setting the value of the overweight variable equal to zero, the value was changed to indicate present if certain thresholds were reached. The overweight thresholds were a body mass index value of 25 for participants up to 36 years old, 27 for participants up to age 56, and 28 for participants up to age 65. The ratio level variable for BMI was not retained for any further analyses since the thresholds were required to interpret the data.

Other general health variables included perceived health status and cholesterol levels. Missing data for health status (item A.8) were replaced with the mode response, very good, in three cases (0.8%). The first survey item on cholesterol (C.1) instructed participants who responded that if they either had never had their cholesterol checked or did not know if it had ever been checked before to skip the item (C.3) that asked about high cholesterol. Item C.1 established that 146 participants (37.3%) had either never had

their cholesterol checked or it was not clear if they had, 130 participants responded no, 15 responded that they did not know, and 1 response was missing. As a result item C.3 was also left blank in 135 cases and 5 cases indicated the don't know response; these were combined for a total of 140 cases (35.8%) with missing data. In order to use item C.3 alone, the missing data in the 140 cases were recoded to zero since they did not indicate they knew what their blood cholesterol was or had not been tested.

Although the survey items for physical activity were not used in the original HRU computations (Bell et al., 1996, 1997), they were used in the general health category for the derived HRU. The three items related to physical activity were frequency of physical activity (D.1), physical work on the job (D.3), and physical work in main activity. Specifically, frequency (D.1) asked how many times in their average week participants had engaged in physical activity (exercise or work) for at least 20 minutes. Job (D.2) asked how much hard physical work was required on their job and main activity (D.3) asked how much hard physical work was required in their main activity (household or non-job activities). If participants read the content of the items carefully and responded accordingly, job and main activity were independent of each other. However, either one or both of them could overlap with frequency. Consequently, a new composite variable was constructed for physical activity with a code of two indicating low physical activity level, the worst state of the variable. Moderate activity level was coded with one if frequency indicated one to two times per week or job indicated a moderate amount or main activity indicated a moderate amount. High activity level was coded with two if frequency indicated at least three times per week, job indicated a great deal or main activity indicated a great deal. While survey item job was left blank in 4 cases (1.0%), the

response for not currently working was indicated in 96 cases (24.6%). Frequency had one missing case (0.3%) and main activity had 51 missing cases (13.0%). Missing data for each physical activity item were coded to reflect the lowest level of activity for the particular variable. There was missing data in 13 cases (3.3%) and each case was recoded to reflect the lowest level of physical activity.

Another variable for the health group was family separation for 30 consecutive days in the past year. It was coded as a dichotomous variable with one for yes and zero for no.

The smoking questions apparently confused some participants by the layout of the items on the HEAR survey form. Most questions on the form were arranged in a double column layout. However, item F.1 was arranged across the whole page in a single column in the middle of the page and was preceded by a heading spanning the whole page and indicating *F. Men's Health.* Underneath item F.1, the layout returned to double columns with the smoking questions and the first four alcohol questions at the bottom of the page. There were 18 participants who skipped the rest of the questions at the bottom of the page after item F.1, 4 who skipped only the left column containing the smoking questions, one who skipped the entire page, and 3 who skipped that page and the facing page on the right. That accounted for all 25 cases (6.4%) in which item G.1 was left blank. Participants who responded that they had never smoked at least 100 cigarettes in their life (G.1) were instructed to skip the remainder of the smoking questions; 158 answered yes and should have answered the remainder of the smoking questions. The most specific information about smoking of any of the smoking questions was obtained from item G.3. This item was selected as a variable at an ordinal level of measurement for the present

research. It inquired about the average number of cigarettes smoked per day and was answered by 83 participants, one of whom endorsed the don't know response. The blank answers and the don't know response were recoded with zero to indicate currently smoking less than one cigarette per day.

The two variables about decrease in activity because of illness or injury in the past two weeks (L.1 and L.2) and the one dichotomous variable about difficulty walking (L.3) were also included in the general health category. Absenteeism (L.2) and days spent in bed (L.1) in the past two weeks were coded as a dichotomous composite item in the original HRU model (Bell et al., 1996, 1997). For the derived HRU model, one variable was constructed from each of the survey items to be considered. There were five possible responses that ranged from none to seven or more days plus one response for don't know. All of the first five responses were coded separately. No participants indicated the don't know response. Item L.1 was left blank in four cases (1.0%), item L.2 was left blank in nine cases (2.3%), and item L.3 was left blank in four cases (1.0%). Missing cases were recoded with a value of zero for absent condition. Zero indicated no difficulty in walking in item L.3.

Emotional health variables. The third group of variables constructed for consideration in deriving a HRU model dealt with a number of emotional and mental health areas. Eighteen emotional health variables were constructed for consideration. The variables covered alcohol use, stress, satisfaction, family problems, family separation, and mental symptoms; they will be discussed in that order.

A separate variable was constructed for all but one of the eight items in the survey relating to alcohol (items H.2 – 8) and all possible responses were coded separately. Item

H.1 was not used because it only asked participants if they had a drink in the past month and instructed them to skip the remaining alcohol items if they answered no; 211 of the 391 participants (54.0%) answered yes and they should have answered the remaining alcohol questions while 180 of the participants (46.0%) should have left the remainder of the alcohol items blank. Using all of the items contrasted with the original HRU model which produced only one alcohol indicator and it was a dichotomous composite variable of the items H.5 – 8 (Bell et al., 1996, 1997). Those four items (H.5 – 8) were

> a variant of the ones used in the CAGE questionnaire (cut down, annoyed by criticism, guilty about drinking, eye-opener drinks) for the detection of alcoholism.... The rationale to using the CAGE questions was that they may be less likely to trigger defensiveness and denial in the alcoholic (Murray and Halpern, 1996, p. 9).

Cases with blank responses to an item and responses such as don't know were recoded to zero. Number of days with a drink in the past two weeks (H.2) was left blank in 158 cases (40.4%); average drinks in a day in the past two weeks when drinking (H.3) was left blank in 159 cases (40.7%) with 5 don't know responses (1.3%); and number of times driving after too much to drink in the past month (H.4) was left blank in 158 cases (40.4%) with 3 participants (0.8%) indicating that they don't drive. The four items that were variants of the CAGE items produced four dichotomous variables. There were 153 blank responses (39.1%) for each of three of those variables: thinking of cutting down on drinking in the past month (H.5), hearing of complaints about their drinking in the past month (H.6), and feeling guilty or upset about drinking in the past month (H.7). There

were 154 blank responses (39.4%) to the question about drinking five or more drinks in at least one day in the past month.

The three survey items on stress were also collapsed into a single dichotomous item on stress in the original HRU model (Bell et al., 1996, 1997). For the derived HRU model, three separate variables were constructed from the three survey items and all possible responses were coded in a manner that resulted in ordinal level data. Responses were missing in ten cases (2.6%) for the current frequency of too much stress (I.1), ten cases (2.6%) for amount of stress in the past two weeks (I.2), and nine cases (2.3%) for amount of effect from stress (I.3). Variables with missing data were recoded to indicate too much stress never (I.1), almost no stress at all (I.2), or hardly any or no effect from stress (I.3).

General satisfaction (J.1) and family problems (J.2) were computed together as a composite score in the original HRU model and were referred to as family satisfaction. However, the content in the general satisfaction question did not mention family. Rather, it mentioned work situation, social activity, and accomplishments. This investigator decided that the two items were really not variations of the same idea, but were actually two different concepts. Subsequently, they were constructed as two separate variables for the derived HRU model. General satisfaction was coded with higher numbers reflecting more satisfaction. However, one of the guidelines employed in this research for the construction of variables was that the lower numbers should indicate the absence of a condition. Since the absence of satisfaction really is dissatisfaction, the coding was reversed for the constructed variables. The coding also had to be reversed from the scanned data for how often family problems were experienced and for family separation

(J.3). Missing data were coded as zero for 10 cases (2.6%) for J.1 items, 10 cases for J.2 items, and 13 cases (3.3%) for J.3 items. The survey item for thirty consecutive days of family separation was a dichotomous item so it was coded accordingly.

The mental symptoms items (K.1 – 5) were coded with zero to indicate absence of the symptom and with one to indicate presence of the symptom. The sixth item, mental health treatment (K.6) was classified under the medical services group. Missing data were found in five cases (1.3%) for little pleasure in doing things (K.1), in seven cases (1.8%) for feeling depressed (K.2), in seven cases (1.8%) for nerves (K.3), and seven cases (1.8%) for worrying (K.4), and five cases (1.3%) for anxiety attack (K.5). The same approach was followed and the cases with missing data on these variables were recoded to zero since they did not definitively indicate the presence of the symptom.

Medical services variables. Seven constructed variables were considered for the medical services group. All but one of the seven were found as separate indicator variables in the original HRU model where they were coded as dichotomous variables. Number of outpatient visits during the past twelve months (M.3) was eliminated for the same reasons of overlapping time period with the dependent variable as discussed under the Original HEAR Survey model. The variable for two or more visits for a musculoskeletal problem (N.16) was not included in the original HRU model. All the responses in the survey items for these seven variables were coded for the derived HRU model. Missing data were found in seven cases (1.8%) for mental health treatment in the past year (K.6), 3 cases (0.8%) for number of current prescription medications (M.1), 12 cases (3.1%) for office visits in the past month (M.2), and 2 cases (0.5%) for ER visits in the past year (M.4). Since some participants responded that they had spent some nights in

the hospital in the past year (M.5), but failed to indicate the number of nights (M.6) and vice versa, those items were coded as a composite score at the ordinal level of measurement. The resulting variable was named hospital stay and was first set equal to zero. Then either an affirmative response to M.5 or a response of one to two nights to M6 was scored as one to two nights. The other three responses to item M.5 were coded as separate responses. There were two missing responses (0.5%) to survey item M.5. The final medical service variable, visits for a musculoskeletal problem (N16), was a dichotomous item on the survey and was coded accordingly. There was missing data in one case (0.3%) for bone or muscle problem.

Disease condition variables. The fifth group contained 14 constructed variables and was called disease conditions. The original HEAR computation scored the hypertension question (B.3) as a dichotomous variable for the HRU sum measure (Bell et al., 1996, 1997). A new variable was computed for hypertension at the ordinal level of measurement. First, the new computation set the value equal to zero to indicate that the condition was absent. That was consistent with the original formulation and the direction of the recoded variables for this research. A yes response to either item B.2, told by a health professional that you had hypertension or high blood pressure, or to item B.3, told two or more times that you had hypertension or high blood pressure, was coded as one to indicate hypertension diagnosed. A yes response either to having ever been prescribed medication for hypertension (B.4) or to currently taking hypertension medication (B.5) was coded as two to indicate hypertension medicated. Item B.6, frequency of taking high blood pressure medication was determined to be an ambiguous question: 48 participants indicated one of the frequency choices, but only 23 participants indicated that they were

currently taking blood pressure medication (B.5). Subsequently, item (B.6) was not used in the new computations. Participants were instructed to skip the remainder of the blood pressure items if they did not answer yes to any one of the items B.2, 4, or 5. Responses were left blank in 3 cases (0.8%) to ever told by a health professional that you had hypertension (B.2), in 296 cases (75.7%) to ever told two or more times that you had hypertension, and in 324 cases (82.9%) to currently taking prescription medications for hypertension (B.5). These blank responses were all recoded to zero since they were not affirmative responses.

All of the items about disease conditions in Section N of the survey instrument were dichotomous items. Variables with a no response were coded with a zero and those with a yes response were coded with a one. Data were missing in six cases (1.5%) for the diabetes question (N.1), in two cases (0.5%) for heart attack (N.3), in two cases (0.5%) for emphysema/bronchitis (N.4), in four cases (1.0%) for arthritis (N.5), and in one case (0.3%) for Parkinson's or other neurological disease (N.6). Data were also missing in two cases (0.5%) for cancer (N.10), in one case (0.3%) for kidney disease (N.13), in five cases (1.3%) for stomach ulcer (N.14), and in two cases (0.5%) for asthma (N.15). There were no missing data for stroke (N.2), depression (N.7), HIV/AIDS (N.8), anxiety or personality disorder (N.9), heart disease or angina (N.11), and liver disease (N.12).

The disease condition variable for anxiety or personality disorder (N.9) was deleted since anxiety disorders and personality disorders are described as conceptually separate categories of disorders in the DSM-IV, which is used to standardize diagnoses of mental disorders (American Psychiatric Association, 1994). The investigator agreed with the original HRU model where heart attack (N.3) and heart disease or angina (N.11) were

combined and computed as one variable for heart disease (Bell et al., 1996, 1997). Since the HIV/AIDS (N.8) variable showed no variation among survey participants, it was not considered further for the derived HRU model.

Family health variables. Two variables which were in the final group of variables were described as Family Health. One variable involved sick family members and the other involved family medical history of heart disease. Data were missing in one case (0.3%) for dependent with a serious medical condition (N.17) and in 12 cases (3.1%) for family history of heart disease (N.18). Responses in these cases were recoded to zero.

Analysis of Data and Statistical Tests

The overall purpose of the analyses was to compare the classification ability of the original HRU model with an alternate HRU model systematically derived. All statistical analyses were performed using the personal computer software package, SPSS® for Windows® (1998). The data were analyzed in several phases. First, Primary Care Manager visit utilization data from the government database were examined using a Pareto analysis. The government data included the number of PCM visits in 1997 by active duty family members enrolled in TRICARE Prime and assigned to a PCM serving under the Navy primary care contract in the Hampton Roads metropolitan area of Virginia. The Pareto analysis identified the cut point for the artificial dichotomy between the low PCM visit utilization group and the high utilization group along the ratio level variable for number of annual visits per person. Participants were assigned to one of the two utilization groups. That created the dependent variable that served as the criterion against which to evaluate the classification ability of the HRU models.

The reliability of the sum score from the original HRU model was examined using Cronbach's alpha to test its internal consistency. The Cronbach's alpha test produces a reliability coefficient that measures the internal consistency of a test (Norušis, 1997). It is used for items within a scale that is designed to measure a common entity. The common measure in this research was health resource utilization. The Kuder-Richardson formula (KR-20) is appropriate when the item scores are dichotomous (Cronbach, 1984). However, "if the data are in dichotomous form, α is equivalent to reliability coefficient KR-20 (Kuder-Richardson-20)" (SPSS, 1988, p. 873). The KR-20 formula also measures internal consistency and "results in an estimate of reliability that is essentially equivalent to the average of the split-half reliabilities computed for all possible halves" (Gay, 1987, p. 140). The reliability estimated by the KR-20 formula is called rationale equivalence reliability.

Using the SPSS® statistical software, the reliability analysis indicated item by item how much the alpha would be increased if an item were removed. If the indicated value was greater than alpha score, the analysis revealed that the reliability would improve. Eliminating the item that indicated the greatest amount of improvement, a step-wise approach was employed to examine the reliability again. This cycle continued until the maximal benefit gained from eliminating items was reached. The results were examined to determine which variables, if any could be removed to improve the reliability of the measure. The indicated variables were removed and the HRU measure and classification were recomputed to determine if the performance of the HRU model improved.

Next, the validity of the classification results from the original HRU model was estimated. A classification table was constructed to examine the ability of the model to correctly classify participants into one of two utilization categories observed in the sample and identified by the Pareto analysis. The classification results were evaluated by comparing the observed utilization groups to the predicted utilization groups with a Kendall's tau-b procedure. "A correlation of a test score with a criterion measure ... is called a validity coefficient" (Cronbach, 1984, p. 136). Therefore correlation between the predicted and observed utilization groups was calculated using Kendall's tau-b as an estimate of validity. The sum score results were examined to determine if the cutpoint between the dichotomous groups could be adjusted and improve the performance of the HRU model.

An HRU model was derived systematically from the HEAR survey items to consider as an alternative to the original HRU model. Subscales were constructed from five of the six groups of the constructed variables: general health-related information, emotional and mental health, medical services, disease conditions, and family health. No reasonably logical approach was readily apparent to construct a subscale for the group of demographic information variables. The reliability of each of the five subscales was estimated with the same step-wise approach using the Cronbach's alpha procedure as described above. Items were deleted as necessary until the maximal improvement to the reliability was achieved. Correlations among the remaining variables within each subscale of the constructed variables were examined to screen for multicollinearity to consider discarding variables with a correlation coefficient greater than .850.

An overall derived HRU sum score was produced when the subscale scores were added together. The cut-point was determined from an examination of the 20% of the highest HRU sum scores. Finally, the performance of the original HEAR model was compared to the performance of the derived HRU model. The computed reliability and validity coefficients were used as the basis of comparison.

CHAPTER IV

RESULTS

This research compared classification results using the original Health Resource Utilization model as designed by the developers of the HEAR survey with classification results using a health resource utilization model derived from variables constructed from HEAR survey items as part of the research. All 82 of the HEAR survey items were examined closely and variables were constructed to make maximal use of the available survey data. Missing data were recoded during the construction of the variables as indicated. The variables were classified into six groups and five groups were used to construct subscales. Five subscales and four demographic variables were available for the derivation of an alternate HRU model. The construction of the variables for the subscales was discussed in detail in Chapter III. The construction of the subscales used for the overall derived HRU scale and classification model is discussed in this chapter. The classification results from the derived HRU model were compared with the classification results from the original HRU model to analyze the performance of the models.

The discussion of the results begins with a discussion of the distribution of Primary Care Manager (PCM) visits observed in the sample of participants. Following is a discussion of the Pareto analysis performed on the actual data of outpatient PCM visits in 1997 to define the dependent variable. The results of the validity and reliability analyses of the original HRU model follows. Additionally, analyses were performed in an effort to improve the performance of the original HRU model. Then, the reliability analyses and construction of the subscales for the derived HRU model will be discussed.

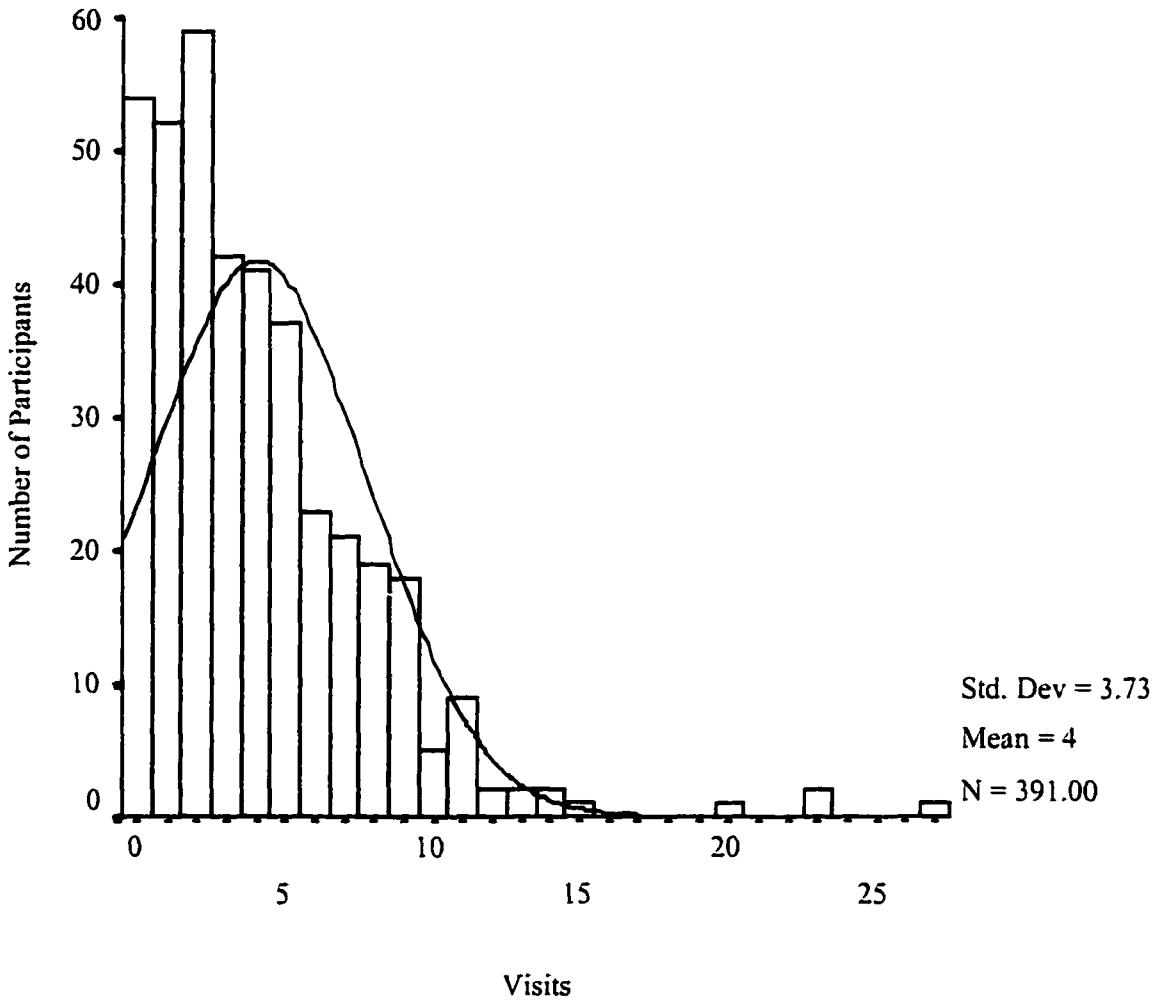
The discussion continues with the procedure for computing the derived HRU sum score and constructing the classification model. The discussion of the results will conclude with a comparison of the reliability and validity of each between the original HRU model and the derived HRU model.

Distribution Of PCM Visit Utilization

Parametric procedures were not chosen for the analysis because the distribution of PCM visits consumed by participants in calendar year 1997 was positively skewed and not normally distributed (skewness = 1.901; kurtosis = 6.701). The distribution of visits is shown in Figure IV-1 and in Table IV-1. The median number of visits was three annual PCM visits while the mode was two visits ($n = 59$; 15.1%); zero annual PCM visits ($n = 54$; 13.8%) was observed as the second most frequent number of visits. Since the distribution of the visits was not normally distributed, other measures of central tendency need to be carefully interpreted; the average number of visits observed was 4.04 annual PCM visits ($SD = 3.73$). Neither the reason nor the duration of each of the PCM visits was available from the government data. It was reported that primary care managers routinely performed annual women's health examinations as part of the Navy contract (D. Nagy, personal communication, August 11, 1999).

Figure IV-1

Histogram of Annual PCM Visits Among Participants with Normal Curve



Pareto Analysis of PCM Visit Utilization

As discussed in detail in Chapter 1, the Pareto principle holds that "in any population that contributes to a common effect, a relative few of the contributors account for the bulk of the effect" (Juran, 1992, p. 57). A Pareto analysis can be performed to identify those relative few whom Juran (1988b, 1992) termed the vital few. The

remainder was termed the trivial many, which he later re-characterized as the useful many. The vital few are the approximately 20% of the contributors who account for about 80% of the common effect. The common effect under study was the consumption of PCM visits in the calendar year 1997. The contributors were the study participants. The vital few was the small number of participants who were responsible for a disproportionate number of outpatient PCM visits consumed during the study year. The overall purpose of this study was to determine how best to use the HEAR survey to identify the participants who belong to the vital few group, the group responsible for high PCM visit utilization. The prediction model is called the Health Resource Utilization (HRU) model.

The Pareto analysis can be approached from two directions. It can begin with the 80% of the visits and solve for the percentage of the persons responsible for those visits. Alternatively, the Pareto analysis can begin with 20% of the persons who consumed the most visits and solve for the overall percentage of visits consumed. The discussion will begin with the first approach, a Pareto analysis to determine the percentage of participants responsible for approximately 80% of the utilization of PCM visits in calendar year 1997. The tabulated results of the Pareto analysis are shown in Table IV-1. The cumulative percentages of persons (x-axis) listed in Table IV-1 are plotted against the cumulative percentages of visits (y-axis) on a line graph in Figure IV-2. The low and high utilization groups comprise the vital few and the useful many respectively. Figure IV-3 shows the disproportionately high use of PCM visits by the vital few, the participants utilizing seven or more visits in a year.

Table IV-1

Pareto Analysis of PCM Visit Utilization per Person to Establish the Dependent Variable

Visits/	Persons			Visits		
person	No.	%	Cum. %	No.	%	Cum. %
27	1	0.3	0.3	27	1.7	1.7
23	2	0.5	0.8	46	2.9	4.6
20	1	0.3	1.0	20	1.3	5.9
15	1	0.3	1.3	15	0.9	6.8
14	2	0.5	1.8	28	1.8	8.6
13	2	0.5	2.3	26	1.6	10.3
12	2	0.5	2.8	24	1.5	11.8
11	9	2.3	5.1	99	6.3	18.0
10	5	1.3	6.4	50	3.2	21.2
9	18	4.6	11.0	162	10.3	31.5
8	19	4.9	15.9	152	9.6	41.1
7	21	5.4	**21.2** [a]	147	9.3	**50.4** [a]
6	23	5.9	27.1	138	8.7	59.1
5	37	9.5	36.6	185	11.7	70.8
4	41	10.5	**47.1** [b]	164	10.4	**81.2** [b]
3	42	10.7	57.8	126	8.0	89.2
2	59	15.1	72.9	118	7.5	96.7
1	52	13.3	86.2	52	3.3	100
0	54	13.8	**100.0**	0	0.0	**100.0**
Total	391			1,579		

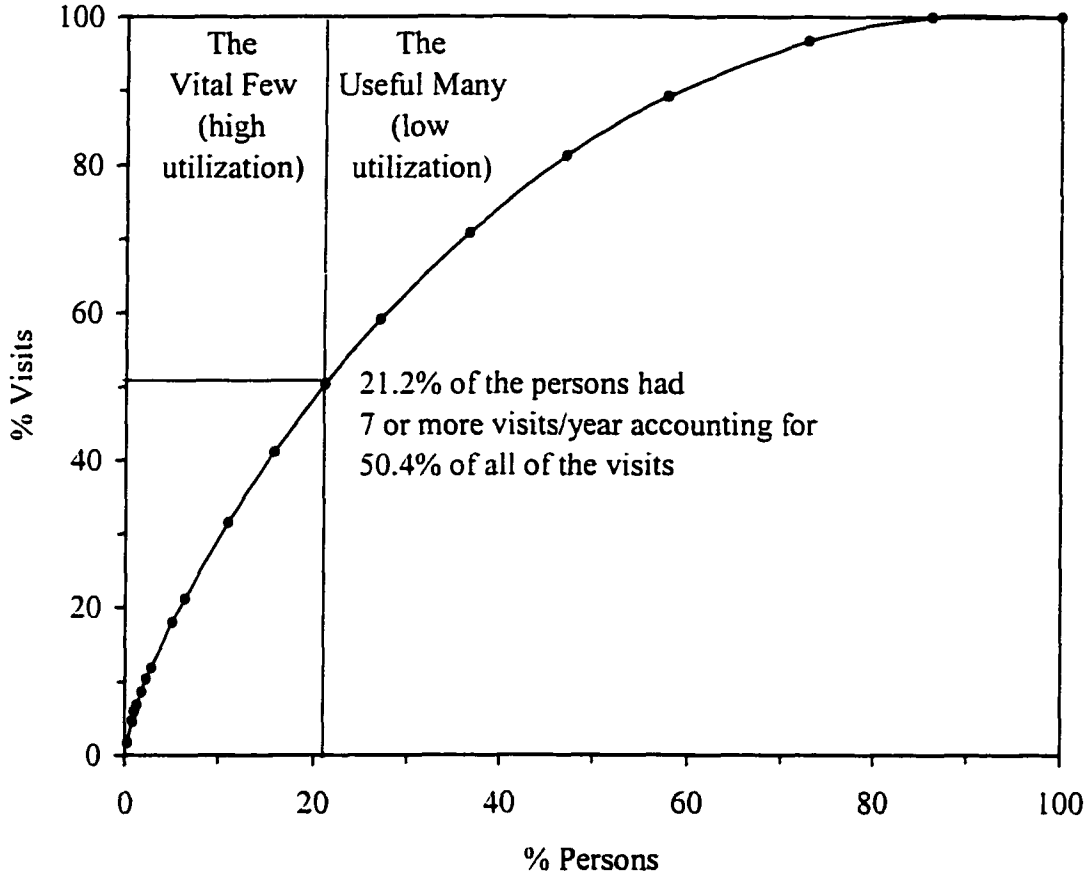
Note. The table is sorted on the first column, visits/person, in descending order. The sample is divided into two utilization groups at the line between three and four visits/person. Visits/person = annual visits to PCM during 1997; Total visits = the quotient of visits/year multiplied by no. of persons.

[a] 21.2% of the persons (n = 83) consumed 50.4% of the health resources as measured by number of visits (n = 796)

[b] 81.2% of the health resources as measured by number of visits (n = 1,283) were consumed by 47.1% of the persons (n = 184)

Figure IV-2

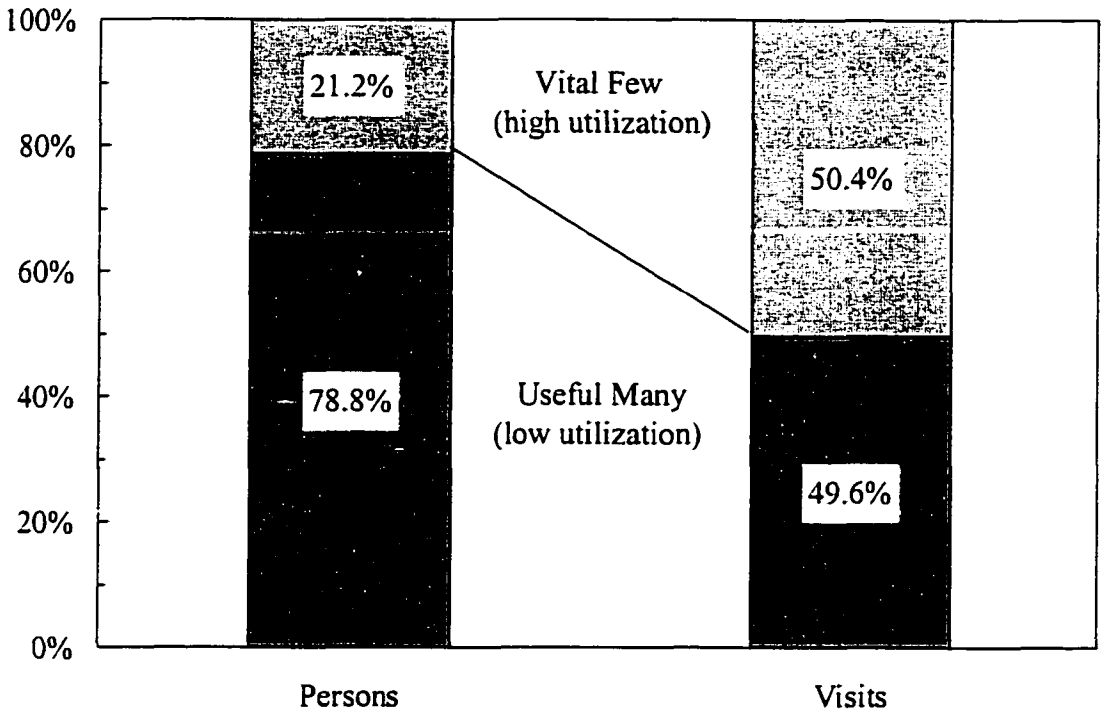
Line Graph of PCM Visit Utilization Pareto Analysis



Note. One data point is plotted from each row of Table IV-1. The x value came from the fourth column of Table IV-1 and the y value came from the seventh column of the table. There is one data point for each rate of utilization expressed as number of primary care visits during the study year.

Figure IV-3

Bar Graph of PCM Visit Utilization Pareto Analysis



The distribution of visits by number of annual visits per person did not strictly follow the Pareto distribution, e.g. the 80/20 rule. The Pareto analysis revealed that 47.0% of the participants in the study sample consumed 81.2% of the total number visits (Table IV-1). Each participant in that high utilization group consumed three or more visits in 1997. It was mentioned earlier that planning for the vital few should proceed on an individual basis (Juran, 1992). Following that line of reasoning it could be quite a challenge for a primary care manager (primary health care provider) to give individual attention to 47.0% of the enrollees assigned to him or her. A group of 47% of the individuals in a target population probably would not fulfill Juran's concept of the vital

few. Consequently, the investigator determined that approaching the Pareto analysis from this angle did not yield a useful result.

The second approach to the Pareto analysis ascertained the percentage of the effect that could be attributed to 20% of the contributors to the effect. That Pareto analysis revealed that 21.2% of the participants accounted for 50.4% of the PCM visits (Table IV-1). That approach yielded a more useful result. While it was noted that the resulting distribution did not strictly follow the 80/20 Pareto distribution, it was also observed that it would be more reasonable to give individual attention to the approximately 20% of the TRICARE Prime enrollees who utilized approximately half of the visits than the approximately 50% who utilized approximately 80% of the visits. The essence of the Pareto principle lies in a disproportionate distribution of effects among a small number of the population of interest. Selecting the 20% utilizing 50% of the visits follows the essence of the principle.

Therefore, the dependent variable in all further analyses was derived from the result of the Pareto analysis whereby 21.2% of the participants ($n = 83$) utilized 50.4% of the primary care manager visits. The dependent variable was the dichotomous variable called utilization group. Members of the high utilization group, the vital few in Juran's terminology, were the 21.2% participants who consumed seven or more visits during 1997. Members of the low utilization group ($n = 308$), the useful many in Juran's terminology, were the remaining 78.8% participants who consumed six or fewer visits during 1997 representing 49.6% of the total number of visits. Participants in the high utilization group consumed a disproportionate amount of visits relative to participants in the low utilization group. Utilization group was a dichotomous variable at the ordinal

level of measurement. "The fundamental difference between nominal and ordinal measurement is in the latter case information concerning not only equivalence and nonequivalence but also concerning relative standing or ordering among objects is implied" (Polit and Hungler, 1987, p. 341).

Original HRU Model

The original HRU model revised for this research computed sixteen dichotomous indicator variables from the HEAR survey items. Scores were added together to arrive at the HRU sum score. That score was used to classify survey participants into one of three utilization categories. Since two categories were desired to be consistent with the Pareto principle, the medium and high utilization groups were combined to create a single high utilization group. The classification resulting from the HRU sum score was examined and compared to the two actual or observed utilization groups identified from the Pareto analysis. The sixteen computed variables used to produce the HRU sum score were entered into the Cronbach's alpha test to estimate reliability.

Bivariate Analyses with the PCM Visit Dependent Variable

The indicator variables in the original HRU model were crosstabulated by utilization group from the Pareto analysis (Tables IV-2 and IV-3). Chi-square tests were performed and the significance values are shown in the first column. Female gender, single marital status, hypertension, stress, and arthritis were significantly associated with the two utilization groups.

Table IV-2

Crosstabulation of Variables in the Original HRU Model by Utilization Groups with Row Percentages

HRU indicator		No. by utilization group			% by utilization group		
Significance		Low	High	Total	Low	High	Total
female gender	no	22	0	22	7.1	0.0	5.6
.012 *	yes	286	83	369	92.9	100.0	94.4
single marital status	no	294	83	377	95.5	100.0	96.4
.048 *	yes	14	0	14	4.5	0.0	3.6
fair/poor health	no	291	74	365	94.5	89.2	93.4
.084	yes	17	9	26	5.5	10.8	6.6
hypertension	no	289	69	358	93.8	83.1	91.6
.002 **	yes	19	14	33	6.2	16.9	8.4
smoking	no	247	65	312	80.2	78.3	79.8
.705	yes	61	18	79	19.8	21.7	20.2
alcohol problem	no	273	75	348	88.6	90.4	89.0
.656	yes	35	8	43	11.4	9.6	11.0
stress	no	239	46	285	77.6	55.4	72.9
.000 ***	yes	69	37	106	22.4	44.6	27.1
family problem	no	282	70	352	91.6	84.3	90.0
.051	yes	26	13	39	8.4	15.7	10.0
mental health	no	170	43	213	55.2	51.8	54.5
.582	yes	138	40	178	44.8	48.2	45.5
work absences	no	305	81	386	99.0	97.6	98.7
.302	yes	3	2	5	1.0	2.4	1.3
no. of medications	no	304	80	384	98.7	96.4	98.2
.158	yes	4	3	7	1.3	3.6	1.8
ER visits	no	306	81	387	99.4	97.6	99.0
.157	yes	2	2	4	0.6	2.4	1.0
hospital visits	no	299	81	380	97.1	97.6	97.2
.802	yes	9	2	11	2.9	2.4	2.8
heart problems	no	305	82	387	99.0	98.8	99.0
.853	yes	3	1	4	1.0	1.2	1.0
emphysema	no	294	79	373	95.5	95.2	95.4
.916	yes	14	4	18	4.5	4.8	4.6
arthritis	no	282	66	348	91.6	79.5	89.0
.002 **	yes	26	17	43	8.4	20.5	11.0

Note. $N = 391$; Low = low utilization group ($n = 308$, 78.8%); High = high utilization group ($n = 83$, 21.2%). Percentage is reflected by column rather than by row. Chi-square performed to produce the significance value results.
*$p < .05$. **$p < .01$. ***$p < .001$.

Table IV-3

Crosstabulation of Variables in the Original HRU Model by Utilization Groups with Column Percentages

HRU indicator		No. by utilization group			% by utilization group		
Significance		Low	High	Total	Low	High	Total
female gender	no	22	0	22	100.0	0.0	100.0
.012 *	yes	286	83	369	77.5	22.5	100.0
single marital status	no	294	83	377	78.0	22.0	100.0
.048 *	yes	14	0	14	100.0	0.0	100.0
fair/poor health	no	291	74	365	79.7	20.3	100.0
.084	yes	17	9	26	65.4	34.6	100.0
hypertension	no	289	69	358	80.7	19.3	100.0
.002 **	yes	19	14	33	57.6	42.4	100.0
smoking	no	247	65	312	79.2	20.8	100.0
.705	yes	61	18	79	77.2	22.8	100.0
alcohol problem	no	273	75	348	78.4	21.6	100.0
.656	yes	35	8	43	81.4	18.6	100.0
stress	no	239	46	285	83.9	16.1	100.0
.000 ***	yes	69	37	106	65.1	34.9	100.0
family problem	no	282	70	352	80.1	19.9	100.0
.051	yes	26	13	39	66.7	33.3	100.0
mental health	no	170	43	213	79.8	20.2	100.0
.582	yes	138	40	178	77.5	22.5	100.0
work absences	no	305	81	386	79.0	21.0	100.0
.302	yes	3	2	5	60.0	40.0	100.0
no. of medications	no	304	80	384	79.2	20.8	100.0
.158	yes	4	3	7	57.1	42.9	100.0
ER visits	no	306	81	387	79.1	20.9	100.0
.157	yes	2	2	4	50.0	50.0	100.0
hospital visits	no	299	81	380	78.7	21.3	100.0
.802	yes	9	2	11	81.8	18.2	100.0
heart problems	no	305	82	387	78.8	21.2	100.0
.853	yes	3	1	4	75.0	25.0	100.0
emphysema	no	294	79	373	78.8	21.2	100.0
.916	yes	14	4	18	77.8	22.2	100.0
arthritis	no	282	66	348	81.0	19.0	100.0
.002 **	yes	26	17	43	60.5	39.5	100.0

Note. $N = 391$; Low = low utilization group ($n = 308$, 78.8%); High = high utilization group ($n = 83$, 21.2%). Percentage is reflected by column rather than by row. Chi-square performed to produce the significance value results.
*$p < .05$. **$p < .01$. ***$p < .001$.

Reliability Analysis with the Original HRU Model

The Cronbach's alpha for the original HRU model was estimated to be .5611 while the standardized item alpha was .5892. "For most purposes, reliability coefficients above .70 are considered satisfactory. In some situations, a higher coefficient may be required, or a lower one may be considered acceptable" (Polit and Hungler, 1987, p. 318). Since the reliability coefficient was .5892 using the standardized item alpha, the original HRU model was not determined to be reliable.

Validity of Classification with the Original HRU Model

The classification results shown in Table IV-4 can be evaluated in terms of sensitivity and specificity. Sensitivity is the true positive rate and specificity is the true negative rate (Waltz, Strickland, and Lenz, 1991). Since the classification models in this research are attempting to detect the identity of the members of the high utilization group, that would be considered the positive. The sensitivity of the original HRU model was observed to be 25.3% (21 true positives) while the specificity was observed to be 91.2% (281 true negatives). Further, the classification resulted in 27 false positives (8.8%) and 62 false negatives (74.7%). Overall, the original HRU model correctly classified 77.0% of the participants into the correct categories.

Table IV-4

Classification Table of PCM Visit Utilization Groups Using Original HRU Model

Actual visit utilization [a]	Original HRU classification [b]		
	low	medium/high	Accuracy (% correct)
low	281 [c]	27 [d]	91.2% specificity
high	62 [d]	21 [c]	25.3% sensitivity
			77.2% overall correct

Note. $N = 391$.
[a] Actual visit utilization is the dependent variable that was determined by the Pareto analysis for the actual number of primary care manager visits utilized by the participants in 1997.
[b] Original HRU classification was determined from the scoring of the original HRU model.
[c] Correctly classified.
[d] Incorrectly classified.

However, the overall correct classification percentage can be misleading. Note that even if the model failed to classify any of the participants into the medium/high HRU category, the overall correct classification would be 78.8% since there were 308 participants actually in the low utilization group and 83 participants in the high utilization group. Stated another way, if the specificity of a model with the study sample was 100% and the sensitivity was 0.0%, the overall classification would be 78.8%. The actual numbers can also be examined from a different perspective. The model classified a total of 48 participants into the high utilization group, but only 21 (43.8%) belonged in the group. The other 56.2% of the participants belonged in the low utilization group. One more perspective can be seen from another examination of Table III-4. The 12.2% of the participants classified into the high utilization group by the original HRU model were responsible for (18.5%) of the PCM visits. That can be compared to the 21.2% of the

participants identified by the Pareto analysis (Table IV-1) who utilized 50.4% of the visits.

The correlation between the predicted and observed utilization groups using Kendall's tau-b was used to estimate validity. The predicted group classification demonstrated a fairly low but significant correlation with the observed group (Kendall's tau-b = .206, $p = .000$). The same comparison with the chi-square demonstrated a significant association between the two groupings ($\chi^2 = 20.03$, $p = .000$, $df = 1$). Since the validity coefficient was determined to be .206, the original HRU model was not determined to be valid.

Initiatives to Improve the Original HRU Model

Recalibration of the Cutpoints in Original HRU Model

A closer examination of Table III-4 revealed a possible approach to improve the original HRU model. It was observed that the original HRU model classified 12.2% of the participants into the high utilization group, however, 21.2% were observed to be in the actual high utilization group according to the Pareto analysis (Table IV-1). Even if all of the 12.2% of the participants were correctly classified into the high utilization group, 9.0% would still be missed. Subsequently, the frequencies of each sum score from the original HRU model were examined (Table III-4) and it was observed that 21.2% of the participants received a sum score of four or more. Rather than using five or more as the cut point to classify participants into the high utilization group as was used with the first analysis of the original HRU, the cut point was recalibrated downwards to four or more visits. The resulting classification was compared with dependent variable and the results

are shown in Table IV-5. The correlation between the predicted and observed utilization groups was used to estimate validity (Kendall's tau-b = .174, $p = .003$). The same comparison with the chi-square demonstrated a significant association between the two groupings ($\chi^2 = 11.848$, $p = .001$, $df = 1$). Consequently, the validity coefficient for this revision of the original HRU model was determined to be .174. This revision failed to improve upon the performance of the original HRU model, which demonstrated a validity coefficient of .206.

Table IV-5

Classification Table of PCM Visit Utilization Groups Using the Original HRU Model With Revised Cutpoints

Actual visit utilization [a]	Original HRU classification [b]		
	low	high	Accuracy (% correct)
low	254 [c]	54 [d]	82.5% specificity
high	54 [d]	29 [c]	34.9% sensitivity
			72.4% overall correct

Note. $N = 391$.
[a] Actual visit utilization is the dependent variable that was determined by the Pareto analysis for the actual number of primary care manager visits utilized by the participants in 1997.
[b] Original HRU classification was determined from the scoring of the original HRU model.
[c] Correctly classified and the last column reflects this as a percentage of the row total
[d] Incorrectly classified.

Original HRU Model With Improved Reliability

An option in the reliability analysis was selected when the procedure was performed that indicated item-by-item how much the alpha would be increased if the item

were removed. Subsequently, a step-wise approach to reliability analysis was employed whereby the item that indicated that its deletion would improve the reliability the most was deleted and the reliability analysis was run again with one fewer item in the analysis. Again the reliability analysis results were examined for potential improvement in the reliability coefficient. This step-wise reliability analysis approach was repeated step-by-step until no further improvement of the reliability coefficient was shown possible. While it was demonstrated that the reliability of the original HRU model could be improved to .6125 by eliminating the four variables: gender, marital status, smoking, and alcohol (Table IV-6), it was still not determined to be reliable.

Table IV-6

Stepwise Reliability Analysis of Original HRU Model

Reliability coefficient	Number of items in the analysis				
	16	15	14	13	12
alpha	.5611	.5818	.5949	.5988	.6125
standardized alpha	.5892	.6164	.6425	.6464	.6627
Variables	**Alpha if item deleted**				
female gender	.5818	—			
single marital status	.5775	.5949	—		
smoking	.5645	.5846	.5988	—	
alcohol use			.5967	.6125	—

Note. The results from each step in the step-wise reliability analysis procedure are listed under the columns headed, "Number of items in the analysis." "—" indicates the variable eliminated at each step in the stepwise analysis. The only variables are listed in the lower half of the table that demonstrated an alpha score greater than the observed alpha in the upper half of the table.

Subsequently, the four variables were removed from the original HRU model and the remaining 12 items were added together to arrive at a new sum score (Table IV-7).

The frequencies of each sum score were examined. It was determined that 16.4% ($n = 64$) of the participants received a sum score of three or more, but they accounted for only 22.9% ($n = 362$) of the visits. That was used as the cut point for the dichotomous classification variable, utilization groups, with the more reliable revision of the Original HRU Model.

Table IV-7

Classification of Participants into Utilization Categories by a Revision of the Original HRU Model with Improved Reliability

		Persons			Visits		
Utilization	HRU Sum	No.	%	Cum. %	No.	%	Cum. %
low	0	156	39.9%	39.9%	556	35.2%	35.2%
low	1	112	28.6%	68.5%	425	26.9%	62.1%
low	2	59	15.1%	83.6%	236	14.9%	77.1%
Sub-Total		327			1,217		
high	3	37	9.5%	93.1%	157	9.9%	87.0%
high	4	14	3.6%	96.7%	103	6.5%	93.5%
high	5	6	1.5%	98.2%	37	2.3%	95.9%
high	6	2	0.5%	98.7%	14	0.9%	96.8%
high	7	5	1.3%	100.0%	51	3.2%	100.0%
Sub-Total		64	16.4%		362	22.9%	
Total		391			1,579		

Note. HRU Sum = score computed from the original HRU model with a range of scores between 0 and 12 possible. The HRU Sum score is used to classify the participants into 3 utilization groups.

Then the classification ability of this revised HRU model was compared with dependent variable and the results are shown in Table IV-8. The correlation between the predicted and observed utilization groups was used to estimate validity (Kendall's tau-b = .176, $p = .003$). The same comparison with the chi-square demonstrated a significant

association between the two groupings ($\chi^2 = 12.118$, $p = .000$, $df = 1$). This revision of the original HRU model yielded a validity coefficient of .176 and a reliability coefficient of .6125. The revision failed to improve upon the performance of the original HRU model.

Table IV-8

Classification Table of PCM Visit Utilization Groups Using Original HRU Model With Improved Reliability

Actual visit utilization [a]	Original HRU classification [b] low	high	Accuracy (% correct)
low	268 [c]	40 [d]	87.0% specificity
high	59 [d]	24 [c]	28.9% sensitivity
			74.7% overall correct

Note. $N = 391$.
[a] Actual visit utilization is the dependent variable that was determined by the Pareto analysis for the actual number of primary care manager visits utilized by the participants in 1997.
[b] Original HRU classification was determined from the scoring of the original HRU model with improved reliability.
[c] Correctly classified and the last column reflects this as a percentage of the row total
[d] Incorrectly classified.

Derived HRU Model

As many variables were constructed from the HEAR survey items as possible. The construction of the variables was detailed in Chapter III. The constructed variables were classified into six groups: demographic information, general health-related information, emotional health, medical services, disease conditions, and family health. The data dictionary detailing variables and the levels of each constructed variable by their

classification group is provided in Appendix E, Table E-1 and their frequencies are provided in Appendix E, Table E-2. Considering all of the variables within each group to be used as items in a subscale for the group, the reliability of the items was estimated using Cronbach's alpha procedure. The scores from the final subscales were added together to produce a derived HRU sum score. Participants were classified into one of two utilization groups by the derived HRU sum score.

Construction of Subscales and Reliability Analyses

Subscales were constructed from five of the six groups of variables: health, mental and emotional health, medical services, and disease conditions. The variables in the demographic information group were not combined into a subscale since they were not conceptually related to each other. None of the three demographic variables were used in the construction of the derived HRU model. Recall that the target population and the resulting study sample included mostly military wives. Since one demographic variable was marital status and a second was gender, these did not discriminate much between the participants. The remaining demographic variables, race/ethnicity and age, did not fit logically into the sum scoring approach. Construction of each of the five subscales began with including all of the constructed variables available from the appropriate groups in a step-wise reliability analysis. Variables were eliminated one at a time from each subscale until the maximum reliability was reached.

Reliability analysis with the subscales. The reliability of each subscale using all of the variables was examined using Cronbach's alpha. Since there are both dichotomous variables as well as ordinal variables with several levels of the variables used in the

subscales, the Cronbach's alpha was the correct procedure to select for the reliability analysis. The same step-wise approach to the reliability analysis that was described under the discussion for the original HRU model was used for the derived HRU model. That approach also guided the construction of the most parsimonious scales as possible. Eliminating items that would result in an improved reliability maximized the reliability of each subscale. The results of the step-wise reliability analyses are shown in Tables IV-9 to IV-13. Both the alpha and the standardized alpha are shown in the tables. The "standardized item variance is the α value that would be obtained if all the items were standardized to have a variance of 1.... If items on the scale have widely differing variances, the two α 's may differ substantially" (Norušis, 1997, pp. 107-108).

Table IV-9

Step-Wise Reliability Analysis With the Variables Available to Measure General Health

Reliability	No. of items in the analysis		
Coefficient	8	7	6
alpha	.4596	.4760	.5009
standardized alpha	.4905	.5025	.5200
Variables	**Alpha if item deleted**		
physical activity	.4760	—	
smoking		.5009	—

Note. The results from each step in the step-wise reliability analysis procedure are listed under the columns headed, "Number of items in the analysis." "—" indicates the variable eliminated at a particular step in the stepwise analysis. The only variables are listed in the lower half of the table that demonstrated an alpha score greater than the observed alpha in the upper half of the table.

Table IV-10

Step-Wise Reliability Analysis With the Variables Available to Measure Emotional Health

Reliability Coefficient	No. of items in the analysis							
	18	17	16	15	14	13	12	11
alpha	.7917	.7998	.8125	.8175	.8214	.8259	.8308	.8362
standardized alpha	.7993	.7941	.7911	.7963	.8028	.8087	.8191	.8300
Variables	**Alpha if item deleted**							
H.2 days drinking	.7998	—						
H.3 drinks/day		.8125	—					
J.3 family separation	.7944	.8035	.8175	—				
H.4 drink drive	.7930	.8021	.8156	.8214	—			
H.5 cut-down		.8008	.8153	.8211	.8259	—		
H.7 guilty	.7921	.8010	.8146	.8203	.8251	.8308	—	
H.6 complaints	.7921	.8010	.8145	.8201	.8246	.8301	.8362	—
H.8 five drinks			.8129	.8186	.8234	.8292	.8355	.8426
K.5 anxiety attacks				.8176	.8221	.8275	.8366	.8404

Reliability Coefficient	No. of items in the analysis							
	16	15	14	13	12	11	10	9
alpha							.8426	.8490
standardized alpha							.8441	.8569
Variables	**Alpha if item deleted**							
H.8 five drinks	.8129	.8186	.8234	.8292	.8355	.8426	—	
K.5 anxiety attacks		.8176	.8221	.8275	.8366	.8404	.8490	—

Note. The results from each step in the step-wise reliability analysis procedure are listed under the columns headed, "Number of items in the analysis." "—" indicates the variable eliminated at a particular step in the stepwise analysis. The only variables are listed in the lower half of the table that demonstrated an alpha score greater than the observed alpha in the upper half of the table.

Table IV-11

Step-Wise Reliability Analysis With the Variables Available to Measure Medical Services

Reliability	No. of items in the analysis		
Coefficient	6	5	4
alpha	.6202	.6340	.6450
standardized alpha	.5953	.6258	.6485
Variables	Alpha if item deleted		
K.6 mental health treatment	.6340	—	
N.16 musculoskeletal visits		.6450	—

Note. The results from each step in the step-wise reliability analysis procedure are listed under the columns headed, "Number of items in the analysis." "—" indicates the variable eliminated at a particular step in the stepwise analysis. The only variables are listed in the lower half of the table that demonstrated an alpha score greater than the observed alpha in the upper half of the table.

Table IV-12

Step-Wise Reliability Analysis With the Variables Available to Measure Disease Conditions

Reliability Coefficient	No. of items in the analysis			
	13	12	11	10
alpha	.4774	.5023	.5059	.5074
standardized alpha	.5297	.5025	.5253	.5210
Variables	**Alpha if item deleted**			
B hypertension	.5023	—		
N.6 neurological disease	.4817	.5059	—	
N.2 stroke		.5030	.5074	—

Note. The results from each step in the step-wise reliability analysis procedure are listed under the columns headed, "Number of items in the analysis." "—" indicates the variable eliminated at a particular step in the stepwise analysis.

Table IV-13

Step-Wise Reliability Analysis With the Variables Available to Measure Family Health

Reliability Coefficient	No. of items in the analysis	
	2	1
alpha	.2340	.3146
standardized alpha	.2768	
Variables	**Alpha if item deleted**	
N.17 sick family member	.3146	—

Note. The results from each step in the step-wise reliability analysis procedure are listed under the columns headed, "Number of items in the analysis." "—" indicates the variable eliminated at a particular step in the stepwise analysis.

The final reliability for each of three subscales (general health, medical services, and disease conditions) was between .5200 and .6485, while a fourth scale (emotional health) demonstrated a reliability of .8569. The fifth subscale, family health, was eliminated from further consideration for the derived HRU model since its reliability could not be improved to be higher than .500. Using the standardized alpha score, the final reliability of the general health subscale was .5200 with six variables: weight, health status, cholesterol, days in bed, days from job, and walking. The reliability was .8569 for the emotional health subscale with nine variables retained: stress frequency, stress amount, stress effect, satisfaction, family problems, disinterest, depressed, anxious, and worry. The reliability coefficient for the medical services subscale was .6485 with four variables retained: prescriptions, visits past month, ER visits past year, and hospitalization. The fourth and final subscale was disease conditions and its reliability coefficient was .5210 with ten variables retained: diabetes, heart disease, emphysema/bronchitis, arthritis, depression, cancer, liver, kidney, ulcer, and asthma.

Measures of association between variables in each subscale. Measures of association between the retained variables were examined using Kendall's tau-b. It was determined a-priori that variables with correlation coefficients greater than .85 would require closer examination. Although a number of significant correlation coefficients were observed, multicollinearity was ruled out since all of them failed to demonstrate a correlation above .500. Several inter-item correlation coefficients greater than .400 were observed among the mental and emotional health variables.

The correlation coefficients for each subscale are shown in Tables IV-14 to IV-17. Inter-item correlation coefficients tended to be low on the Health subscale except for the

correlation between days in bed and days out of the job (Kendall's tau-b = .533, p = .000). Items on the Emotional Health subscale tended to be well correlated with each other. The amount of stress and the frequency of stress showed the highest correlation (Kendall's tau-b = .684, p = .000). Correlation coefficients between items on the Medical Service subscale ranged from .111 to .427. Correlation coefficients on the Disease Condition subscale also were fairly low with the highest correlation observed between kidney disease and liver disease (Kendall's tau-b = .328, p = .000). The screening for multicollinearity failed to identify any items to be considered for elimination from any of the subscales.

Table IV-14

Correlation Between Variables in the Health Subscale

	(1)	(2)	(3)	(4)	(5)
1. weight	1.000				
	.				
2. health status	.243**	1.000			
	.000	.			
3. cholesterol	.009	.156**	1.000		
	.860	.001	.		
4. days in bed	.055	.112*	.076	1.000	
	.272	.017	.130	.	
5. days from job	.129**	.101*	.028	.533**	1.000
	.010	.031	.572	.000	.
6. walking	.078	.113	-.014	.050	.133**
	.126	.017	.785	.304	.008

Note. The first row for each variable is the test statistic for Kendall's Tau-b and the second row is the significance level.
*p < .05. **p < .01.

Table IV-15

Correlations Between Variables in the Mental and Emotional Health Sub-Scale

	(1)	(2)	(3)	(4)	(5)	(6)
1. stress frequency	1.000					
	.					
2. stress amount	.684**	1.000				
	.000	.				
3. stress effect	.471**	.458**	1.000			
	.000	.000	.			
4. satisfaction	.428**	.419**	.399**	1.000		
	.000	.000	.000	.		
5. family problems	.390**	.363**	.368**	.420**	1.000	
	.000	.000	.000	.000	.	
6. disinterest	.263**	.244**	.317**	.315**	.297**	1.000
	.000	.000	.000	.000	.000	.
7. depressed	.324**	.359**	.334**	.355**	.294**	.479**
	.000	.000	.000	.000	.000	.000
8. anxious	.390**	.428**	.357**	.263**	.250**	.294**
	.000	.000	.000	.000	.000	.000
9. worry	.429**	.470**	.382**	.297**	.274**	.260**
	.000	.000	.000	.000	.000	.000

	(7)	(8)
7. depressed	1.000	
	.	
8. anxious	.494**	1.000
	.000	.
9. worry	.420**	.483**
	.000	.000

Note. The first row for each variable is the test statistic for Kendall's Tau-b and the second row is the significance level.
*$p < .05$. **$p < .01$.

Table IV-16

Correlations Between Variables in the Medical Service Subscale

	(1)	(2)	(3)
1. prescriptions	1.000		
	.		
2. visits past month	.427**	1.000	
	.000	.	
3. ER visits past year	.135**	.176**	1.000
	.005	.000	.
4. Hospitalization	.111*	.131**	.245**
	.021	.006	.000

Note. The first row for each variable is the test statistic for Kendall's Tau-b and the second row is the significance level.
*$p < .05$. **$p < .01$.

Table IV-17

Correlations Between Variables in the Disease Condition Subscale

	(1)	(2)	(3)	(4)	(5)	(6)
1. diabetes	1.000					
	.					
2. heart disease	.196**	1.000				
	.000	.				
3. emphysema/	.105**	-.022	1.000			
bronchitis	.038	.659	.			
4. arthritis	.198**	.208**	.118*	1.000		
	.000	.000	.020	.		
5. depression	.098	.131**	.044	.173**	1.000	
	.054	.010	.381	.001	.	
6. cancer	-.037	.154**	.129*	.164**	.114*	1.000
	.459	.002	.011	.001	.024	.
7. liver	.106*	-.009	-.019	.063	.066	.182**
	.037	.860	.703	.215	.195	.000
8. kidney	.106*	-.009	-.019	.156**	-.030	-.013
	.037	.860	.703	.002	.552	.790
9. ulcer	.118*	.107	.201**	.092	.098	.227**
	.020	.034	.000	.068	.053	.000
10. asthma	.142**	-.038	.249**	.188**	.148	-.007
	.005	.447	.000	.000	.004	.896

	(7)	(8)	(9)
7. liver	1.000		
	.		
8. kidney	.328**	1.000	
	.000	.	
9. ulcer	.130*	-.018	1.000
	.010	.720	.
10. asthma	-.033	.055	.078
	.511	.275	.124

Note. The first row for each variable is the test statistic for Kendall's Tau-b and the second row is the significance level.
*$p < .05$. **$p < .01$.

Scoring of the subscales. Subscale scores were determined by adding the values of the variables in the scales together. Since the coding of each variable was anchored at zero to indicate the absence of a condition or the better state of the variable, the minimum

possible score for each subscale was zero. The coding of each variable increased by consecutive integers up to the maximum value of the variable. The total score possible on each subscale was determined by adding the maximum value possible on each of the variables in each subscale. The maximum possible values for each variable in each subscale is displayed in Table IV-18 along with the total possible score for each subscale.

Table IV-18

Maximum Score of Each Subscale in the Derived HRU Model

General health subscale		Medical services subscale	
Max.	Variable	Max.	Variable
1	weight	3	prescriptions
4	health status	4	visits past month
1	cholesterol	4	ER visits past year
4	days in bed	4	Hospitalization
4	days from job	15	Total possible score
1	walking		
15	Total possible score		

Emotional health subscale		Disease conditions subscale	
Max.	Variable	Max.	Variable
3	stress frequency	1	diabetes
3	stress amount	1	heart disease
2	stress effect	1	emphysema/ bronchitis
3	satisfaction	1	arthritis
3	family problems	1	depression
1	disinterest	1	cancer
1	depressed	1	liver
1	anxious	1	kidney
1	worry	1	ulcer
18	Total possible score	1	asthma
		10	Total possible score

Note. Max. = maximum possible value of the particular variable.

Construction of Derived HRU Model

The design of the derived HRU model was similar to the design of the original HRU model. First, a derived HRU sum score was calculated and then the sum score was used to classify participants into one of two utilization groups. By adding the maximum values of the subscales together, the maximum possible value of the sum score was determined to be 58 (Table IV-18). The derived HRU sum score was calculated for each participant by adding the subscale scores together. The frequencies of each sum score were examined and it was determined that 19.9% of the participants received a sum score of 19 or more and accounted for 27.2% of the visits (Table IV-20). That most closely matched the dependent variable determined from the Pareto analysis. Subsequently it was decided to classify participants scoring 19 – 58 to the high utilization group (Table IV-20) and participants scoring 0 – 18 to the low utilization group (Table IV-19). The Cronbach's alpha for the derived HRU model using the variables from four subscales was determined to be .8162. Since the reliability coefficient was .8162 using the standardized item alpha, the derived HRU model was not determined to be reliable.

Table IV-19

Low Utilization Group - Frequencies of Derived HRU Model Sum Scores with Persons and Visits

dHRU	Persons			Visits		
Sum score	No.	%	Cum. %	No.	%	Cum. %
0	1	0.3	0.3	2	0.1	0.1
1	1	0.3	0.5	0	0.0	0.1
2	4	1.0	1.5	17	1.1	1.2
3	11	2.8	4.3	34	2.2	3.4
4	7	1.8	6.1	4	0.3	3.6
5	10	2.6	8.7	30	1.9	5.5
6	24	6.1	14.8	98	6.2	11.7
7	16	4.1	18.9	41	2.6	14.3
8	20	5.1	24.0	72	4.6	18.9
9	19	4.9	28.9	46	2.9	21.8
10	28	7.2	36.1	115	7.3	29.1
11	28	7.2	43.2	120	7.6	36.7
12	27	6.9	50.1	107	6.8	43.4
13	20	5.1	55.2	61	3.9	47.3
14	34	8.7	63.9	143	9.1	56.4
15	21	5.4	69.3	83	5.3	61.6
16	16	4.1	73.4	41	2.6	64.2
17	15	3.8	77.2	91	5.8	70.0
18	11	2.8	80.1	45	2.8	72.8
SubTotal	313			1,150		

Table IV-20

High Utilization Group - Frequencies of Derived HRU Model Sum Scores with Persons and Visits

dHRU	Persons			Visits		
Sum score	No.	%	Cum. %	No.	%	Cum. %
19	18	4.6	84.7	83	5.3	78.1
20	8	2.0	86.7	25	1.6	79.7
21	9	2.3	89.0	29	1.8	81.5
22	6	1.5	90.5	33	2.1	83.6
23	5	1.3	91.8	32	2.0	85.6
24	5	1.3	93.1	28	1.8	87.4
25	4	1.0	94.1	19	1.2	88.6
26	4	1.0	95.1	27	1.7	90.3
27	1	0.3	95.4	4	0.3	90.6
28	1	0.3	95.7	14	0.9	91.5
29	5	1.3	96.9	40	2.5	94.0
30	3	0.8	97.7	10	0.6	94.6
31	2	0.5	98.2	19	1.2	95.8
32	3	0.8	99.0	29	1.8	97.7
34	1	0.3	99.2	5	0.3	98.0
38	1	0.3	99.5	11	0.7	98.7
39	1	0.3	99.7	1	0.1	98.7
40	1	0.3	100.0	20	1.3	100.0
Total	78	19.9		429	27.2	

Note. dHRU = derived HRU. $N = 391$. Total number of visits = 1,539. The dHRU Sum score is used to classify the participants into 2 utilization groups. dHRU Sum Score was computed from the derived HRU model with a range of scores between 0 and 58 possible.

Validity of Classification with the Derived HRU Model

The ability of the derived HRU model to correctly classify the participants into utilization categories is displayed in Table IV-21. The Pareto principle puts the most attention on the vital few, which is the high utilization category in this research. The correct classification for that category is displayed in the right lower quadrant. The sensitivity of the derived HRU model was observed to be 34.9% (29 true positives) while

the specificity was observed to be 84.1% (259 true negatives). Further, the classification resulted in 49 false positives (15.9%) and 54 false negatives (65.1%). Overall, it was 73.7% successful in assigning the participants to the correct utilization group. The actual numbers can be examined from a different perspective. The derived HRU model classified a total of 78 participants into the high utilization group, but only 29 (27.1%) belonged in the group. The other 72.9% of the participants belonged in the low utilization group. One more perspective can be seen from another examination of Table IV-20. The 19.9% of the participants classified into the high utilization group by the derived HRU model were responsible for (27.1%) of the PCM visits. That can be compared to the 21.2% of the participants identified by the Pareto analysis (Table IV-1) who utilized 50.4% of the visits.

Table IV-21

Classification Table of PCM Visit Utilization Groups Using the Derived HRU Model

Actual visit utilization [a]	Derived HRU classification [b]		
	low	high	Accuracy (% correct)
low	259 [c]	49 [d]	84.1% specificity
high	54 [d]	29 [c]	34.9% sensitivity
			73.7% overall correct

Note. $N = 391$.

[a] Actual visit utilization is the dependent variable that was determined by the Pareto analysis for the actual number of primary care manager visits utilized by the participants in 1997.

[b] Derived HRU classification was determined from the scoring of the derived HRU model.

[c] Correctly classified.

[d] Incorrectly classified.

The predicted group classification demonstrated a fairly low correlation with the observed group (Kendall's tau-b = .195, $p = .001$). The same comparison with the chi-square demonstrated a significant association between the two predicted and observed groups ($\chi^2 = 14.828$, $p = .000$, $df = 1$). Subsequently, the validity coefficient for the derived HRU model was determined to be .195.

Comparison of the Original HRU Model with the Derived HRU Model

A Pareto analysis was performed on the PCM utilization data and revealed that 21.2% of the participants consumed 50.4% of the total visits. Those participants were considered to be the true members of the high utilization group. The ability of the original HRU model identify the members of the high utilization group was compared with the ability of the derived HRU model to predict the true members of the high utilization group (Table IV-22).

Table IV-22

Comparison of the Performance of the Original and Derived HRU Models

	HRU Model	
Metric	Original	Derived
Persons identified	12.2%	19.9%
& visits utilized	18.5%	27.2%
Specificity	91.2%	84.1%
Sensitivity	25.3%	34.9%
Reliability coefficient	.589	.816
Validity coefficient	.206	.195

The original Health Resource Utilization model was determined to have a reliability coefficient of .589 and a validity coefficient of .206. The sensitivity of the

original HRU model was observed to be 25.3% with a 91.2% specificity rate resulting in an overall accuracy rate of 77.2% of the participants correctly classified. Within the high utilization group only 21 of the 49 (42.9%) of the participants were correctly classified.

The derived Health Resource Utilization model was determined to have a reliability coefficient of .816 and a validity coefficient of .195. The sensitivity of the derived HRU model was observed to be 34.9% with a 84.1% specificity rate and an overall accuracy rate of 73.7% of the participants correctly classified. Within the high utilization group only 29 of the 78 (37.2%) of the participants were correctly classified.

CHAPTER V

CONCLUSIONS

The research examined the ability of the Health Enrollment Assessment Review (HEAR) survey instrument to classify participants into groups according to their anticipated health resource utilization (HRU). The Assistant Secretary of Defense for Health Affairs issued guidance that the HEAR survey should be implemented worldwide for all of the 8.4 million Military Health System beneficiaries who choose to enroll in TRICARE Prime, the military's adaptation of the HMO approach to managed care (Josephs, 1996). The HEAR survey is administered to TRICARE Prime enrollees when they enroll. The Health Resource Utilization (HRU) classification is scored from the HEAR survey and is designed to classify participants into one of three potential utilization groups.

Problem and Research Questions

Although the HEAR is being utilized system-wide, the HRU classification model was "developed with the expectation that they [it] would be validated sometime after deployment" (U. S. Air Force Office of Prevention and Health Services Assessment. 1999, p. 3). Neither of the two technical reports on the HEAR development (Halpern et al., 1994; Murray & Halpern, 1996) established that the utilization of health resources actually reflected the Pareto principle. While the HRU classification is routinely being reported to primary care managers on their assigned TRICARE Prime enrollees, it can be concluded that its reliability and validity had not been established to date. However, it was anticipated that the HRU measure would be validated at some point in the future.

This research attempted to fill the gap. Further, since the original HRU model was developed from the developer's review of the literature and expert opinion, an additional HRU model was derived systematically. The performance of the investigator-derived HRU model was compared to the performance of the original HRU model. The research questions identified for the study are repeated below:

1. Does a Pareto analysis of utilization of primary care manager office visits produce a meaningful criterion against which to evaluate the ability of the HEAR survey to classify participants into utilization groups?
2. What is the reliability and validity of the original Health Resource Utilization (HRU) model to classify participants into utilization groups?
3. Can a measure of health resource utilization be derived systematically from the HEAR survey data? If so, what is the reliability and validity of the derived Health Resource Utilization (HRU) model to classify participants into utilization groups?
4. Which, if either, of the models is best suited for use with the target population?

Assumptions and Limitations

Before considering conclusions, some of the limitations of the research will be summarized. The post-hoc or retrospective design of the research was a significant limitation. The survey was administered after the health services had been delivered. While the data reflected visits utilized during calendar year 1997, the HEAR survey was administered during the first half of calendar year 1998. The purpose of the HEAR survey was to predict which participants would be in a higher utilization group in the following year on a prospective basis. However, the retrospective design was considered to be of

value as a pilot study for a larger national study using similar utilization data as a criterion. Further, the retrospective research design was very similar to Benjamin-Coleman and Alexy (1999) who administered a survey to predict a dichotomous variable, past utilization. The investigators concluded that the SF-36 could be used to screen members of a variety of populations to identify members at-risk for future hospitalization.

The selection of the target population presented another limitation that resulted from the scope of the study being limited to one metropolitan area. While random sampling was used with a sufficient sample size to generalize to the target population, the target population was limited to the Hampton Roads metropolitan area of Southeastern Virginia. Parente et al. (1996) found that primary care delivered by practices in metropolitan areas were more expensive, generated more referrals, and spent more on laboratory tests than practices in rural areas.

The primary care utilization data itself were limited in scope. All visits were counted as one unit because length of visit data were not available. The primary care managers (PCM; primary health care providers) were not typical of other PCMs in the Military Health System. Neither active duty service members nor federal employees, the PCMs were civilian providers directly managed by a Navy contractor. However, the majority of TRICARE Prime enrollees across the country are assigned to military treatment facility primary care managers. Three of the contracted PCM locations were located in commercial space in the local community and not located on a military installation. Not only were they more conveniently located to many beneficiaries than typical military providers, the beneficiaries did not have to share any of the cost of the care out-of-pocket just as if they were receiving care from a military provider. Thus, there

were fewer financial incentives to discourage beneficiaries from seeking professional health services. Subsequently, caution is advised when considering generalizing the research beyond the target population.

Administering a survey by mail has inherent limitations. Two mailings produced 391 usable surveys resulting in an overall survey response (20.7%) that was less than originally anticipated. Although that response rate may appear to be relatively low, it could be considered reasonable taking into account that the participants were health services consumers who were randomly sampled to participate anonymously in a mailed survey. The return rate from general consumers can be expected to be lower than a return rate achieved with a more controlled administration of a survey instrument. In fact, the response rate reported for production use of the HEAR survey in several of the TRICARE regions ranged from 17% to 38% (T. Baker, personal communication, July 1, 1999).

Nonetheless, the possibility for response bias was analyzed. Comparing the sample to the target population with non-identifying data available on both revealed some significant differences. The study sample had significantly more members ($n = 308$, 21.2%) in the high utilization group than the target population did ($n = 2{,}608$, 17.2%). Since the study sample was not fully representative of the target population, it was concluded that it would be inappropriate to generalize the research results beyond the current study to the target population.

Collecting self-reported survey data was a limitation. It was assumed that participants who returned the surveys were consenting to participate. The cover letter and the notification of additional information that was mailed with the survey instruments clearly revealed that assumption to the potential participants. Further, it was assumed that

they would complete the survey and that their self-reported responses would be an honest reflection of reality. However, Aday et al. (1981) reported that clinical conditions are often under-reported in self report surveys.

Coding of the data and the statistical analyses required assumptions to be made. A non-parametric procedure, Kendall's tau-b, was used for correlations and it required few assumptions about the data (Norušis, 1997). The assumption of normality was not required. Assumptions were made throughout the coding of the variables. Each survey item and its available responses were examined for face validity. Assumptions were made when specifying the good or normal state of the responses. It was assumed that more robust variables with higher and more numerous levels of measurement would be more discriminating. A fundamental assumption was that higher scores on the HRU sum would correlate with higher utilization of primary care manager visits.

Pareto Analysis of Primary Care Manager Visit Utilization

In the health care field, it was found that the Pareto analysis procedure is often performed as part of a quality improvement initiative. Once a problem has been identified, a Pareto analysis can be used to sort the causes of the problem into priorities for quality improvement. The utilization of primary care visits was the focus in this research. Utilization rates in the sample can be compared to rates in the literature to determine if a utilization problem existed in the sample.

The grouping in the report of national ambulatory utilization rates that can be most closely compared to the sample was the category for females in the 25 – 44 year-old bracket (Schappert, 1998). The sample was 94.4 % female and 5.6% male (Table III-2).

The average age of the participants in the research at the time the first mailing began was 34.7 years of age (*S. D.* = 7.5) with 88.8% of the participants between 23 and 47 years of age. In the national report, females in the 25 – 44 year-old bracket had 3.6 visits per year while males in the same age bracket had 1.9 visits per year. Since visits in this group were not further divided, the percentage of visits to physician offices reported for females of all ages (83.0%) will be applied to the overall ambulatory utilization rate. That revealed that females in the 25 – 44 year-old bracket utilized 2.99 physician office visits per year . However, the utilization rate in the report for visits to physician offices included both primary and specialty care. The utilization rate in the study (4.04 primary care manager visits per year) was 35.1% higher than the office visit utilization rate (2.99 physician office visits per year) in the report by Schappert (1998). While the PCM visit data in the study reflected primary care only, it did include women's preventive health visits (D. Nagy, personal communication, August 11, 1999). Since specialty care visits were excluded from the sample utilization rate, but not the report utilization rate, the true gap between utilization rates was probably even greater. It was concluded that office visit utilization was higher in the sample than was evident in the national utilization statistics. The utilization rate in the sample demonstrated a potential for quality improvement and a Pareto analysis could be one of the quality improvement tools utilized.

The PCM visit data were subjected to a Pareto analysis to assign the participants to either a high or a low PCM visit utilization group. The Pareto principle holds that approximately 20% of the contributors to an effect are responsible for approximately 80% of the effect (Caldwell, 1994; Juran, 1992). The number of visits of each participant with his or her primary care manager was determined from a government database. The

participants were grouped by the number of visits that they had with their Primary Care Manager in calendar year 1997. Starting with the highest number of visits per person, a Pareto analysis was performed.

The Pareto analysis of the study sample revealed that 21.2% of the participants ($n = 83$) accounted for 50.4% of the PCM visits. While that did not strictly adhere to the 80/20 rule, it did demonstrate that the utilization of PCM visits was disproportionately distributed among participants in the sample. The literature supported the value of applying the Pareto principle to quality improvement initiatives in health service delivery even when the distribution varied from a 80/20 proportion (Carey & Teeters, 1995; Clark et al., 1998; Fields & Siroky, 1994; Juran, 1994; Ziegenfuss & McKenna; 1995). One study reported 20.0% of the diagnoses for otherwise unassigned deaths were responsible for 52.9% of the trauma deaths (Clark et al., 1998). Freeborn et al. (1990) reported that 26.1% of the study sample represented high users and accounted for 51% of all ambulatory care contacts and 47% of office visits. That contrasted with Yen et al. (1994) who used total health expenditures as the dependent variable and found that 20% of the participants accounted for 90.9% to 92.6% of the total costs. The distribution in the study was very similar to the distribution observed in Clark et al. and Freeborn et al., but not Yen et al.

The answer to the first research question was that the Pareto analysis of utilization of primary care manager office visits did produce a meaningful criterion against which to evaluate the ability of the HEAR survey to classify participants into utilization groups. Although a few participants were responsible for only half of the PCM visits consumed rather than the bulk, half of the visits was still a significantly disproportionate amount. It

could be beneficial to identify those participants prospectively. Individual attention to only 20% of a primary care provider's panel of patients could potentially impact 50% of the visit utilization. Members of the high utilization group utilized seven or more visits during 1997. They would merit a closer look by their PCM to determine if their health service needs were being satisfied in an optimal manner.

Alternatively, the distribution of visits also revealed that 36.6% of the participants utilized 70.8% of the primary care manager visits. That could be rounded to 35/70 or even more roughly to one-third / two-thirds. Assuming that approximately 1,500 beneficiaries might typically be assigned to one primary care manager, 36.6% would represent 549 beneficiaries. However, it was also observed that 13.8% of the beneficiaries (n = 207) did not have any visits. So, while a primary care manager may be responsible for 1,500 beneficiaries overall, only 1,293 are users of visits. Subsequently, the 549 beneficiaries would represent 42.5% of the patients that primary care managers see in their offices in a given year. While it has been noted that the Pareto principle does not strictly follow a 80%/20% distribution, considering 42.5% to be a vital few is a significant departure from the general intent of the Pareto principle. It can be recalled that Juran (1992) recommended that planning for the vital few proceed on an individual basis while planning for the useful many should proceed on a group basis. Targeting 42.5% of the users calls for broader population-based action rather than action targeted towards individuals. The Pareto principle is used to target individual action toward a manageable few. The 21.2% observed in the sample (n = 318) represents 24.6% of the users which is much closer to the intent of the Pareto principle.

Original Health Resource Utilization Model

The ability of the original Health Resource Model to classify participants into a group characterized by high utilization of PCM visits as designed by the HEAR developers was evaluated. A minor revision to the scoring design in the original HRU model was required by the retrospective design of the study; the variable for outpatient visits in the past year was eliminated. That reduced the maximum possible sum score to 16 from 17. For the study sample, the original Health Resource Utilization model was determined to have a reliability coefficient of . 589 and a validity coefficient of .200. The response to the second research question is that the original Health Resource Utilization (HRU) model produced an unreliable measure that was not valid. The sensitivity of the original HRU model was observed to be 25.3% with a 90.9% specificity rate and an overall rate of 77.0% of the participants correctly classified. Within the high utilization group only 21 of the 49 (42.9%) of the participants were correctly classified.

According to the Pareto principle, identification of the high utilization category is the most important category; they are the vital few to be targeted for individual attention. While the original HRU model classified participants correctly into the low utilization category, the ability to classify participants into the high utilization category was poor. Consequently, the ability to classify participants correctly into the high utilization was the aim of the exercise and the original formulation of the HEAR was not successful in achieving that aim.

Derived Health Resource Utilization Model

Next, the content of each HEAR survey item was carefully examined and variables were constructed to take maximal advantage of the measurement in the available data. The variables were used to construct five subscales. Only four subscales achieved a reliability of greater than 0.500: general health, emotional health, medical services, and disease conditions. The subscale sum scores were added together to produce the derived HRU sum score. The frequencies of the derived HRU sum scores were examined and it was determined that 19.9% of the participants received a sum score of 19 or more (Table IV-20). That most closely matched the dependent variable determined from the Pareto analysis that revealed that approximately 20% of the participants utilized about 50% of the primary care manager visits. Subsequently it was decided to classify participants scoring 19 – 63 into the high utilization group and participants scoring 0 – 18 into the low utilization group.

The derived Health Resource Utilization model was determined to have a reliability coefficient of .816 and a validity coefficient of .195. The sensitivity of the derived HRU model was observed to be 34.9% with an 84.1% specificity rate and an overall rate of 73.7% of the participants correctly classified. Within the high utilization group only 29 of the 78 (37.2%) of the participants were correctly classified. The response to the third research question is similar to the response to the second research question. The HRU model derived from the HEAR data produced a measure that was reliable, but not valid.

Individual Examination of the Top Users

It was recommended that the vital few identified by a Pareto analysis be singled out for individual attention (Juran, 1992). As an example, the top nine users identified by the Pareto analysis were examined individually and all were female. Four of the nine top users were classified as members of the high utilization group by both the original and HRU models. In fact, the top three users were classified into the low utilization group by both models.

The survey responses for the top four users will be reported in detail. The highest use was 27 visits utilized by one participant. Her survey responses revealed frequent stress, moderate stress, much effect from stress, frequent family problems, family separation greater than 30 days, and cancer. The second highest user (23 visits) was characterized by one to two days in bed during the past two weeks, stress sometimes, moderate stress, and some effect from the stress. Tied with second highest user, another participant was characterized by family separation and asthma, but had no ER visits. Both HRU models identified the fourth highest user (20 visits). She was characterized by fair health, high cholesterol, greater than seven days in bed and absent from work, five to six ER visits past year, visits for musculoskeletal problems, and arthritis. She also revealed stress sometimes, moderate stress, much effect from the stress, family problems sometimes, family separation, lack of interest, anxiety, worry, mental health treatment, and depression.

Several observations can be made from the survey results of the top four users. First, the profiles of the users vary widely. Subsequently, the clinical approach to each of these users would be quite different. Freeborn et al. (1990) found that the high users

tended to have chronic diseases with treatable symptoms and mentioned arthritis in particular. One of the users described had cancer, one had asthma, and one had depression and arthritis while one did not report any chronic conditions. All endorsed emotional health indicators. Freeborn et al. (1990) and McFarland et al. (1985) also found that mental health conditions were more highly represented among high users. The HEAR survey revealed a large number of indicators with the fourth highest user, but not many with the top three users.

Overall Conclusions and Recommendations

The answer to the fourth research question was that neither the original HRU model nor the derived HRU model was suited for the intended use with the target population (see Table IV-22). The research failed to establish the reliability and validity of either the original or the derived Health Resource Utilization models in classifying HEAR survey participants into a high utilization group for the purpose of individualized planning. This conclusion was consistent with the findings of the U. S. Air Force Office of Prevention and Health Services Assessment (1999) with the original HRU model. It is recommended that a moratorium be placed on reporting HRU classification results on routine HEAR reports until a reliable and valid measure can be established. The SF-36 survey could be considered for further research as a HRU classification tool. The use of an established health risk appraisal instrument could be reconsidered.

A Pareto analysis is inherently a retrospective process. Prospectively identifying the vital few beneficiaries who will utilize a high number of primary care visits in the following year may simply not be possible to accomplish with a self-reported survey

instrument in a manner that is valid. While it may be recommended to abandon using the HEAR survey to identify high users, the Pareto principle remains a compellingly simple approach for targeting action to decrease overall utilization of primary care manager visits. The Pareto analysis yielded a useful finding in that about 20% of the beneficiaries in the target population utilized about 50% of the visits. The literature generally supported the notion that past high utilization is a good predictor of future high utilization (Ash et al., 1989; Epstein & Cumella, 1988; Gruenberg et al., 1996; Newhouse et al., 1989; Van Vliet & Van De Ven, 1992).

The threshold where beneficiaries in the sample (as well as the target population) cross over into the high utilization group was determined to be seven primary care visits in a year. Concurrent observation to detect when beneficiaries approach or cross that threshold could be a quite simple, reliable, and valid methodology for identifying the vital few. That could be accomplished by a simple record review for all beneficiaries when they go to a visit with their primary care manager, whether it is a routine or acute visit. Further, McFarland et al. (1985) found that two to three years of high utilization by an individual was sufficient to establish him or her as a consistently high user. So even a retrospective review of records could be used to identify high users.

An entirely different approach to the Pareto principle could be beneficial as well. A Pareto analysis can only identify the vital few who consume a disproportionate amount of the visits. It could be appropriate for persons with significant medical problems to utilize a high number of visits in a year. As mentioned, Freeborn et al. (1990) found that high users tended to have chronic diseases with treatable symptoms. The Pareto analysis cannot distinguish when utilization is appropriate and when it is inappropriate. What is

probably of more concern than a high number of visits per se is utilization of professional primary care services when there is an appropriate alternative. A number of the TRICARE Regions in the Department of Defense employ strategies generally referred to as demand management strategies to offer alternatives to beneficiaries when indicated rather than a visit to a health professional. These include toll-free health care information services and providing self-care manuals to TRICARE Prime enrollees, one per household.

Recommendations for Future Research

Future research could be directed either toward using an instrument to classify participants into potential utilization groups or toward abandoning the use of a survey instrument in favor of concurrent observation. The SF-36 survey instrument could be considered as a classification instrument. Either approach could benefit by conducting qualitative analysis to characterize high users identified by a Pareto analysis. Having adequately characterized high users with qualitative analyses, research could continue with quantitative analyses using the identified descriptors. Then, a multi-modal approach could be considered.

Federal Policy Implications

In the fall of 1998, the Department of Defense formally established an Integrated Product Team to review the HEAR program. The team was established under the auspices of the TRICARE Management Activity, a field activity of the Assistant Secretary of Defense for Health Affairs. It is anticipated that this research and the validation study conducted by the U. S. Air Force Office of Prevention and Health

Services Assessment (1999) will provide input to the policy decision process currently underway about the future of the HEAR survey.

REFERENCES

Aday, L. A., Sellers, C., & Andersen, R. M. (1981). Potentials of local health surveys: A state-of-the-art summary. American Journal of Public Health, 71, 835-840.

Amsden, D. M., Butler, H. E., & Amsden, R.T. (1991). SPC simplified for services: Practical tools for continuous quality improvement. White Plains, NY: Quality Resources.

Ash, A., Porell, F., Gruenberg, L., Sawitz, E., & Besier, A. (1989). Adjusting medicare capitation payments using prior hospitalization data. Health Care Financing Review, 10(4), 17-29.

Bell, M., Rosebrough, B., & Wall, S. (1996). System documentation: Project HEAR - health enrollment assessment review. Battelle Memorial Institute, Statistics and Data Analysis Systems; 505 King Avenue; Columbus Ohio 43201-2693

Bell, M., Rosebrough, B., & Wall, S. (1997). User guide: Project HEAR - health enrollment assessment review version 1.2. Battelle Memorial Institute, Statistics and Data Analysis Systems; 505 King Avenue; Columbus Ohio 43201-2693

Beebe, J., Lubitz, J., & Eggers, P. (1985). Using prior utilization to determine payments for Medicare enrollees in health maintenance organizations. Health Care Financing Review, 6(3), 27-38.

Benjamin-Coleman, R. & Alexy, B. (1999). Use of the SF-36 to identify community dwelling rural elderly at risk for hospitalization. Public Health Nursing, 16, 223-227.

Borus, J. F., Olendzki, M. C., Kessler, L., Burns, B. J., Brandt, U. C., Broverman, C. A., & Henderson, P. R. (1985). The 'offset effect' of mental health treatment on ambulatory medical care utilization and charges. Archives of General Psychiatry, 42, 573-580.

Buczo, W. (1989). Hospital utilization and expenditures in a medicaid population. Health Care Financing Review, 11(1), 35-47.

Cafferata, G. L. (1987). Marital status, living arrangements, and the use of health services by elderly persons. Journal of Gerontology, (42), 613-618.

Caldwell, L.J. (1994). What's in your toolbelt. Industrial Management, 36(1), 12-13.

Carey, R. G., & Teeters, J. L. (1995). CQI case study: Reducing medication errors. The Joint Commission Journal on Quality Improvement, 21, 232-237.

Chetwynd, J., & Rayner, T. (1986, April 9). Impact of smoking on health care resource use. New Zealand Medical Journal, 230-232.

Clark, D. E., Cushing, B. M., Bredenberg, C. E. (1998). Monitoring hospital trauma mortality using statistical process control methods. Journal of the American College of Surgeons, 186, 630-635.

Cronbach, L. J. (1984). Essentials of Psychological Testing (4th ed.). Cambridge, MA: Harper & Row.

Diehr, P., Price, K., Williams, S. J., and Martin, D. P. (1986). Factors related to the use of ambulatory mental health services in three provider plans. Social Science Medicine, 23, p. 773-780.

Epstein, A. M., & Cumella, E. J. (1988). Capitation payment: Using predictors of medical utilization to adjust rates. Health Care Financing Review, 10(1), 51-69.

Fields, W. L. & Siroky, K. A. (1994). Converting data into information. Journal of Nursing Care Quality, 8(3), 1-11.

Freeborn, D. K., Pope, C. R., Mulloly, J. P., & McFarland, B. H. (1990). Consistently high users of medical care among the elderly. Medical Care, (28), 527-540.

Gay, L. R. (1987). Educational research: Competencies for analysis and application (3rd ed.). New York: Macmillan.

Gruenberg, L., Kaganova, E., & Hornbrook, M. C. (1996). Improving the AAPCC with health-status measures from the MCBS. Health Care Financing Review, 17(3), 59-75.

Gyrna, F. M. (1988). Quality Improvement. In Juran, J. M., & Gyrna, F. M. (Eds.), Juran's Quality Control Handbook (4th ed., pp. 22.1-22.74). New York: McGraw-Hill. TS156.J87 1988

Halpern, M. T., Murray, M. I., Palmer, C. S., Reblando, J. A., & Rust, S. W. (1994). Project HEAR: Health Enrollment Assessment Review, Phase One: Literature Review, Analysis, and Recommendations (Air Force Armstrong Laboratory Report Number AL/PS-TR-1995-0012). Air Force Armstrong Laboratory, Office for Prevention and Health Services Assessment, 2601 West Road, Suite 3, Brooks Air Force Base, TX, 78235-5241.

Isaac, S., & Michael, W. B. (1981). Handbook in Research and Evaluation (2nd ed.). San Diego, CA: EdITS.

Josephs, S. C. (1996). Policy for TRICARE Health Enrollment Assessment Review (HEAR) Survey (Health Affairs Policy 97-003). Washington DC: Office of the

Assistant Secretary of Defense for Health Affairs. Available: http://www.tricare.osd.mil/policy/fy97/hear9703.html.

Josephs, S. C. (1997). Policy Memorandum to Refine Policy for Priority Use of Medical Treatment Facilities by TRICARE Prime Enrollees. (Health Affairs Policy 97-041). Washington DC: Office of the Assistant Secretary of Defense for Health Affairs. Available: http://www.tricare.osd.mil/policy/fy97/mtf9741.html.

Juran, D. (1994). Achieving sustained quantifiable results in an interdepartmental quality improvement project. Journal on Quality Improvement, 20, 105-119.

Juran, J. M. (1964). Managerial breakthrough: A new concept of the manager's job. New York: McGraw-Hill.

Juran, J. M. (1975). The non-pareto principle: Mea culpa. Quality Progress, 17(5), 8-9.

Juran, J. M. (1988a). Companywide planning for quality. In Juran, J. M., & Gyrna, F. M. (Eds.), Juran's Quality Control Handbook (4th ed., pp. 6.1-6.52). New York: McGraw-Hill.

Juran, J. M., & Gyrna, F. M. (Eds.). (1988). Juran's Quality Control Handbook (4th ed.). New York: McGraw-Hill.

Juran, J. M. (1988b). Juran on planning for quality. New York: Macmillan.

Juran, J. M. (1992). Juran on quality by design: The new steps for planning quality into goods and services. New York: The Free Press, a Division of Macmillan.

Juran, J. M., & Gyrna, F. M. (Eds.). (1988b). Juran's Quality Control Handbook (4th ed.). New York: McGraw-Hill.

Kongstevdt, P. R. (1993). Preface. In Kongstevdt, P. R. (Ed.), The Managed Care Handbook (2nd ed.). Gaithersburg, MD: Aspen.

Kosloski, K., Austin, C., & Borgatta, E. (1987). Determinants of VA utilization. Medical Care, 25, 830-846.

Manton, K. G., & Stallard, E. (1992). Analysis of underwriting factors for AAPCC. Health Care Financing Review, 14(1), 117-132.

Martin, E. D. (1996). Policy for Surveys within the Military Health Services System (MHSS) (Health Affairs Policy 97-012). Washington DC: Office of the Assistant Secretary of Defense for Health Affairs. Available: http://www.tricare.osd.mil/policy/fy97/surv9712.html

McFarland, B. H., Freeborn, D. K., Mullooly, J. P., & Pope, C. R. (1985). Utilization patterns among long-term enrollees in a prepaid group practice health maintenance organization. Medical Care, 23, 1221-1233.

Murray, M. I., & Halpern, M. T. (1996). Project HEAR: Health Enrollment Assessment Review, Phase Two: Literature Review, Analysis, and Recommendations (Contract Number 200-92-0534). Air Force Armstrong Laboratory, Office for Prevention and Health Services Assessment, 2601 West Road, Suite 3, Brooks Air Force Base, TX, 78235-5241.

Newhouse, J. P., Manning, W. G., Keeler, E. B., & Sloss, E. M. (1989). Adjusting capitation rates using objective health measures and prior utilization. Health Care Financing Review, 10(3), 41-54.

Norušis, M. J., (1997). SPSS Professional Statistics 7.5. Chicago: SPSS.

Oberle, J. (1990). Quality gurus: The men and their message. Training: The magazine of human resources development, 27(1), 47-52.

Office of the Assistant Secretary of Defense for Health Affairs. (1999). Corporate Executive Information System [electronic database]. 5111 Leesburg Pike, Skyline 5, Suite 809, Falls Church, VA 22041.

Parente, S. T., Weiner, J. P., Garnick, D. W., Fowles, J, Lawthers, A. G., & Palmer, R. H. (1996). Profiling resource by primary-care practices; Managed medicare implications. Health Care Financing Review, 17(9), 23-42.

Pfadt, A., & Wheeler, D. J. (1995). Using statistical process control to make data-based clinical decisions. Journal of Applied Behavior Analysis, 28, 349-370.

Polit, D. F., & Hungler, B. P. (1987). Nursing research: Principles and methods (4th ed.). New York: Lippincott.

Quesenberry, C. P., Caan, B., and Jacobson, A. (1998). Obesity, health services use, and health care costs among members of a health maintenance organization. Archives of Internal Medicine, (158), 466-472.

Radecki, S. E., Kane, R.L., Solomon, D. H., Mendenhall, R. C., & Beck, J. C. (1988). Do physicians spend less time with older patients? Journal of the American Geriatrics Society, 36, 713-718.

Roos, N. P., Roos, L. L., Mossey, J., & Havens, B. (1988). Using administrative data to predict important health outcomes: Entry to hospital, nursing home, and death. Medical Care, 26, 221-239.

Saag, K. G., Doebbeling, B. N., Rohrer, J. E., Kolluri, S., Peterson, R., Hermann, M. E., and Wallace, R. B. (1998). Variation in tertiary prevention and health service utilization among the elderly. Medical Care, 36, 965-976.

Schappert, S. M. (1997). Ambulatory care visits to physician offices, hospital outpatient departments, and emergency departments, 1995. [Vital and health statistics series 13(129), DHHS publication no. (PHS) 97-1790]. Hyattsville, MD: National Center for Health Statistics.

Schappert, S. M. (1998). Ambulatory care visits to physician offices, hospital outpatient departments, and emergency departments, 1996. [Vital and health statistics series 13(134), DHHS publication no. (PHS) 98-1795]. Hyattsville, MD: National Center for Health Statistics.

SPSS. (1988). SPSS-X™ user's guide (3rd ed.) [Computer software manual]. Chicago: Author.

Statement on the status of military medicine: Hearings before the Subcommittee on Defense, of the House Committee of Appropriations, 105th Cong., 2nd Sess. (1998). (testimony of Edward D. Martin). Available: http://www.tricare.osd.mil/planning/congress/dahcstate.html.

Swift, J. A., Ross, J. E., & Omachonu, V. K. (1998). Principles of total quality (2nd ed.). Boca Raton, FL: St. Lucie Press.

TRICARE Management Activity (1999). TRICARE regions. Available: http://www.tricare.osd.mil/tricare/trimap2.html.

U.S. Air Force Office for Prevention and Health Services Assessment (1999). Health enrollment assessment review 1.0 (HEAR): High resource utilization (HRU) validation study final report. 2601 West Road, Suite 3, Brooks Air Force Base, TX, 78235-5241: Author.

Van Vliet, R. C. J. A. and Van De Ven, W. P. M. M. (1992). Towards a capitation formula for competing health insurers: An empirical analysis. Social Science Medicine, (9),1035-1048.

Vogt, T. M., & Schweitzer, S. O. (1985). Medical costs of cigarette smoking in a health maintenance organization. American Journal of Epidemiology, 122, 1060-1066.

Waltz, C. F., Strickland, O. L., & Lenz, E. R. (1991). Measurement in Nursing Research (2nd ed.). Philadelphia: F. A. Davis.

Wilson, J. (1963). Thinking with concepts. Cambridge: Cambridge.

Yen, L. T., Edington, D. W., & Witting, P. (1992). Prediction of prospective medical claims and absenteeism costs for 1284 hourly workers from a manufacturing company. Journal of Occupational Medicine, 34, 428-435.

Yen, L. T., Edington, D. W., & Witting, P. (1994). Corporate medical claim cost distributions and factors associated with high cost status. Journal of Occupational Medicine, 36, 505-515.

Ziegenfuss, J. T., & McKenna, C. K. (1995). Ten tools of continuous quality improvement: A review and case example of hospital discharge. American Journal of Medical Quality, 10, 213-220.

APPENDIX A

SAMPLE HEAR PCM REPORT

This sample Primary Care Manager Report informs the PCM of the HEAR responses of an individual and identifies possible chronic conditions, risk factors and preventative service recommendations.

888-00404
Primary Care Manager Report
Thursday, May 15, 1997

DMIS: 0100
PCM: 55555555555555

Tupper Gman
2602 Donothing Rd

Brooks, TX 78235

Import Date:	5/14/97
Name:	LAST, FIRST
D.O.B. (Age)	4/28/65 (32)
Social Security #	123-40-0404
Sponsor's SSN:	987-60-0404
Gender:	Female
Phone No:	

1. Clinical Preventive Services

Based on the Health Enrollment Assessment Review (HEAR) questionnaire the following clinical services are recommended for your patient. The following table indicates the currency of recommended clinical preventive screening tests. Please review and use your clinical judgement in addressing these preventive service requirements.

Clinical Preventive Service	Recommended Intervals	Last Exam Date	Current	Not Current	Never Performed
Blood Pressure Screening	q 2yrs.	3 or more years ago		X	
Cholesterol Screening	q 5 yrs.	1-2 years ago	X		
Tetanus Shot	q 10 yrs.		X		
Clinical Breast Exam	q 1 yr. (> age 39)	2 years ago	X		
Mammography	q 2 yrs. (40-49); q 1 yr. > 49	1 year ago	X		
PAP Smear	q 1 yr.**	3 or more years ago		X	
Rectal Exam	q 1 yr. (> age 39)	3 or more years ago	X		

* Based on TRICARE Prime Clinical Preventive Services Benefits package

** Every 3 years after 3 consecutive satisfactory normal annual pap smears

2. Counseling Services

Based on the information from the HEAR questionnaire your patient indicated the following high-risk behaviors.

The following table lists the patient's risk factors and risk factor result.

Risk Factor	Risk Factor Result
Smoking	1-10 per day
Alcohol Consumption	Patient reports frequent and/or excessive alcohol consumption
Drinking and Driving	Patient reports drinking and driving
Stress	Patient reports considerable stress at work and/or home
Family Separation/Family Problems/Marital Problems	Patient reports family separation, family problems, and/or marital problems
Depression	Patient reports depression symptom(s)
Anxiety	Patient reports anxiety symptom(s)
Smoking Cessation	Patient reports moderate readiness to stop smoking
High Blood Cholesterol	Patient reports being told of high blood cholesterol

888-00404
Primary Care Manager Report
Thursday, May 15, 1997

DMIS: 0100
PCM: 55555555555555

Tupper Gman
2602 Donothing Rd

Brooks, TX 78235

Import Date:	5/14/97
Name:	LAST, FIRST
D.O.B. (Age)	4/28/65 (32)
Social Security #	123-40-0404
Sponsor's SSN:	987-60-0404
Gender:	Female
Phone No:	

3. Chronic Conditions/Impairments
Based on the HEAR questionnaire your patient indicated the presence of the following chronic disease conditions.

No chronic conditions reported

4. Please Note

Risk Factor	Level
Patient self-reported health status	Very Good
Potential risk for family separation	No
Potential alcohol abuse	No
Number of prescription medications being taken	6 or more medications
Number of outpatient visits in the past year	16-20 visits
Number of hospitalizations in the past year	2-3 times
Family member with a serious illness	No
Primary Care Level [1]	Level 3
Risk for high resource utilization [2]	Level 3

[1] Primary Care Level indicates the complexity and level of medical care a patient will require.
Level 1 = Least Complex
Level 2 = Moderately Complex
Level 3 = Most Complex

[2] High Resource Utilization indicates the frequency and intensity of use of the medical delivery system.
Level 1 = Low resource utilization
Level 2 = Moderate resource utilization
Level 3 = High resource utilization

5. Missing/Incomplete Information

Your patient completed all portions of the HEAR questionnaire.

The layout of sample Primary Care Manager Report was modified minimally to fit within the margins of the paper.

APPENDIX B

COVER LETTER

TRICARE
Mid-Atlantic

5425 Robin Hood Road, Suite 203
Norfolk, Virginia 23513
(757) 677-6440
Fax (757) 677-6481

March 3, 1998

Dear TRICARE Prime Enrollee,

I am a student in the Ph.D. program in Urban Services at Old Dominion University. I am conducting a field test of the "Health Enrollment Assessment Review" (HEAR) survey for my doctoral dissertation research. You are invited to participate in the field test. The HEAR survey was designed by the Department of Defense as a part of the national TRICARE Prime program. Later this year, the HEAR survey will become part of the TRICARE Prime program in your local area. Then you may receive another HEAR survey to complete when you renew your TRICARE Prime enrollment.

Enclosed you will find your HEAR survey. The HEAR survey asks questions about your personal health background and current health. **Please begin with question A1 on page 5. Please do not complete the section on pages 3 and 4 that asks for name, address, and social security numbers.** The survey form has been pre-coded with a random number so that your survey can be analyzed anonymously. It will take approximately 10 - 15 minutes to complete the survey.

Please complete the HEAR survey and return it in the enclosed postage-paid envelope by Friday, March 13, 1998, if possible. If you cannot mail it by that time, please send it as soon as you can. By returning the survey, you are consenting to participate in this research project. Please be sure to read the "Notification of Additional Information – HEAR Field Test Research Project" on the back. I appreciate you taking the time to complete and return the HEAR survey. If you have any questions about your HEAR survey or anything else about the field test research, please contact me, Jody Donehoo, at (757) 314-6474 or (757) 482-8660 and I will be happy to assist you.

Sincerely,

JODY W. DONEHOO
Doctoral Student, Old Dominion University

ENDORSEMENT

I fully support this field test of the HEAR survey and encourage you to participate. The results of the HEAR survey you complete for this field test will only be used for this research and will not affect your health benefits in the TRICARE Program in any way.

G. R. HARMEYER
Director, TRICARE Mid-Atlantic
CAPTAIN, Nurse Corps, U. S. Navy

Notification of Additional Information – HEAR Field Test Research

This research has been approved by the Naval Medical Center Portsmouth and by Old Dominion University in accordance with all applicable protocols and laws. The principal investigator is Mr. Jody W. Donehoo. This research will serve as his doctoral dissertation. Under the direct supervision of his ODU dissertation committee, this dissertation research is being conducted in cooperation with the Naval Medical Center Portsmouth and TRICARE Mid-Atlantic.

Since the HEAR survey is already in full use in several of the TRICARE regions, this is not an experimental type of research and you are not undergoing any experimental procedures. Approximately 1,000 people will be invited to participate in this field test. The research feature of this field test is to try to understand of how effective the HEAR survey is in determining health care needs of TRICARE Prime enrollees. In particular, this research examines the ability of the HEAR to predict how much care they may need from their Primary Care Manager (PCM) during the year. For comparison, records will be reviewed to determine how many visits you had with your Primary Care Manager at the TRICARE Prime Clinic during 1997.

The questions are of a personal nature. It asks about your health, medical history, and personal health behaviors. The possible risks or discomforts are the same as with any HEAR survey. Information you provide is protected by the Privacy Act and other laws applicable to this research. Your anonymity is guaranteed in all publications and presentations resulting from this research project. The results of your individual survey will be used only for this research project. Authorized Navy Medical Department personnel, and authorized Old Dominion University personnel may have access to the research file in order to verify that your rights have been safeguarded. The results of your individual survey will not be provided to anyone else, added to any other database, or provided for your medical record.

If you have any questions regarding this research project including the results of your survey, you may contact Mr. Jody W. Donehoo at (757) 314-6474 or (757) 482-8660. If you have any questions regarding ethical considerations of this survey or your rights as an individual while participating, you may contact the Chairman, Institutional Review Board, Naval Medical Center Portsmouth. Alternatively, you may contact Captain M. Zajdowicz, Head, Clinical Investigation and Research Department at (757) 953-5939. You may also contact Valerian Derlega, Ph.D., Chairman, Human Subjects Review Board, Old Dominion University, at (757) 683-3118.

Your participation in this project is voluntary and your refusal to participate will not involve any penalty or loss of benefits to which you are entitled under applicable regulations. If you choose to participate by returning the survey, you are free to ask questions or to withdraw from the project at any time.

You may withdraw from this study at any time without prejudice to your future care. If you choose to withdraw, you need to notify Mr. Jody Donehoo at (757) 314-6474 or (757) 482-8660 to ensure an orderly termination process. Your withdrawal will involve no loss of benefits to which you are entitled.

APPENDIX C

HUMAN SUBJECTS REVIEW APPROVAL

No: 98-010
(reference 98-007 amended)

OLD DOMINION UNIVERSITY
HUMAN SUBJECTS INSTITUTIONAL REVIEW BOARD
NOTIFICATION OF EXEMPT RESEARCH

TO: Jody W. Donehoo
Principal Investigator

DATE: 17 Feb 97
~~98-010~~ *IRB Decision Date*

RE: Validation of the Health Resource (HRU) Scale of the HEAR Survey and Opportunities for Improvement
Name of Project

- Please be informed that your research proposal has been considered by the Institutional Review Board and was found to be EXEMPT from review for the following reason(s):

- Under 45 CFR 46.101(b)(2)(i)
(So long as information obtained is free of identifiers –)

- Note: Multiple institutions require review from multiple IRB's

- Contact the IRB for clarification of the terms of your research or if you wish to make any change to your research proposal.

- Your project expires in one year from the IRB Decision Date. You must seek re-approval if you wish to continue beyond that date.

Valerie J. Sribney 683-4439
Chairperson *phone*

6500
Ser 00B3/1258
4 March 98

MEMORANDUM

From: Head, Medical Analysis and Research Center (MARC)
To: Mr. Jody Donehoo, TRICARE Mid-Atlantic
Subj: APPROVAL AND FUNDING OF CLINICAL INVESTIGATION STUDY #P98-035, ENTITLED "VALIDATION OF THE HEALTH RESOURCE UTILIZATION (HUR) SCALE OF THE HEALTH ENROLLMENT ASSESSMENT REVIEW (HEAR) SURVEY AND OPPORTUNITIES FOR IMPROVEMENT"

Ref: (a) HSETCINST 6000.41A
(b) SECNAVINST 3900.39B
(c) SECNAVINST 5212.5C

Encl: (1) Revised Protocol and Consent Form
(2) Signatures for the Record of 27 February 1998
(3) Adverse Event Report
(4) Travel Request Letter
(5) Trip Report

1. Your CIP proposal was approved by the Commander, NMCP on 27 February 1998. Approval is documented in the minutes of the Institutional Review Board, DOD #40003. A copy of the revised proposal is forwarded as enclosure (1). Please be sure to use the approved consent form. A copy of the approval is forwarded as enclosure (2).

2. Reporting requirements of reference (a) are as follows:

a. **ADVERSE EVENT REPORT**, enclosure (3). Federal law **requires** you to report any adverse event a subject experiences during the course of your study. Reporting an adverse event does **not** reflect negatively on you. It ensures that all side effects, problems, and mortality are reported to regulatory agencies.

b. **PROGRESS**. A progress report to NSHS Code OC is required semi-annually on your study. Clinical Investigation Division will contact you to schedule and conduct a continuing review on your study if it involves human use. Failure to file this report upon request may result in immediate termination of your study. Patients names and sponsors' SSNs are to be included with this report. Appropriate record keeping for this is described in item 3, "Maintenance of Records".

c. **CHANGE**. Submit changes in your study; for example, change in investigator, unusual delays, change in objectives, or new funding requirements. A change report may be combined with the progress report provided it does not result in a delay of more than

30 days. Any change in the consent form must be approved by the IRB.

d. **TERMINATION REPORT**. Submit when the objectives are not met or when directed by NSHS. The reasons for termination should be fully justified and submitted within 30 days of the termination. Include the actual total dollar outlay (O&MN and OPN) since the study's inception through the proposed termination date and a summary report of the study's findings.

e. **COMPLETION REPORT**. Submit within 30 days of the study's completion. This should be an accurately prepared detailed summary containing an abstract and complete bibliography of all publications and presentations resulting from the study in a format suitable for publication.

3. **Maintenance of Records**

a. Research records are to be maintained as required by references (a), (b), and (c).

b. For human use studies, it is your responsibility to keep a copy of the **SIGNED** consent form for your records, give a copy to the subject, and file a copy in the subject's record. This includes the Privacy Act Statement, which is the last page of the consent. The consent must be signed by the subject, a witness, **and** either the principal investigator, associate investigator, or a sub-investigator.

c. Clinical Investigation Division must be notified in writing with an endorsement by your Department Head 30 days in advance of your separation from the Command. A progress/final report must be forwarded with that memorandum. All research records must be turned in to Clinical Investigation 30 days before your separation unless you are being replaced as the principal investigator. If a new principal investigator is assuming the responsibility for the study, he/she must be identified before your departure. Upon completion or termination o your study for any reason, the research records must be turned in to CIRD within 30 days.

4. Continuing review will be conducted at least annually on studies involving human use. You will be notified by Clinical Investigation a the time the review is due and given the appropriate forms to complete At that time, you should be prepared to provide copies of all of the consent forms, a list of patient identifiers for the study, and, if appropriate, an updated list of sub-investigators.

5. Publications and Presentations

The research data generated as a result of this approved CIP study are the property of the Department of the Navy and may not be released from within the Department without prior written approval. All abstracts, presentations, manuscripts, and review articles are to be forwarded for approval at the directorate level and MARC ***prior to***

submission for publication. These approvals can be obtained simultaneously.

a. Acknowledgment Credit

(1) Presentations and/or publications resulting from your study shall acknowledge and identify the Bureau of Medicine and Surgery Clinical Investigation Program as the sponsor of your study. The assigned CIP number shall also be included in any presentation, publication, or written reference to the study.

(2) The acknowledgment is to be written as follows: "This study has been sponsored and supported by the Bureau of Medicine and Surgery Clinical Investigation Program No. P98-L-HOOOOO-035:A.

6. Funding

a. Approval of your study does not automatically guarantee that funds are available. All requests for purchase of supplies, equipment, etc. must be submitted to your Department Head , or in some cases, MARC for processing. ANY ITEM, INCLUDING REPRINTS, ORDERED DIRECTLY FROM A SUPPLIER BY YOU IS CONSIDERED AN UNAUTHORIZED COMMITMENT FOR WHICH YOU WILL BE RESPONSIBLE. Plan ahead and contact Mrs. Janie Slade, X3-7001 for further information.

b. Travel funds for presentations are available from NSHS within allocated resources. Your request for travel funds must be submitted in the format of enclosure (4) to NSHS via MARC 8 weeks prior to travel dates. A trip report, enclosure (5), must be completed and forwarded to NSHS-CIP within 30 days after completion of your travel. Contact Mrs. Davidson, X3-5412, for further information.

M. ZAJDOWICZ
CAPT, MC, USN

Copy to:
MARC Serialization File
Protocol Record
NSHS

APPENDIX D

HEALTH ENROLLMENT ASSESSMENT REVIEW (HEAR)

SURVEY INSTRUMENT

DEPARTMENT OF THE AIR FORCE
AIR FORCE MEDICAL OPERATIONS AGENCY
BROOKS AIR FORCE BASE, TEXAS 78235-5249

12 MAR 1999

AFMOA/SGOH
2602 Doolittle Road, Bldg 804
Brooks AFB TX 78235-5249

Mr. Jody W. Donehoo
TRICARE Mid-Atlantic Region Office
5425 Robin Hood Road, Suite 203
Norfolk, VA 23513-2441

Dear Mr. Donehoo

The Health Enrollment Assessment Review (HEAR), copyrighted by the Air Force in 1997, may be referenced and reproduced in the appendix of your dissertation, "Bureau of Medicine and Surgery Clinical Investigation Program No. P98-LH00000-35:A." Inclusion of the survey requires complete reproduction, to include the copyright and Privacy Act statements, as well as the question set.

The Office for Prevention and Health Services Assessment would appreciate a final copy of the dissertation for inclusion in the HEAR Program documentation.

If you have any questions, please call Ms. Kathleen Sotello at 210-536-6515 or email: kathleen.sotello@ophsa.brooks.af.mil.

Sincerely

Russell W. Eggert

RUSSELL W. EGGERT, Lt Col, USAF, MC, FS
Chief, Office for Prevention and Health
Services Assessment

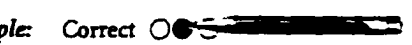
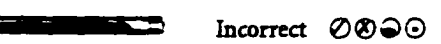
Health Enrollment Assessment Review

This questionnaire was developed by the Office for Prevention and Health Services Assessment (OPHSA), the National Center for Environmental Health (NCEH), and the Battelle Memorial Institute for TRICARE Regions 4 and 6 through a Memorandum of Agreement between Armstrong Laboratory, Human Systems Center, United States Air Force Materiel Command, and the Centers for Disease Control and Prevention (CDCP).

Marking Instructions

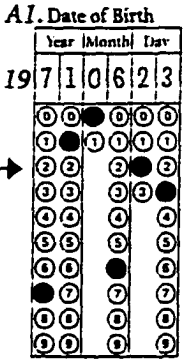
Health Enrollment Assessment Review (HEAR) Instructions

Please *use a No. 2 pencil* or darker to complete the survey. Make dark black marks that fill the response circles completely. If you make a mistake, erase the incorrect mark and fill in the correct circle.

Example: Correct  Incorrect 

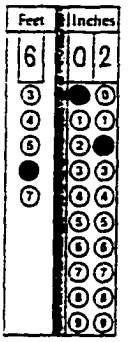
Here is an example of how someone born on June 23, 1971 would answer question A1. Notice that it is easier to darken the response circles if you write your answer in the boxes first. →

A1. Date of Birth



Here is an example of how someone 6 feet 2 inches tall would answer question A6. →

A6. About how tall are you, without shoes?



Please *answer all appropriate questions* and complete the entire survey. However, you should skip questions where the survey says to do so. For example, males should not answer the female questions, and non-smokers should not answer the smoking questions.

Example: In the illustration below, we have answered "not at all" to question G2. Therefore we will skip the rest of the G section questions and go directly to question H1.

G2. Do you *NOW* smoke cigarettes every day, some days, or not at all?

○ Every day ○ Some days ● Not at all *(go to H1)*

Do not fold or staple the survey pages.

Privacy Act Statement:

Authority: 10 U.S.C., 8013

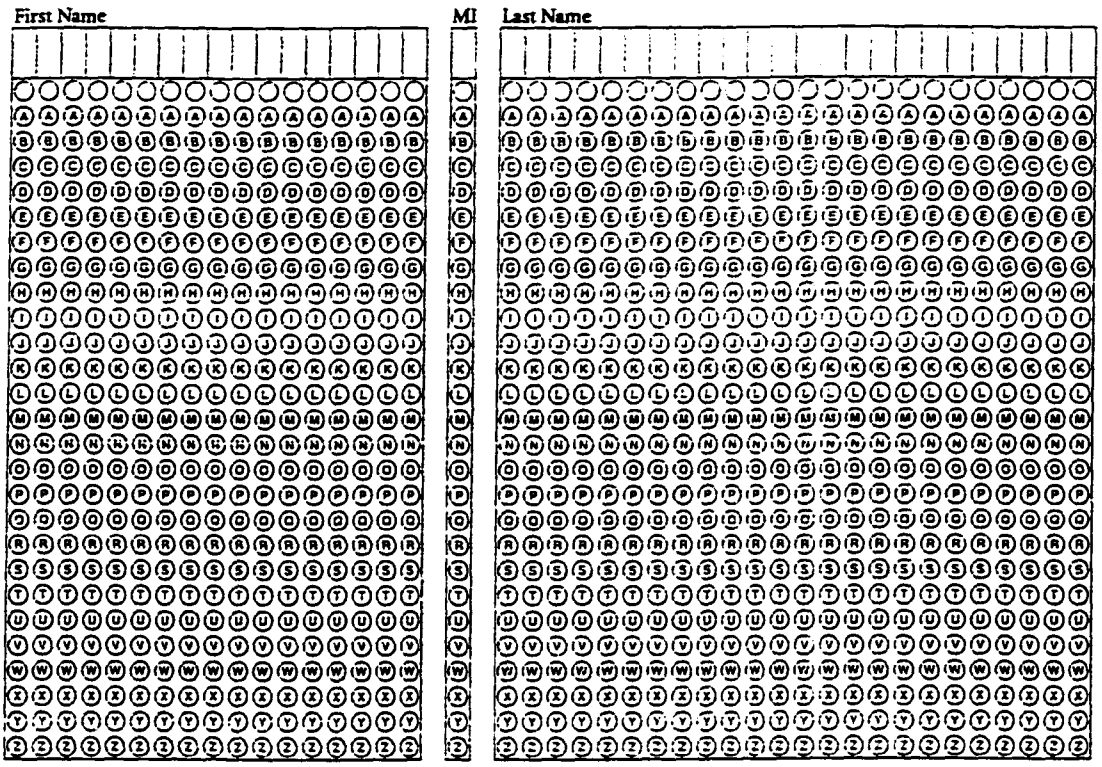
Purpose: The Health Enrollment Assessment Review (HEAR) survey is designed to collect personal information from military health services system beneficiaries.

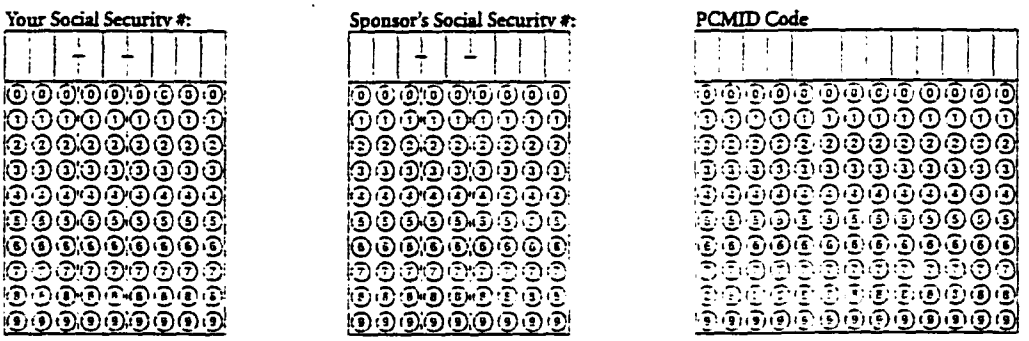
Routine Uses: This information is used primarily by health care personnel to plan health care delivery needs. Information used in this survey will be sent only to you and your Primary Care Manager (PCM) and kept in your medical record. Other results from this survey will be provided only in combination with results from other enrollees and cannot be used to identify you.

Disclosure: Completion of information in this survey is highly desirable, but not mandatory. Completion of the survey information will help your PCM design a plan of care. Preexisting medical conditions and other risk factors will in no way affect enrollment eligibility.

Thank you for completing the survey.

Health Enrollment Assessment Review Questionnaire

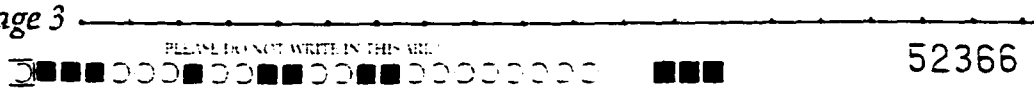




For Office Use Only

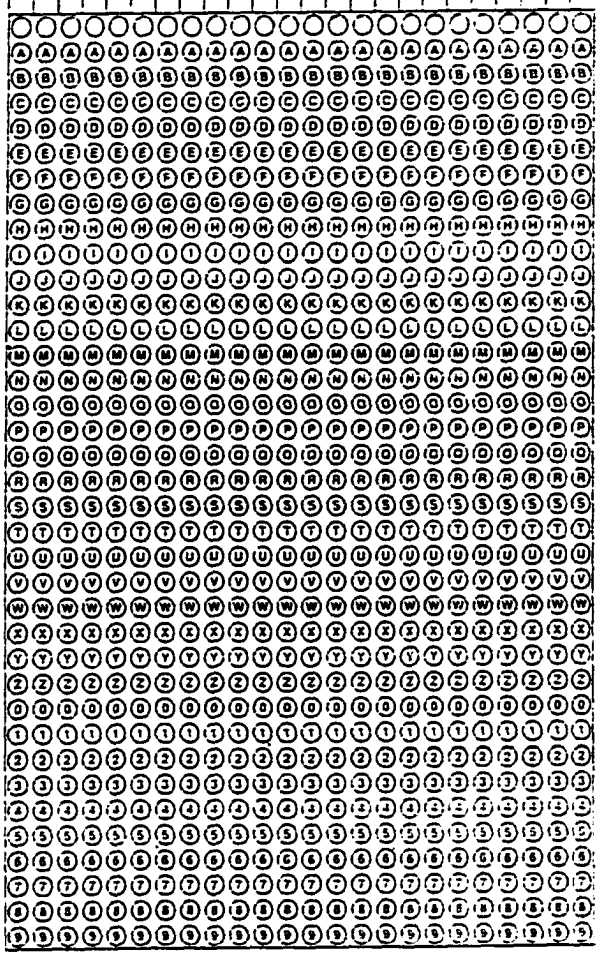
Please do not write or mark in this area

Page 3

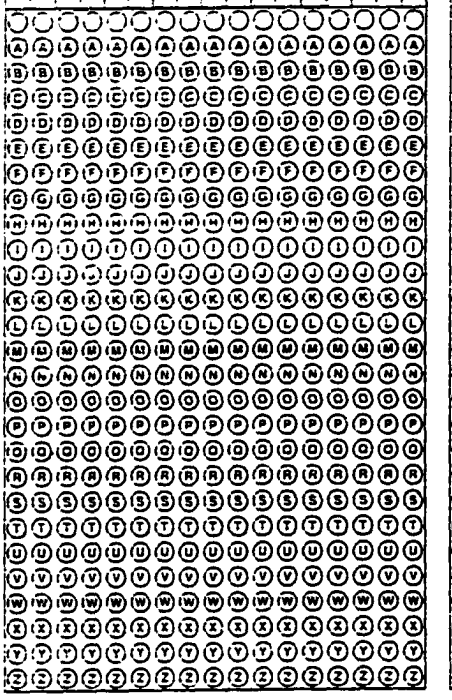


Health Enrollment Assessment Review Questionnaire

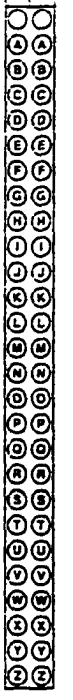
Street Address (include apartment #)



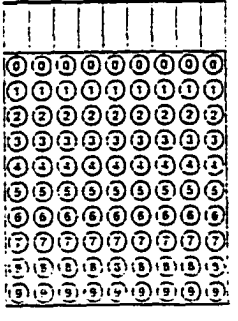
City



State



Zip Code + 4



For Office Use Only

Please do not write or mark in this area

Health Enrollment Assessment Review Questionnaire

A1. Date of Birth

Year | Month | Day

19 ☐☐ ☐☐ ☐☐

A2. Gender
- ○ Male
- ○ Female

A3. Marital Status
- ○ Never married
- ○ Married
- ○ Separated
- ○ Divorced
- ○ Widowed

A4. Racial/Ethnic Background
- ○ Amer. Indian or Alaska Native
- ○ Asian/Oriental
- ○ Black, Hispanic
- ○ Black, Non-Hispanic
- ○ Pacific Islander
- ○ White, Hispanic
- ○ White, Non-Hispanic
- ○ Other

A5. Are you:
- ○ Active duty service member
- ○ Retired service member

or Family Member of:
- ○ Active duty service member
- ○ Retired/deceased service member

or
- ○ Other

A6. About how tall are you, without shoes?

Feet | Inches

☐ ☐☐

A7. About how much do you weigh, without shoes?

Pounds

☐☐☐

A8. Would you say that your health in general is...
- ○ Excellent
- ○ Very good
- ○ Good
- ○ Fair
- ○ Poor

B1. About how long has it been since you last had your blood pressure taken by a doctor, nurse, or other health professional?
- ○ Less than 1 year ago
- ○ 1 year ago
- ○ 2 years ago
- ○ 3 or more years ago
- ○ Never
- ○ Don't know

B2. Have you ever been told by a doctor or other health professional that you had hypertension, sometimes called high blood pressure?
- ○ Yes *(go to B3)*
- ○ No *(go to C1)*
- ○ Only during pregnancy *(go to C1)*

B3. Have you been told two or more different times that you had hypertension or high blood pressure?
- ○ Yes
- ○ No
- ○ Don't know

B4. Has any medicine ever been prescribed by a doctor for your hypertension or high blood pressure?
- ○ Yes
- ○ No *(go to C1)*
- ○ Don't know *(go to C1)*

B5. Are you now taking any medicine prescribed by a doctor for your hypertension or high blood pressure?
- ○ Yes
- ○ No *(go to C1)*
- ○ Don't know *(go to C1)*

B6. How regularly do you take your high blood pressure medicine?
- ○ Always
- ○ Most of the time
- ○ About half the time
- ○ Less than half the time
- ○ Never

C1. Blood cholesterol is a fatty substance found in blood. Have you ever had your blood cholesterol checked?
- ○ Yes *(go to C2)*
- ○ No *(go to C4)*
- ○ Don't know *(go to C4)*

C2. About how long has it been since you last had your blood cholesterol checked?
- ○ Less than 1 year ago
- ○ 1-2 years ago
- ○ 3-4 years ago
- ○ 5 years ago
- ○ More than 5 years ago
- ○ Don't know

C3. Have you ever been told by a doctor or other health professional that your blood cholesterol is high?
- ○ Yes
- ○ No
- ○ Don't know

C4. About how long has it been since you had a rectal exam?
- ○ Less than 1 year ago
- ○ 1 year ago
- ○ 2 years ago
- ○ 3 or more years ago
- ○ Never
- ○ Don't know

C5. During the past ten years, have you had a tetanus shot?
- ○ Yes
- ○ No
- ○ Don't know

D1. In an average week, how many times do you engage in physical activity (exercise or work which lasts at least 20 minutes without stopping and which is hard enough to make you breathe heavier and your heart beat faster)?
- ○ Less than 1 time per week
- ○ 1-2 times per week
- ○ At least 3 times per week

Health Enrollment Assessment Review Questionnaire

D2. How much hard physical work is required on your job? Would you say...

- ○ A great deal
- ○ A moderate amount
- ○ A little
- ○ None
- ○ Not currently working

D3. How much hard physical work is required in your main daily activity (household or other non-job activities)? Would you say...

- ○ A great deal
- ○ A moderate amount
- ○ A little
- ○ None

E. Women's Health (men go to F1)

E1. About how long has it been since you had a breast examination by a doctor or other health professional?

- ○ Less than 1 year ago
- ○ 1 year ago
- ○ 2 years ago
- ○ 3 or more years ago
- ○ Never
- ○ Don't know

E2. A mammogram is an X-ray of each breast to look for breast cancer. Have you ever had a mammogram?

- ○ Yes
- ○ No *(go to E4)*
- ○ Don't know *(go to E4)*

E3. How long has it been since you had your last mammogram?

- ○ Less than 1 year ago
- ○ 1 year ago
- ○ 2 years ago
- ○ 3 or more years ago
- ○ Don't know

E4. A Pap smear is a test for cancer of the cervix. Have you ever had a Pap test (or Pap smear)?

- ○ Yes
- ○ No *(go to G1)*
- ○ Don't know *(go to G1)*

E5. How long has it been since you had your last Pap smear?

- ○ Less than 1 year ago
- ○ 1 year ago
- ○ 2 years ago
- ○ 3 or more years ago
- ○ Don't know

F. Men's Health (women go to G1)

F1. How long has it been since you had a testicular examination by a doctor or other health care professional?

- ○ Less than 1 year ago
- ○ 1 year ago
- ○ 2 years ago
- ○ 3 or more years ago
- ○ Never
- ○ Don't know

G1. Have you smoked at least 100 cigarettes in your entire life? (Note: 1 pack = 20 cigarettes)

- ○ Yes
- ○ No *(go to H1)*

G2. Do you *NOW* smoke cigarettes every day, some days, or not at all?

- ○ Every day
- ○ Some days
- ○ Not at all *(go to H1)*

G3. On the average, about how many cigarettes a day do you now smoke?

- ○ Less than 1 per day
- ○ 1-10 per day
- ○ 11-20 per day
- ○ 21-40 per day
- ○ 41 or more per day
- ○ Don't know

G4. Are you seriously intending to quit smoking in the next 6 months?

- ○ Yes
- ○ No

G5. Are you planning to quit smoking in the *next month?*

- ○ Yes
- ○ No

G6. Have you tried to quit smoking in the *past 12 months?*

- ○ Yes
- ○ No

H1. During the *past month*, have you had at least one drink of any alcoholic beverage such as beer, wine, wine cooler, or liquor?

- ○ Yes
- ○ No *(go to I1)*
- ○ Don't know

H2. In the past two weeks, on how many days did you drink any alcoholic beverages, such as beer, wine, or liquor?

- ○ None *(go to H4)*
- ○ 1-2 days
- ○ 3-4 days
- ○ 5-6 days
- ○ 7 or more days
- ○ Don't know

H3. A drink is 1 can or bottle of beer, 1 glass of wine, 1 can or bottle of wine cooler, 1 cocktail, or 1 shot of liquor. During the past 2 weeks, on the days when you drank, how many drinks did you drink on average?

- ○ 1-2 drinks
- ○ 3-4 drinks
- ○ 5-6 drinks
- ○ 7 or more drinks
- ○ Don't know

H4. During the past month, how many times have you driven when you've had perhaps too much to drink?

- ○ None
- ○ 1-2 times
- ○ 3-4 times
- ○ 5-6 times
- ○ 7 or more times
- ○ Don't drive
- ○ Don't know

PLEASE DO NOT WRITE IN THIS AREA

52366

Health Enrollment Assessment Review Questionnaire

H5. During the past month, have you thought you should cut down on your drinking of alcohol?
○ Yes ○ No

H6. During the past month, has anyone complained about your drinking?
○ Yes ○ No

H7. During the past month, have you felt guilty or upset about your drinking?
○ Yes ○ No

H8. During the past month, was there at least one day on which you had five or more drinks of beer, wine, or liquor?
○ Yes ○ No

I1. How often do you feel that your present work or lifestyle is putting you under too much stress?
○ Often ○ Sometimes ○ Seldom ○ Never

I2. During the *past 2 weeks*, would you say that you experienced...
○ A lot of stress ○ Relatively little stress
○ A moderate amount of stress ○ Almost no stress at all

I3. In the past year, how much effect has stress had on your health?
○ A lot ○ Some ○ Hardly any or none

J1. In general, how satisfied are you with your life (e.g., work situation, social activity, accomplishing what you set out to do)?
○ Not satisfied ○ Mostly satisfied
○ Somewhat satisfied ○ Totally satisfied

J2. How often do you have any serious problems dealing with your husband or wife, parents, friends, or with your children?
○ Often ○ Sometimes ○ Seldom ○ Never

J3. During the past year, have you been separated from your family for a block of at least 30 days?
○ Yes ○ No

In the past month, have you often been bothered by...

K1. ...little interest or pleasure in doing things?
○ Yes ○ No

K2. ...feeling down, depressed, or hopeless?
○ Yes ○ No

K3. ..."nerves" or feeling anxious or on edge?
○ Yes ○ No

K4. ...worrying about a lot of different things?
○ Yes ○ No

K5. During the past month, have you had an anxiety attack (suddenly feeling fear or panic)?
○ Yes ○ No

K6. During the past 12 months, have you seen a mental health professional?
○ Yes ○ No ○ Don't know

L1. During the past two weeks, how many days did you stay in bed for more than half of the day because of illness or injury?
○ None ○ 3-4 days ○ 7 or more days
○ 1-2 days ○ 5-6 days ○ Don't know

L2. During the past two weeks, how many days did you miss more than half of the day from your job or business because of illness or injury?
○ None ○ 3-4 days ○ 7 or more days
○ 1-2 days ○ 5-6 days ○ Don't know

L3. Do you have difficulty walking such as hobbling, shuffling, or not being able to walk a straight line?
○ Yes ○ No

M1. How many different prescription medications are you currently taking?
○ None ○ 6 or more medications
○ 1-2 medications ○ Don't know
○ 3-5 medications

M2. & **M3.** Excluding visits for pregnancy, medication refills, and dental care, how many times did you see a doctor, nurse, or other health care professional for an office visit or clinic appointment? (Include both civilian and military health care professionals. Only include visits for yourself.)

during the PAST MONTH	*during the 12 PAST MONTHS*
○ None	○ None
○ 1-2 visits	○ 1-5 visits
○ 3-4 visits	○ 6-10 visits
○ 5-6 visits	○ 11-15 visits
○ 7 or more visits	○ 16-20 visits
○ Don't know	○ 21 or more visits
	○ Don't know

Health Enrollment Assessment Review Questionnaire

M4. During the past 12 months, how many times have you gone to an emergency room or urgent care clinic?

- ○ None
- ○ 1-2 visits
- ○ 3-4 visits
- ○ 5-6 visits
- ○ 7 or more visits
- ○ Don't know

M5. During the past 12 months, have you spent one or more nights in the hospital? (Do not include hospitalization for deliveries.)

- ○ Yes
- ○ No (go to N1)

M6. During the past 12 months, how many nights have you spent in the hospital?

- ○ 1-2 nights
- ○ 3-4 nights
- ○ 5-6 nights
- ○ 7 or more nights
- ○ Don't know

M7. During the past 12 months, on how many different occasions have you entered the hospital and stayed for at least one night?

- ○ 1 time
- ○ 2-3 times
- ○ 4 or more times
- ○ Don't know

Have you ever been told by a health care provider that you have...

N1. ...diabetes or sugar diabetes?	○ Yes	○ No	○ Don't know
N2. ...had a stroke?	○ Yes	○ No	○ Don't know
N3. ...had a heart attack?	○ Yes	○ No	○ Don't know
N4. ...emphysema/chronic bronchitis?	○ Yes	○ No	○ Don't know
N5. ...arthritis?	○ Yes	○ No	○ Don't know
N6. ...Parkinson's disease or other neurologic disease?	○ Yes	○ No	○ Don't know
N7. ...depression?	○ Yes	○ No	○ Don't know
N8. ...HIV or AIDS?	○ Yes	○ No	○ Don't know
N9. ...anxiety or personality disorder?	○ Yes	○ No	○ Don't know
N10. ...cancer?	○ Yes	○ No	○ Don't know
N11. ...heart disease or angina?	○ Yes	○ No	○ Don't know
N12. ...liver disease?	○ Yes	○ No	○ Don't know
N13. ...kidney disease?	○ Yes	○ No	○ Don't know
N14. ...a stomach ulcer?	○ Yes	○ No	○ Don't know
N15. ...asthma?	○ Yes	○ No	○ Don't know

N16. During the past 12 months, have you seen a health care provider on 2 or more occasions for a bone, joint, back, or muscle problem

- ○ Yes
- ○ No

N17. Do you have a dependent family member less than 18 years old with a serious medical condition?

- ○ Yes
- ○ No

N18. Do you have a close family member (parent, brother/sister, or child) who has or had angina, a heart attack, or other heart disease?

- ○ Yes
- ○ No
- ○ Don't know

PLEASE DO NOT WRITE IN THIS AREA

52366

APPENDIX E

VARIABLES CONSTRUCTED FOR CONSIDERATION IN THE DERIVED HRU MODEL

Table E-1

Data Dictionary for Variables Constructed for the Derived HRU Model

Variable & survey item	Variable levels & coding	Description	Original HRU
Demographic Information			
age	continuous	age as of March 1, 1998	
gender	0 male		1
A.2	1 female		
marital status	1 never married		2
A.3	2 married		
	3 separated		
	4 divorced		
race/ethnicity	1 American Indian or Alaska native		
A.4	2 Asian/Oriental		
	3 black, Hispanic		
	4 black non-Hispanic		
	5 Pacific islander		
	6 white, Hispanic		
	7 white, non-Hispanic		
	8 other		
General Health			
BMI	continuous	body mass index	
overweight	0 no	by body mass index adjusted by age	
A.6, 7	1 yes		
health status	0 excellent		3
A.8	1 very good		
	2 good		
	3 fair		
	4 poor		
cholesterol	0 normal		
C.3	1 high		
activity level	0 high	exercise 3 times/week or a great deal of hard physical work	
D.1, 2, 3	1 moderate	exercise 1-2 times/week or moderate hard physical work	
	2 low	exercise less than 1 time/week or little or no hard physical work	

Table E-1 "Continued"

Variable & survey item	Variable levels & coding	Description	Original HRU
smoking	0 <1/day	cigarettes per day	5
G.3	1 1-10/day		
	2 11-20/day		
	3 21-40/day		
	4 41+/day		
days in bed	0 none		10
L.1	1 1-2 days		
	2 3-4 days		
	3 5-6 days		
	4 7+days		
days from work	0 none		10
L.2	1 1-2 days		
	2 3-4 days		
	3 5-6 days		
	4 7+days		
walking	0 normal		
L.3	1 difficulty		
Emotional Health			
days drinking	0 none	past 2 weeks	
H.2	1 1-2 days		
	2 3-4 days		
	3 5-6 days		
	4 7+ days		
drinks per day	0 none	past 2 weeks	
H.3	1 1-2 drinks		
	2 3-4 drinks		
	3 5-6 drinks		
	4 7+ drinks		
drinking driving	0 none		
H.4	1 1-2 times		
	2 3-4 times		
	3 5-6 times		
	4 7+ times		
cut down	0 no		6
H.5	1 yes		
complaints	0 no		6
H.6	1 yes		
guilt	0 no		6
H.7	1 yes		

Table E-1 "Continued"

Variable & survey item	Variable levels & coding	Description	Original HRU
5 drinks in a day H.8	0 no 1 yes		6
stress frequency I.1	0 never 1 seldom 2 sometimes 3 often		7
stress amount I.2	0 none 1 little 2 moderate 3 high		7
stress effect I.3	0 no effect 1 some effect 2 much effect		7
satisfaction J.1	0 totally satisfied 1 mostly satisfied 2 somewhat satisfied not satisfied		
family problems J.2	0 never 1 seldom 2 sometimes 3 often		8
separation J.3	0 < 30 days 1 30+ days	consecutive separation in past year	
disinterested K.1	0 absent 1 present		9
depressed K.2	0 absent 1 present		9
anxious K.3	0 absent 1 present		9
worrying K.4	0 absent 1 present		9
anxiety attack K.5	0 absent 1 present		9
Health Services			
mental health treatment K.6	0 none 1 treated	seen by a mental health professional in past 12 months	9

Table E-1 "Continued"

Variable & survey item	Variable levels & coding	Description	Original HRU
prescriptions	0 none		11
M.1	1 1-2 drugs		
	2 3-5 drugs		
	3 6+ drugs		
month visits	0 visits	office visits in past month	
M.2	1 1-2 visits		
	2 3-4 visits		
	3 5-6 visits		
	4 7+ visits		
year visits	0 visits	office visits in past year	12
M.3	1 1-2 visits		
	2 3-4 visits		
	3 5-6 visits		
	4 7+ visits		
emergency visits	0 visits	emergency room or urgent care clinic in past year	13
M.4	1 1-2 visits		
	2 3-4 visits		
	3 5-6 visits		
	4 7+ visits		
hospital stay	0 nights		14
M.6	1 1-2 nights		
	2 3-4 nights		
	3 5-6 nights		
	4 7+ nights		
musculoskeletal visit	0 0-1 visits		
N.16	1 2+ visits		
Disease Conditions			
hypertension	0 normal BP		4
B.2, 3, 4, 5	1 hypertension	told twice by professional	
	2 medicated	on hypertension medications	
diabetes	0 absent		
N.1	1 present		
stroke	0 absent		
N.2	1 present		
heart disease	0 absent		15
N.3, 11	1 present	heart attack, heart disease, or angina	

Table E-1 "Continued"

Variable & survey item	Variable levels & coding	Description	Original HRU
emphysema			16
bronchitis	0 absent		
N.4	1 present		
arthritis	0 absent		17
N.5	1 present		
neurological disease	0 absent		
N.6	1 present		
depression	0 absent		
N.7	1 present		
HIV or AIDS	0 absent		
N.8	1 present		
anxiety or personality disorder	0 absent	deleted, anxiety disorder and personality disorder are separate classes of mental disorders	
N.9	1 present		
cancer	0 absent		
N.10	1 present		
liver disease	0 absent		
N.12	1 present		
kidney disease	0 absent		
N.13	1 present		
ulcer	0 absent		
N.14	1 present		
asthma	0 absent		
N.15	1 present		
Family Health			
sick dependent	0 yes		
N.17	1 no		
family history heart disease	0 yes		
N.18	1 no		

Note. Variable & survey item = variable name followed by the number of the item(s) on the HEAR survey that was used to construct the variable; variable levels and coding = all the levels that were coded for the variable and the numeric code value assigned to each level; Description = comments about the methodology used to code the particular variable; Original HRU = reference number to comparable item in Table III-3 from the original HRU model.

Table E-2

Frequencies of Variables Constructed for the Derived HRU Model

Variable & survey item	Variable levels & coding	Frequency	%	Cum. %
Demographic Information				
age	continuous	NA	NA	NA
gender	1 female	369	94.4	94.4
A.2	0 male	22	5.6	100.0
marital status	2 married	366	93.6	93.6
A.3	1 never married	14	3.6	97.2
	3 separated	9	2.3	99.5
	4 divorced	2	0.5	100.0
race/ethnicity	7 white, non-Hispanic	273	69.8	69.8
A.4	4 black non-Hispanic	53	13.6	83.4
	2 Asian/Oriental	21	5.4	88.8
	8 other	15	3.8	92.6
	6 white, Hispanic	10	2.6	95.2
	5 Pacific islander	9	2.3	97.5
	3 black, Hispanic	8	2.0	99.5
	1 American Indian or Alaska native	2	0.5	100.0
General Health				
BMI	continuous	NA	NA	NA
overweight *	0 no	231	59.1	59.1
A.6, 7	1 yes	160	40.9	100.0
health status *	1 very good	175	44.8	44.8
A.8	2 good	110	28.1	72.9
	0 excellent	80	20.5	93.4
	3 fair	26	6.6	100.0
	4 poor	0	0.0	100.0
cholesterol *	0 normal	329	84.1	84.1
C.3	1 high	62	15.9	100.0
activity level	0 high	190	48.6	48.6
D.1, 2, 3				
	1 moderate	145	37.1	85.7
	2 low	56	14.3	100.0
smoking	0 <1/day	319	81.6	81.6
G.3	1 1-10/day	31	7.9	89.5
	2 11-20/day	25	6.4	95.9
	3 21-40/day	15	3.8	99.7
	4 41+/day	1	0.3	100.0

Table E-2 "Continued"

Variable & survey item	Variable levels & coding	Frequency	%	Cum. %
days in bed *	0 none	359	91.8	91.8
L.1	1 1-2 days	23	5.9	97.7
	2 3-4 days	6	1.5	99.2
	3 5-6 days	1	0.3	99.5
	4 7+days	2	0.5	100.0
days from work *	0 none	359	91.8	91.8
L.2	1 1-2 days	22	5.6	97.4
	2 3-4 days	5	1.3	98.7
	3 5-6 days	1	0.3	99.0
	4 7+days	4	1.0	100.0
walking *	0 normal	376	96.2	96.2
L.3	1 difficulty	15	3.8	100.0
Emotional Health				
days drinking	0 none	234	59.8	59.8
H.2	1 1-2 days	112	28.6	88.4
	2 3-4 days	26	6.6	95.0
	4 7+ days	10	2.6	97.6
	3 5-6 days	9	2.3	100.0
drinks per day	0 none	232	59.3	59.3
H.3	1 1-2 drinks	122	31.2	90.5
	2 3-4 drinks	24	6.1	96.6
	3 5-6 drinks	7	1.8	98.4
	4 7+ drinks	6	1.6	100.0
drinking driving	0 none	385	98.5	98.5
H.4	1 1-2 times	6	1.5	100.0
	2 3-4 times	0	0.0	100.0
	3 5-6 times	0	0.0	100.0
	4 7+ times	0	0.0	100.0
cut down	0 no	371	94.9	94.9
H.5	1 yes	20	5.1	100.0
complaints	0 no	385	98.5	98.5
H.6	1 yes	6	1.5	100.0
guilt	0 no	384	98.2	98.2
H.7	1 yes	7	1.8	100.0

Table E-2 "Continued"

Variable & survey item	Variable levels & coding	Frequency	%	Cum. %
5 drinks in a day	0 no	363	92.8	92.8
H.8	1 yes	28	7.2	100.0
stress frequency *	2 sometimes	155	39.6	39.6
I.1	1 seldom	114	29.2	68.8
	3 often	71	18.2	87.0
	0 never	51	13.0	100.0
stress amount *	2 moderate	176	45.0	45.0
I.2	1 little	100	25.6	70.6
	3 high	68	17.4	88.0
	0 none	47	12.0	100.0
stress effect *	0 no effect	168	43.0	43.0
I.3	1 some effect	165	42.2	85.2
	2 much effect	58	14.8	100.0
satisfaction *	1 mostly satisfied	200	51.2	51.2
J.1	2 somewhat satisfied	107	27.4	78.6
	0 totally satisfied	61	15.5	94.1
	3 not satisfied	23	5.9	100.0
family problems *	1 seldom	189	48.3	48.3
* J.2	2 sometimes	118	30.2	78.5
	0 never	64	16.4	94.9
	3 often	20	5.1	100.0
separation	0 < 30 days	271	69.3	69.3
J.3	1 30+ days	120	30.7	100.0
disinterested *	0 absent	307	78.5	78.5
* K.1	1 present	84	21.5	100.0
depressed *	0 absent	288	73.7	73.7
* K.2	1 present	103	26.3	100.0
anxious *	0 absent	250	63.9	63.9
* K.3	1 present	141	36.1	100.0
worrying *	1 present	225	57.5	57.5
* K.4	0 absent	166	42.5	100.0
anxiety attack	0 absent	364	93.1	93.1
K.5	1 present	27	6.9	100.0

Table E-2 "Continued"

Variable & survey item	Variable levels & coding	Frequency	%	Cum. %
Health Services				
mental health		345	88.2	88.2
treatment	0 none			
K.6	1 treated	46	11.8	100.0
prescriptions	0 none	203	51.9	51.9
* M.1	1 1-2 drugs	153	39.1	91.0
	2 3-5 drugs	28	7.2	98.2
	3 6+ drugs	7	1.8	100.0
month visits	0 visits	226	57.8	57.8
* M.2	1 1-2 visits	127	32.5	90.3
	2 3-4 visits	26	6.6	96.9
	4 7+ visits	7	1.8	98.7
	3 5-6 visits	5	1.3	100.0
emergency visits	0 visits	289	73.9	73.9
M.4	1 1-2 visits	88	22.5	96.4
	2 3-4 visits	10	2.6	99.0
	3 5-6 visits	3	0.8	99.8
	4 7+ visits	1	0.2	100.0
hospital stay	0 nights	353	90.3	90.3
M.6	1 1-2 nights	22	5.6	95.9
	2 3-4 nights	5	1.3	97.2
	3 5-6 nights	3	0.8	98.0
	4 7+ nights	8	2.0	100.0
musculoskeletal visit	0 0-1 visits	304	77.7	77.7
N.16	1 2+ visits	87	22.3	100.0
Disease Conditions				
hypertension	0 normal BP	348	89.0	89.0
B.2, 3, 4, 5	2 medicated	31	7.9	96.9
	1 hypertension	12	3.1	100.0
diabetes	0 absent	369	94.4	94.4
N.1	1 present	22	5.6	100.0
stroke	0 absent	390	99.7	99.7
N.2	1 present	1	0.3	100.0
heart disease	0 absent	387	99.0	99.0
N.3, 11	1 present	4	1.0	100.0
emphysema		373	95.4	95.4
bronchitis	0 absent			
N.4	1 present	18	4.6	100.0

Table E-2 "Continued"

Variable & survey item	Variable levels & coding	Frequency	%	Cum. %
arthritis	0 absent	348	89.0	89.0
N.5	1 present	43	11.0	100.0
neurological disease	0 absent	387	99.0	99.0
N.6	1 present	4	1.0	100.0
depression	0 absent	350	89. 5	89. 5
N.7	1 present	41	10.5	100.0
HIV or AIDS	0 absent	391	100.0	100.0
N.8	1 present	0	0.0	100.0
cancer	0 absent	382	97.7	97.7
N.10	1 present	9	2.3	100.0
liver disease	0 absent	388	99.2	99.2
N.12	1 present	3	0.8	100.0
kidney disease	0 absent	388	99.2	99.2
N.13	1 present	3	0.8	100.0
ulcer	0 absent	375	95.9	95.9
N.14	1 present	16	4.1	100.0
asthma	0 absent	342	87.5	87.5
N.15	1 present	49	12.5	100.0
Family Health				
sick dependent	0 yes	366	93.6	93.6
N.17	1 no	25	6.4	100.0
family history		268	68.5	68.5
heart disease	0 yes			
N.18	1 no	123	31.5	100.0

Note. * = used in the final Derived HRU Model

APPENDIX F

CURRICULUM VITA

Jody W. Donehoo

Current Experience

As Health Program and Policy Specialist of the TRICARE Mid-Atlantic Region, represented the region and provided leadership in all aspects of the procurement and program management of the Managed Care Support Contract awarded for $3 billion dollars. Continue to serve as the chief technical advisor to the Administrative Contracting Officer for TRICARE Program issues. Represent the region as a member on the Integrated Product Team established by the TRICARE Management Activity, a field office of the Office of the Assistant Secretary of Defense for Health Affairs to provide leadership on a re-bid of the contract.

Professional Employment and Military Service

Mar 93 - Present	Health Program & Policy Specialist. Department of Defense, TRICARE Mid-Atlantic Lead Agent Office, Norfolk, VA.
Sep 87 - Present	From Lieutenant to Commander rank. Selected (drilling) Reservist. Naval Reserve.
Jul 91 - Feb 93	Director, Psychiatric Care Center. Obici Hospital, Suffolk, VA.
Jun 89 - Jul 91	Mental Health Clinician, Comprehensive Mental Health Services. City of Virginia Beach.
Sep 87 - May 89	Flex-pool Staff Nurse. Tidewater Psychiatric Institute, Norfolk, VA. Full-time graduate student concurrently.
Sep 88 - Apr 89	Masters Student Counseling Intern. Center Psychiatrists, Chesapeake, VA.
Feb 86 - Aug 87	Lieutenant rank. Staff Nurse, Relief Division Officer. Naval Medical Clinic, Quantico, VA.
Oct 81 - Feb 86	From Ensign to Lieutenant rank. Staff Nurse, Relief Division Officer. Naval Medical Center, Portsmouth, VA.

Research

1989	Masters Thesis – *Religious Life Profile of Christians with a Borderline Personality Disorder*

Education

1987	M.A. Counseling, Summa Cum Laude, Regent University. Dean's Scholarship. Leadership award.
1981	B.S.N., Oral Roberts University, Deans List. Attended Georgia Institute of Technology.

Professional Licensure, & Certification

- Registered Nurse since 1981. Virginia licensure.
- Clinical Nurse Specialist since 1992. Virginia licensure.
- Certified Specialist in Adult Psychiatric and Mental Health Nursing (R.N.,C.S.) since1991. American Nurse's Credentialing Center.