To the Graduate Council:

I am submitting herewith a dissertation written by Umagasen Chinnasamy Naidoo entitled "The Impact of Physician Profiling in Modifying Physician Utilization Practice Patterns." I have examined the final copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Human Ecology.

Robert J. Pursley, Major Professor

We have read this dissertation
And recommend its acceptance:

Accepted for the Council:

Associate Vice Chancellor and
Dean of The Graduate School

THE IMPACT OF PHYSICIAN PROFILING IN MODIFYING PHYSICIAN UTILIZATION PRACTICE PATTERNS

A Dissertation
Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Umagasen Chinnasamy Naidoo
August 2000

DEDICATION

This dissertation is dedicated to my wife and son

Tricia Lynn

And

Devin Allen

for the sacrifices they made and the support they provided through all my educational endeavors.

ACKNOWLEDGEMENTS

Sir Isaac Newton once said that if he sees farther than others "it is by standing on the shoulders of giants". This research would not have been possible without the contributions of the many teachers and educators who preceded me and who imparted to me their knowledge. I am, as is society, indebted to their gift, the gift of knowledge and the pursuit of truth. I am profoundly grateful to my dissertation committee for their expertise and enthusiasm that served as a beacon of light.

To Dr. Robert J. Pursley whose global vision kept me on track to answer the appropriate questions and whose subtle humor sustained me in this journey; to Dr. Debra C. Wallace whom I admire for her practical and yet thought provoking approach to respected scientific research; to Dr. Robert H. Kirk whose keen sense of detail and methodological rigor made this research more succinct and informative; to Dr. Bill C. Wallace for having faith in my abilities and whose encouragement propelled me to strive for more.

To The Plan, especially the Medical Information Management team, for giving me the opportunity to evaluate the physician profiling process and access to important utilization data that were beneficial to scientific investigation and the pursuit of vastly important process improvements.

Finally, to my family and friends who have offered their support. To my parents and grandparents for teaching me invaluable lessons of life, for their love, and for their patience. To my in-laws who always offered their support and encouragement in a caring and loving manner, and whose presence and humor always brightened my day.

ABSTRACT

Even though the results of before-and-after studies on physician profiling vary, profiles continue to be used widely as a vehicle for information feedback to change physician utilization practice patterns. The purpose of this study was to assess the impact of physician profiling and information sharing in modifying practice patterns among HMO primary care physicians within a managed care plan in the state of Tennessee.

All commercial HMO primary care physicians from the plan were included in the study, as part of a pre-intervention/post-intervention quasi-experimental design. The experimental group (physicians receiving profiles) consisted of 463 physicians, whereas the control group (physicians not receiving profiles) consisted of 1,213 physicians. The pre-intervention period was calendar year 1998 and the post-intervention period was calendar year 1999. In addition, rural and urban physicians were analyzed within each of the groups.

Non-parametric (Wilcoxon Rank Sum, Wilcoxon Signed Rank, and Chi-square) test-statistics were used to test for differences between and within groups across 7 utilization measures (primary care visits PMPY, specialist visits PMPY, number of prescriptions PMPY, proportion of generic prescriptions written, inpatient admissions/1000 members PY, emergency room visits/1000 members PY, outpatient surgeries/1000 members PY) and 3 quality measures (cholesterol screening rate, mammography screening rate, and Hemoglobin A1c screening rate). All utilization measures were case-mix adjusted using the Johns Hopkins Adjusted Clinical Groups (ACGs).

Results showed significant improvement ($p < 0.05$) in 5 of the 7 utilization measures, from pre-intervention to post-intervention, in the experimental group. Only 3 of the 7 utilization measures showed significant improvement, from pre-intervention to post-intervention, in the control group. The rural group showed significantly lower ($p < 0.05$) utilization than the urban group. However, urban physicians were more responsive to physician profiling than rural physicians. No significant improvement was seen in the quality measures from pre-intervention to post-intervention across all groups.

The research concludes that physician profiling is effective in changing practice patterns for the utilization measures, but not for the quality measures. Additional interventions are needed to improve the quality screening measures. Potential under-utilization among the rural physicians needs to be addressed.

TABLE OF CONTENTS

LIST OF TABLES

TABLE NUMBER **PAGE**

LIST OF FIGURES

FIGURE NUMBER **PAGE**

CHAPTER I

FORMULATION AND DEFINITION OF THE PROBLEM

Introduction

Health care costs have increased from 9.1% of the gross national product in 1980 to almost 14% in 1999, far more than any other country in the world (Schwartz, et al., 1999). Many solutions have been offered for as long as costs have been escalating. It is widely believed that success in the containment of health care costs will result from team efforts in which many health care professionals contribute to the delivery of higher quality, more efficient care. Any anticipated success requires the active participation on the part of physicians, as they receive approximately 20% of health care dollars but make decisions that create 80% of medical care costs (Evans & Lindsay 1996).

U.S. physicians are becoming increasingly aware of the need to improve utilization of the services they deliver and recommend, thereby participating in the health care cost improvement effort. Two types of review could be used to improve utilization: assessing the appropriateness of specific decisions, and assessing the broad pattern of decisions made over time. Because the case-by-case approach is burdensome on physicians and does not easily accommodate valid exceptions to practice guidelines, physicians have shown an interest in devising utilization improvement mechanisms that are based on their average behavior over time (Spoeri & Ullman, 1997).

The idea of concentrating on patterns of care instead of specific clinical decisions has been termed "physician profiling". The practice pattern of a single physician or group of physicians is expressed as a rate: this rate is a measure of resource consumption during a specific period for the population being served. The resulting profile is then compared

with a norm that is either based on practice (e.g. profiles of other physicians) or based on standards (e.g. practice guidelines).

The purpose of this chapter was to outline the need and purpose for the study of physician profiling, and to state the assumptions, delimitations, limitations, and definitions associated with the study.

Need for the Study

Presenting peer-comparison information (profiles) feedback to physicians is designed to stimulate consensus and allow physicians to make their own decisions. More than half of the physicians in the United States are subjects of either clinical or economic profiling (Emmons and Wozniak, 1994). Even though the results of before-and-after studies on profiling vary, profiles continue to be used widely as a vehicle for information feedback. Profiling reports, in most cases, compare individual physicians directly with their peers. In their meta-analysis of randomized clinical trials, Balas, et al, (1996) have discovered that while some randomized clinical trials of information feedback have been successful in changing practice patterns, several other trials indicated inconclusive or nonsignificant results (Balas, et al., 1996). This indicated a need for further research to determine the effectiveness of physician profiling.

In enacting Medicare physician-payment reform in 1989, both the American College of Physicians and Congress have recommended physician profiling. In the early 1990s, the Physician Payment Review Commission found that most profiling studies were limited to a single facility and to the use of a specific treatment or service (generally pharmaceutical agents or laboratory tests) instead of covering all the different services

associated with a clinical encounter. However, the problem of high costs is closely associated with the provision of multiple services that may or may not be related. Insurers are carrying out much of the profiling that incorporates measures on these multiple services, with the information being proprietary. Insurer's profiling methods are hence not being peer-reviewed, because the details are not publicly available. It is thus difficult for physicians to feel comfortable with the methodology, and for them to learn how comprehensive profiles may actually be implemented in practice (Spoeri & Ullman, 1997).

A new and controversial aspect of physician profiling is whether the results can be shared publicly, i.e. to employer clients and members. To help members choose a physician, some plans on the West Coast have started publicly releasing physician-specific, performance-based information. The release of this information is also intended to create an incentive for physicians to improve in the reported aspects of care. If these are to be trends for the millennium, there is certainly a need for physicians in every health plan across the country to become familiar with profiling measurements and methodologies for their individual practices (Evans, et al., 1995).

Most existing intervention studies on physician profiling have been performed in the inpatient service (training programs and teaching hospitals) and with medical interns and residents as subjects (Martin, Tul, & Sox, 1985; Berwick & Coltin, 1986). The results of such studies are difficult to apply to ambulatory care, nonteaching settings, and to experienced practitioners as opposed to trainees. In addition, measurements of effectiveness in most existing intervention studies include only rates of utilization of specific tests involved, but do not look at variation among individual physicians as an

outcome (Berwick & Coltin, 1986). Thus, there is a need to investigate profiles that measure changes in the absolute rates of utilization, and in variation among individuals in the same type of group practice.

Most prior studies of physician profiling tend to have "serious methodologic limitations that greatly restrict the strength of their conclusions" (Epstein, 1991). There have also been few investigations analyzing differences in the effectiveness of profiling under different circumstances. Epstein (1991) suggests the need to "identify factors that determine the effectiveness of different interventions, and the complexity of achieving behavioral changes among physicians".

Physician profiling as an analytic tool is still in its early stages of development. Profiling efforts will need to move beyond the inpatient population. If profiles are to be studied in relation to patient outcomes, new methods to adjust for case mix will need to be developed. Profiling as an analytic method will be encouraged for the following reasons: there will be continued pressure to do something about escalating health care costs; there is a strong desire to avoid the strategy of micromanagement of physicians (Spoeri & Ullman, 1997).

Purpose of the Study

The purpose of the study was to assess the impact of physician profiling and information sharing in modifying practice patterns among HMO primary care physicians within a managed care plan (hereafter referred to as The Plan) in the state of Tennessee.

In order to address the purpose of the study, the following research questions were formulated:

1. How do physician utilization practice patterns compare between HMO primary care physicians who receive physician profiles and those who do not?
2. How do physician utilization practice patterns compare between rural and urban HMO primary care physicians who receive physician profiles?
3. How do physician utilization practice patterns compare between rural and urban HMO primary care physicians who do not receive physician profiles?

Assumptions

The basic assumptions made regarding the study were:

1. Health insurance claims data are a reliable source of information to estimate physicians' utilization of services, and to monitor practice patterns.
2. Physicians understand the use of percentiles, which is the statistic used to measure their utilization of services on the profile.
3. Physicians understand the use of case-mix adjustments, which were applied to their profiles to effectively adjust for the illness burden of the physician's caseload.

Delimitations

For purposes of this study the researcher applied the following delimitation:

The study was delimited to primary care physicians in the State of Tennessee who participated in The Plan's commercial HMO, and who were responsible for managing the care of at least 50 commercial HMO members during the study period of the profile.

Limitations

For the purposes of the study, the following limitations were noted:

1. Each physician in the study may not have been exposed to the same set of members for the entire study period, since members in the plan are allowed to change their primary care physician at any time with proper notification to the plan.
2. Some physicians practice under multiple specialties. Only physicians who had one of the following specialties listed as their primary specialty were included in the study: pediatrics, internal medicine, general practice, and family practice.
3. Unlike the typical HMO that uses capitation to reimburse its primary care physicians, this Plan's HMO paid most primary care physicians on a fee for service basis. This limited the potential for under-utilization.

Definitions

Specific terms defined for the purposes of this study:

Physician Profiling – physician-specific reports with patient-level and procedure-specific detail, outlining precisely how physicians vary from their peers in the way they provide care based on specialty, practice, region, or other relevant metrics.

Primary Care Physician – In an episode of illness, a primary care physician is in most cases the first person of contact that the patient has with the health care system. Internists, pediatricians, family physicians, and general practitioners are referred to as primary care physicians.

HMO – Health Maintenance Organization. HMOs are organized health care systems that are responsible for both the financing and the delivery of a broad range of

comprehensive health services to an enrolled population. An HMO can be viewed as a combination of a health insurer and a health care delivery system that is responsible for providing health care services to their covered members through affiliated providers, who are reimbursed under different methods.

Practice Pattern – manner in which a physician practices medicine through the use of health care services.

Screening Tests – those preventive services in which a test or standardized examination procedure is used to identify patients requiring special intervention.

Impact – the effect physician profiling has on changing physician practice patterns, by comparing specific utilization and quality measures between physicians who receive profiles and those who do not receive profiles, to determine increases or decreases in measures.

Information Sharing – the distribution of knowledge and facts in the form of data and reports to physicians.

Summary

In this chapter, the researcher addressed the framework, the purpose, and the need for the study. An outline of the research questions, the basic assumptions, delimitations, limitations, and definition of key terms were also addressed.

The purpose of this study was to assess the impact of physician profiling and information sharing in modifying practice patterns among HMO primary care physicians in the state of Tennessee. Collating quality utilization measures into comprehensive physician profiles has recently becoming a useful approach to assessing physicians'

effectiveness and efficiency in providing health care. In the next chapter it will be shown how the need for the investigation was supported by a review of literature and research that concluded that in spite of the frequently described potential and widespread use of information feedback, the results of before- and after- studies on physician profiling vary. The ensuing chapters will present the methods and procedures used in the study, a detailed analysis of the results, and finally a discussion of the findings, conclusions, and recommendations.

CHAPTER II

REVIEW OF RELATED LITERATURE AND RESEARCH

Introduction

In spite of uncertainty about how to define and measure quality in our health care system, activities relating to quality management and outcomes are on the rise. This increasing interest in quality stems from recognizing that one can only achieve value in health care by balancing cost and quality.

Changing the inappropriate utilization of clinical procedures is one way to control costs and improve the quality of health care. As practice standards evolve, there is a call for using some procedures less and others more often. Provider uncertainty and differing beliefs about the value or efficacy of procedures have been stated as the underlying causes of variation in utilization (Balas, et al., 1996). Information is becoming an increasingly recognized intervention for changing the utilization rates of clinical procedures (Avorn, et al., 1992).

Information feedback, i.e. reporting on past patient care activities to influence future clinical decisions, is a common and fairly inexpensive intervention (Balas, et al., 1996). A popular approach is presenting peer-comparison information profiles to physicians without suggesting any further interpretation or benchmark. This physician profile is intended to encourage consensus and enable physicians to make their own decisions (Keller, Chapin, & Soule, 1990). A 1994 American Medical Association survey indicated that more than half of the physicians in the United States are subjects of clinical or economic profiling (Emmons & Wozniak, 1994). The results of before-and-after

studies on profiling vary, in spite of the frequently described potential and widespread use of information feedback (Balas, et al., 1996).

The purpose of this chapter was thus to explore the literature as exhaustively as possible to assess the effectiveness of physician profiling as an informational tool in both improving practice patterns and forging successful partnerships with physicians. The literature and research reviewed in this chapter were selected based on the following: theoretical framework, literature and research related in content, literature and research related in methodology, and literature and research related in content and methodology.

Theoretical Framework

Despite uncertainty about how to define and measure quality in our health care system, activities relating to quality management and outcomes are on the rise. This increasing interest in quality stems from recognizing that one can only achieve value in health care by balancing cost and quality. For a managed-care organization (MCO), measuring and monitoring quality has even broader implications.

Retention of existing customers and acquisition of new customers is of paramount importance to the MCO. Health care purchasers have increasing interest in quality management and improvement. As a result, organizations that monitor, evaluate, and manage quality are in a better position to compete in an increasingly competitive marketplace.

A Brief History of Quality in Health Care

It was the Bell System that was the leader in the early modern history of quality assurance. In the early 1900s, the Western Electric Company (a Bell operating company) sought the assistance of a group of individuals from its inspection department to research and develop new approaches in providing quality telephone services across the country. It was this group – Walter Shewhart, Harold Dodge, George Edwards, and W. Edwards Deming – who were the early pioneers of quality assurance. In fact, it was this group that coined the term *quality assurance* (Evans & Lindsay, 1996).

Walter Shewhart brought in the era of *quality control.* Quality control goes beyond simply preventing defects in manufactured goods or errors in service operations by inspection – it is focused on eliminating the problems that cause defects. During World War II, the U.S. Military began using statistical sampling techniques to impose strict standards on its suppliers. Free training courses in statistical methods (developed by the Bell system) were offered to all its suppliers. Even though this had a minimal impact on the wartime production effort, it created quality specialists who began using and extending these tools within their organizations. Statistical quality control had thus become widely known and slowly adopted by many manufacturing industries.

In the early 1950s, after World War II, the shortage of civilian goods in the U.S. made production a priority, and the push to produce large quantities of goods led to a decline in quality. During this time, Joseph Juran and W. Edwards Deming introduced statistical quality control techniques to the Japanese, who integrated quality into their management practices and developed a culture of continuous quality improvement (CQI). By the 1970s, Japanese companies were penetrating western markets with their superior

quality goods. This made the U.S. consumers, industry, and government more aware of quality, with considerable changes and improvements occurring in the 1980s (Evans & Lindsay, 1996).

Until the 1980s most of the quality focus has been in the manufacturing arena. Less attention was given to service industries, with health care possibly receiving the least attention. The concepts of CQI were not initially accepted among health care administrators and providers of care – most were still thinking in terms of quality assurance rather than quality improvement. Rather than concentrating on defective processes, quality assurance concentrated on identifying poor providers. Instead of looking to their customers, providers initially looked to themselves to determine what should be improved. They also battled with measurement issues. In summary, the health care field was slow to commit time and resources to understanding CQI theory and tools. Most of the early efforts in the late 1980s went into organizing teams, even after some health care leaders began to look earnestly at CQI. Far less effort was put into measuring the achievement of teams in improving outcomes (Lloyd & Carey, 1995).

When accrediting bodies and purchasers of care began to urge providers to document quality, a new group of individuals called "health care quality consultants" were almost immediately formed. Most of these consultants came from the manufacturing industry and understood the principles of CQI theory. However, they were less familiar with the unique problems of health care. Also, most health care administrators had not been exposed to CQI theory in their training programs. Most physicians still do not receive CQI training in medical school. Thus, both physicians and

administrators often find it hard to use CQI tools to measure the success or failure of their efforts.

The concepts of quality measurement that have been used in industry are now growing in demand for applications in health care. Consultants who could easily present examples of quality measurement and quality improvement in the manufacturing industry have been less successful in applying CQI techniques to measure improvement in delivering babies, increasing immunization rates, and lowering mortality rates. There is a definite need among health care administrators and providers to apply CQI tools to measure the success of efforts to improve health care outcomes. Because practicing professionals understand health care better, it would be most appropriate for these individuals to familiarize themselves with CQI techniques and apply them in pertinent situations (Lloyd & Carey, 1995).

The Deming Cycle

The physician profiling process in this study will be based on a CQI model developed by Walter Shewhart, called the Deming Cycle. The model was initially called the Shewhart Cycle and was later renamed the Deming Cycle by the Japanese in 1950, after William Edward Deming's work in that country (Evans & Lindsay, 1996).

The Deming Cycle is made up of four stages: plan, do, study, and act. Most of the focus of the Deming Cycle is on implementation. The plan stage consists of studying the current situation, gathering data, and planning for improvement. Its activities include defining the process, its inputs, outputs, customers, and suppliers; understanding customer expectations; identifying problems; testing theories of causes; and developing

solutions. In the do stage, the plan is implemented on a trial basis, for example, in a laboratory, a pilot disease management program, or with a small group of customers. This small-scale implementation is an experiment to evaluate a proposed solution and provide objective data. The study stage determines whether the trial plan is working correctly and if any further problems or opportunities are found. Sometimes, a proposed solution needs to be redesigned or eliminated. New solutions are proposed and evaluated by returning to the do stage. In the act stage, the final plan is implemented and the improvements become standardized and practiced continuously. This process then leads back to the plan stage for further diagnosis and improvement. The cycle is never ending – it is focused on continuous improvement. The improved process is a springboard for further improvements. The Deming Cycle is based on the premise that improvement comes from the application of knowledge. When looking at the Deming Cycle, one needs to consider three basic questions:

1. What are we trying to accomplish?
2. What changes can we make that will result in improvement?
3. How will we know that a change is an improvement?

Through a process of learning, knowledge is developed (Evans & Lindsay, 1996).

Some managed care organizations (MCOs) have used the Deming Cycle to improve processes (both clinical and non-clinical) in a few situations: to design an antibiogram to improve physician's prescribing of antibiotic therapy (Gerald & Arnold, 1995); to improve the management of aggressive behavior in a behavioral health care facility (Joint Commission on Accreditation of Healthcare Organizations, 1996); to assess, implement, and evaluate an employee assistance program (Monfils, 1995). The

cycle may also be used to measure and improve the physician profiling process at The Plan. The *plan* and *do* phases of the cycle have already been completed for the physician profiling process. The *study* (measure whether the profiling intervention has been successful in improving practice patterns) and act (how to improve the process based on the results from the *study* phase) phases form a part of this research.

Physicians who receive profiles also have the opportunity to use the Deming Cycle to improve processes. For example, a physician could choose a measure on the profile that needs improvement (e.g., mammography screening rates), and start a cycle that would have as an outcome increased mammography screening rates. In doing this, the physician would scrutinize their process for recommending mammography screenings to the eligible patient population, thereby improving the process not only for patients in the plan that provided the information, but for all patients that visit their practice.

There is little research and literature that have shown how health care organizations have used CQI tools (in particular, the Deming Cycle) to improve processes. However, there is a growing need to apply CQI concepts and theories to improve health outcomes.

Review of Literature Related in Terms of Content

The purpose of this section was to report on research and literature that evaluates physician profiling efforts in various settings. If appropriately structured, physician profiling stimulates improvements in practice patterns, assuming that providers can offer input into the methodology and thus readily accept the results. However, many physicians

remain wary of profiling because they have seen the resulting reports used as sticks instead of carrots (Bernstein, 1998).

Joseph E. Biskupiak (1998) stated that instead of looking at performance, which is the most important benchmark, many health plans and provider groups mainly engage in economic profiling to determine which physicians provide care at the lowest cost. The problem with this approach was that the focus was on cost, not outcomes. He says, "from a patient's perspective, cost is not the issue. Patients don't care if their physician is an outlier. The objective of physician profiling should be to determine the most reasonable cost for the best possible outcome. Once you identify low-cost providers with good outcomes, you can study their practices, develop guidelines, and disseminate that information throughout your network to move other physicians in that direction" (Bernstein, 1998).

In his article on medical practice in the new millenium, Alan B. Bernstein (1998) suggests that health care consumers and payers are now demanding accountability for health outcomes. For example, there is a need for health plans to demonstrate efforts to raise mammography and immunization rates, to improve eye and foot care for diabetics, and to ensure that patients with congestive heart failure are on proper medication. Bernstein further states that outcome measures used to quantify the value of health care fall into four categories:

- Financial outcomes – these include appropriate utilization of services for specific patient cohorts and the cost of providing these services.
- Performance outcomes – these include internal measures of organizational performance and external evaluations based on patient, provider, and payer

satisfaction. The National Committee for Quality Assurance (NCQA) offers a comprehensive patient satisfaction survey that assesses various aspects of care.

- Clinical outcomes – these quantify the effectiveness and appropriateness of the type of care and treatment provided. For example, specific surgical procedures such as carotid endarterectomy and coronary artery bypass grafting (CABG) have been studied in large-scale trials in an attempt to determine predictive factors for positive outcomes.
- Perceived health outcomes – these serve to evaluate a patient's perception of his or her own health status and perceived quality of life. The SF-36 instrument developed by John Ware and associates is the prototype tool used to evaluate health outcomes in a patient population.

With the shift in emphasis to outcome-based quality, recent physician profiling models have started to examine individual practices that impact outcomes such as immunization rates, hemoglobin A_{1c} levels in diabetic patients, and mammography rates for women over age 50 (Metfessel, 1998).

One of the hallmarks of competent profiling is to effectively adjust for the illness burden of the provider's patient load. Physicians who are confident that they are practicing high-quality care are not likely to accept evidence to the contrary unless their populations' underlying illness burden is considered. Otherwise, these physicians are likely to excuse their performance with the assertion that "my patients are sicker, older, and different". A number of case-mix adjustment techniques are available that deal with the problem of differing illness burdens among physicians and allow comparisons. Methods include diagnosis clusters, ambulatory visit groups, ambulatory care groups, and

episodes of care. Currently, the most popular is the use of adjusted clinical groups (ACGs), which are a population-based case-mix adjuster that use ICD-9 codes, age, and gender for classification of patients. Episodes of care are a useful case-mix indicator, since care can be tracked over time and allows analysis of visit rates, procedures performed, presence of complications, and length of treatment, within the context of specific medical conditions (Metfessel, 1998; Seidman & Greenfield, 1998).

Peter R. Kongstvedt (1996) states that physician profiles are of no use unless the results are compared with some type of standard. The most common way of comparing profiling results is to provide data for each individual practice in comparison with one or more of the following:

- Plan average results – This standard is simply the average for the entire health plan, and is the crudest method of comparison.
- Specialty or peer group – This compares each practice only within its own specialty, e.g. pediatricians are compared with other pediatricians.
- Peer group adjusted for age, sex, and severity of illness – This is the most complicated approach, as described earlier, but provides the most meaningful comparisons.

In 1998, Donald E. Lighter, Director of Quality Management and an expert in physician profiling systems, with a large health insurer, suggested the following as important features for physician profiles:

- From the physician's perspective, a consistent report format will allow easier interpretation of the results. Numeric reports are rarely as effective as graphic output,

and so graphs and charts should be included as much as possible. Physicians also appreciate concise reports that convey a great deal of information for rapid analysis.

- A method for the physician to contact the plan for additional information should be included in the profile so that physicians have a resource for improving practice patterns through the use of more specific information.
- Once outlier physicians are discovered based on high-level profiling, drill-down (or analytic pathway) capability is essential to determine if these physicians are different because of noise – or random variation – or because their practice patterns substantially differ from the norm.
- Traditionally, profiles have provided static information about a physician's performance, usually by summarizing data quarterly or even less frequently. As techniques of Continuous Quality Improvement (CQI) and Statistical Process Control (SPC) are applied to health claims data for improving the cost/quality profile, different analytic approaches will be used to evaluate patterns and cost of care.

Measuring and monitoring quality of care can provide managed-care organizations with many benefits. One of them is improved patient care through proactive patient and provider communication. Managed care organizations have an increasing concern about underutilization of medical services, particularly with changing provider-reimbursement methods, such as capitation. By improving patient care and medical outcomes, managed care organizations can have a better-managed, healthier patient population, which in turn can lead to lower costs associated with covering such a population (Clinical Resource Management System, 1997).

Kongstvedt (1996) states that the most important use of physician profiles is producing physician feedback reports to help the physicians modify their own behavior. In providing the profiles, health plans intend to build good working relationships with physicians.

In 1998, Anne Freiburger stated that to help gain physician buy-in to the profiling process and to create solid measures, it is essential to include medical directors and respected community physicians in the development of measures to be used in the profile.

Joseph Carver (1999) states that it is essential that plan medical directors make 'office calls' to physicians who appear to have poor practice patterns. However, he states that "we don't pay anyone more or less if they are out of whack, and we don't remove them from the network of physicians." The profiles are meant to be purely educational and the visit helps develop a good relationship between the plan medical director and the physician.

Another use for physician profiles is that you can see which specialty groups had the highest costs and how these groups performed in relation to clinical guidelines. This analysis helps focus on areas that need most attention. Then information can be provided to physicians at a detailed, patient level. Also, data can be sorted to find detailed physician information (e.g. which physician demonstrated the most overutilization, highest cost per service, most services per patient, etc.) to ensure good network and physician management (Clinical Resource Management Systems, 1997).

In 1998, Robert W. Dubois suggests that in the quest for efficiency and cost effectiveness, profiling lends itself to the following four key functions:

- Selecting providers – health plans use physician profiles in combination with other quality information to make decisions about including or excluding physicians from their networks.
- Compensating providers – some health plans use physician profiles to reward accountable physicians. Health plans also use profiling to allocate bonuses or risk pool incentive funds.
- Offering intangible rewards – Physicians with good practice profiles may be exempted from certain precertification or utilization review mandates. This incentive is offered to providers with excellent practice patterns, to reduce their administrative burden.
- Identifying opportunities for improvement – through profiling, health plans and provider organizations can collaborate to identify physicians who achieve quality outcomes on a cost-effective basis. These practice patterns are then modeled throughout the provider network.

Basically, physician profiling should be an ongoing exercise to allow organizations not only to select and compensate excellent physicians, but also to identify opportunities for improvement throughout the organization (National Health Information, LLC, 1998).

Research and literature reported a number of efforts that evaluate physician profiling in various settings, and provided suggestions for producing and sharing profiles optimally.

Review of Literature Related in Terms of Methodology

The purpose of this section was to report on different methods used to evaluate the efficacy of physician profiling and information sharing. A large percentage of research in the affective domain used controlled trials to determine efficacy.

A study on education as a tool for modifying physician use of the laboratory was conducted by Marton, Tul, and Sox (1985) entitled "Modifying test-ordering behavior in the outpatient medical clinic. A controlled trial of two educational interventions." The researchers compared two interventions by assigning 56 medical house officers into four groups: control group; feedback group, which received feedback concerning its use of tests; manual group, which received a manual concerning cost-effective laboratory use; and manual plus feedback group, which received both interventions. A baseline measurement of laboratory utilization, through the use of encounter forms, was taken prior to the start of the study. Each house officer was also asked to fill out a questionnaire to measure knowledge about the use of laboratory tests in various clinical situations, before the study began. After the introduction of the interventions, i.e. feedback and formal instruction, house staff test use was monitored for seven and one-half months. The effects of the intervention were assessed by monitoring test ordering at every clinic visit, and by means of intervention and post-intervention questionnaires. Mean laboratory charges, and mean number of tests ordered, per patient visit per physician for each of the study groups both before and after the interventions were compared. Two-way analysis of variance tests to compare group means, and comparisons using both the Kruskal-Wallis test on ranks and Student's t test were also used. Paired comparisons were also made for

each group before and after the intervention. The researchers concluded that by using simple techniques, staff use of the outpatient laboratory could be modified.

A second study of quality concerning physician feedback, entitled "Feedback reduces test use in a health maintenance organization", was done by Berwick and Coltin (1986).

In a cross-over design controlled clinical trial, three interventions were tested for their impact on the rate of use of 13 commonly ordered blood tests among 35 internists in a health maintenance organization (HMO) that had three different ambulatory centers. The 13 tests were divided into three test groups that were balanced for type of test and rates of test use. Three interventions were developed for use in a modified Greco-Latin square design involving crossover of interventions, test groups, and center. The three interventions were as follows: (1) Test-Specific Education – over a two-week period, two departmental meetings were devoted to a discussion of appropriate use of tests in each one of the three test groups; (2) Peer Comparison Feedback on Cost of Test Use – clinicians received individualized reports on the rates at which they and their colleagues within their center were ordering each of the tests within a particular test group; (3) Peer Comparison Feedback on Yield of Tests – individuals received reports, and learned their own rank in their own center, on the rate of abnormal test results for each test in the appropriate test group. In the cross-over design, each test group was subjected to each of the three interventions in a different center. Rates of test use (expressed as tests ordered per 1,000 encounters) and variation (expressed as the coefficient of variation) in rates among physicians within centers were measured during baseline and intervention years. Effects of the intervention were measured by studying changes in rates during intervention periods, compared with immediately preceding nonintervention periods. Intervention effects on the change scores between the two periods were analyzed using analysis of variance and Kruskal-Wallis procedures.

A prospective, randomized, controlled study, was conducted at a Veterans Administration Hospital by Steele, et al (1989), to compare the cost effectiveness of two behaviorally oriented interventions designed to reduce physician's drug prescribing costs.

Physicians participating in the study were randomized into one of three groups using a table of random numbers. Clinical pharmacists visited one group of 11 physicians during weekly clinic sessions to counter-detail the prescribing of certain targeted drugs. A second group of 10 physicians were given data each week comparing their individual prescribing costs to those of their colleagues (peer-comparison feedback). A control group of 10 physicians received neither intervention. Written patient-specific suggestions for cost-effective prescribing were employed in both intervention groups. Baseline

prescribing costs were monitored through the hospital's computer during a three-month, pre-intervention period and a seven-month intervention period. Physician-specific drug cost data including cost per fill and number of prescriptions were collected from the hospital's pharmacy system. One-way analysis of variance with the Fisher LSD (least-significant difference) test was used to determine levels of significance for inter-group differences between the three study groups. Multiple regression analysis was used to describe the relationship between the number of clinical pharmacist suggestions and the percentage of prescribing changes adopted by physicians. It was concluded that personal visits by a doctor of pharmacy are a cost-effective technique for reducing physicians' prescribing costs, and are superior to peer-comparison feedback.

In 1991, Frazier, et al, conducted a prospective controlled trial in an internal medicine teaching clinic to determine whether an educational program using a drug cost manual could assist physicians in reducing their patients' out-of-pocket prescription drug expenses. Thirty one interns received a manual of comparative drug prices annotated with prescribing advice, two feedback reports, and weekly prescribing reminders. A control group of twenty interns concurrently participated in a manual-based cholesterol management education program. In addition, feedback reports were generated from the carbon copies of the prescriptions written by physicians, and were distributed to intervention group physicians only. Each report contained the physician's own data with averages for all physicians in the intervention group for comparison. Copies of 3012 prescriptions written over a period of eight months were analyzed. It was found that the intervention group physicians prescribed less expensive drugs within certain classes of drugs. The change in drug price score per prescription was –0.15. A score of 3 was assigned to the most expensive drug, 2 was assigned to intermediate-priced, and 1 was assigned to the least expensive drug. An increase of 0.74 months' supply of medication was dispensed per prescription, reducing dispensing fees. Outcomes for the intervention were compared with those of the control group using the PROC-GLM procedure in SAS.

Raw, unadjusted means and standard deviations were reported for the individual groups; an adjusted difference and its associated P value from a weighted analysis of covariance were reported for differences or comparisons between groups. The program was well accepted by physicians. The authors concluded that a relatively simple educational intervention can help physicians to reduce their patients' drug expenses and may serve as a model for incorporating cost information into the routine practice of medicine.

A study to describe and evaluate a controlled intervention to achieve more rational and economic drug therapy in general practice was conducted by Larsen and Kristensen (1992) at the University of Copenhagen's Institute of General Practice in Denmark. The study was based on peer comparison feedback and encouragement of local peer group discussions. Seven districts comprising 53 practices and six districts comprising 55 practices constituted the intervention and control groups, respectively. The process was evaluated by questionnaires to the general practitioners. Eighty eight percent of the physicians found the feedback relevant, and 74% expressed a wish for information about prescribing of particular drugs. The outcome evaluation was based on computerized registration of prescriptions. During the study period of six months the median costs per prescription increased significantly in the control group, while there was no significant change in the intervention group. There were no significant effects on the prescribed amounts, leading the authors to believe that the intervention was not effective.

McPhee, et al (1989), conducted a randomized controlled trial to determine effective methods of promoting routine cancer screening. Sixty two internal medicine residents were randomly assigned to one of three groups: residents to receive cancer screening reminders (computer-generated lists of overdue tests at patients' visits); audit

with feedback (monthly seminars about screening, with feedback about their performance rates); or no intervention (control group). Half of the residents in each group were also randomized to receive patient education (patients received literature and notices of overdue tests). The authors reviewed a sample of each physician's medical records to assess performance on seven tests during 9-month periods before and after initiating the interventions. Cancer screening reminders increased performance on six of seven tests; audit with feedback improved performance on four of seven tests; and patient education improved performance on one of two targeted breast cancer screening tests. The results of the study show that the cancer screening reminders strategy was the most effective in promoting the performance of routine cancer screening tests.

In 1989, Pugh, et al, conducted a controlled trial to determine the effect of daily feedback on inpatient charges and physician knowledge and behavior. Concurrent charge feedback is widely used as a method of minimizing hospitals' losses under the Medicare prospective payment system despite the fact that its effect on patient outcomes, physician behavior, or charges has not been studied in great detail. This study used two medical wards in an academic medical center to determine the effect of daily charge feedback on charges. Sixteen teaching physicians and sixty-eight house staff participated in the trial during a 35-week period, while taking care of 1057 eligible patients. To test for significant differences, the authors used t-tests. Since 45% of patients had planned protocol admissions on which the house staff had little change to impact, a subgroup analysis was performed, excluding these patients. A highly significant reduction in mean total charges (17%), length of stay (18%), room charges (18%), and diagnostic testing (20%) was found in the remaining patients. Patient outcomes in the two wards were

similar when comparing in-hospital mortality and preventable readmission within 30 days. The authors concluded that charge feedback alone is effective in a teaching hospital for decreasing charges.

A study to determine the effect of a physician education program on hospital length of stay and total patient charges was conducted by Johnson, et al, (1993).

An experimental-control group study was carried out in a large community hospital in which the experimental variable consisted of a one-time exposure of physicians to clinical and financial specifics regarding their individual treatment and management of pneumonia patients. A computerized report generating technology was used to create individualized reports for physicians, detailing their practice patterns with respect to pneumonia patients. Analysis of variance and t-tests were used for comparisons between the intervention and control groups and to test for significant differences. It was found that providing physicians with specific information about their practice behavior resulted in these same physicians using fewer hospital resources as expressed in total billed charges and length of stay. At the same time there was no compromise in outcomes for patients as measured by mortality, readmission rates, and infection/complication rates. The improvement in resource utilization was observed for two years following the provision of practice specific data to the experimental physicians.

Tierney, Miller, and McDonald (1990) studied the effect of informing physicians of the charges for outpatient diagnostics on their ordering of such tests in a primary medical care practice. The 121 physicians in the study all ordered tests at computer workstations. For the intervention group (half of the physicians), charges for the test being ordered and total charges for tests for that patient on that particular day were displayed on the computer screen. The control group of physicians (the other half) also used the computers, but did not receive a message about charges. The number of tests ordered and the average charge for tests per patient visit were the outcomes measured. The study included all patients and all visits to physicians, and also included all outpatient diagnostic tests. A questionnaire to determine the physicians' knowledge of test charges was administered once prior to the intervention and once after the

intervention (6 months later). For each physician, the mean number of tests ordered and the mean charge for tests per patient visit were calculated for each study period (preintervention, intervention, and postintervention). When comparing the mean values for the physicians in the intervention and control groups in the preintervention period, a weighted analysis of variance was used; for comparisons within the intervention and postintervention periods, a weighted analysis of covariance, with each physician's preintervention mean entered as a covariate, was used. To determine the accuracy of the physician's knowledge of test charges, the absolute value of the percent deviation of each physician's estimate of the charge for each test was calculated. This score was used to compare the knowledge of test charges of the physicians in the intervention and control groups at base line and the degree of improvement after the intervention. The authors concluded that the intervention was effective in significantly reducing the number and cost of tests ordered, especially for patients with scheduled visits. However, they noted that the effects of the intervention did not persist after it was discontinued.

In 1984, Winickoff, et al, conducted a study to improve physician compliance with colorectal cancer screening. The standard used required a digital examination and stool test for occult blood at annual check-ups of patients aged 40 years and older. Three intervention strategies to improve compliance with the standard were implemented during a 3 ½ -year period: educational meeting, retrospective feedback of group compliance rate, and retrospective feedback of individual compliance rate compared with that of peers. To evaluate the first two intervention strategies, a pre-test/post-test design was used. None of these two strategies resulted in a significant improvement in compliance. The third intervention (ranking with peers) was then implemented in a randomized clinical trial

with crossover design. The chi-square test was used to compare performance between groups and between periods. During the first 6-month period, physicians receiving feedback (group 1) improved from 66.0% to 79.9%, while the control group (group 2) also improved from 67.5% to 76.6%. Group 2 received feedback and group 1 did not in the second 6-month period. Group 1 maintained their rate at 80% while group 2 continued to improve from 76.6% to 84.0%. Even up to 12 months after the intervention, behavior changes persisted.

A study was conducted by C.C. Johnson and M. Martin to determine the effectiveness of a physician education program in reducing consumption of hospital resources in elective total hip replacement.

Orthopedic surgeons at a medical center in South Carolina were presented with verbal and written physician-specific materials during a 7-week period (intervention). Data on all patients with hip replacement surgery scheduled and completed in 1992 and 1993 were collected from hospital records. Xbar and R charts (control charts) were constructed to monitor any effects of the educational program on the overall consumption of hospital resources used in providing elective hip replacement as measured by length of stay and average total charges for each of the subgroups (control and intervention). Xbar charts plot the data points along with upper and lower control limits (95%) and a center line. These charts, along with R charts, provide immediate graphic indicators of the stability of a process. Length of stay and total charges was plotted on Xbar charts as a measure of process performance. R charts were created to provide measures of process variability. In addition, physicians were provided with surgeon-specific data profiling their own practice and that of their peers. Two-sample t-tests were used to test for statistical differences in mean length of stay and total charges after the educational materials were presented to the surgeons. Length of hospital stay for patients receiving total hip replacement dropped from 13.7 days to 9.9 days. Adjusted total charges dropped from $22,103 to $18,607. The variance for length of stay and total charges dropped by one half or more. Physician education programs are effective in reducing consumption of hospital resources. Statistical process control data formatting appears to make the data more usable to these physicians.

The research literature reported a number of methods to evaluate the efficacy of physician profiling and information sharing. It was found that a large percentage of

research in this area was conducted at hospitals, using controlled trials to determine efficacy.

Review of Literature Related in Terms of Content and Methodology

The purpose of this section was to report on research and literature that both evaluates physician profiling as a tool, and evaluates the efficacy of physician profiling and information sharing.

In 1994, H.G. Welch, et al, analyzed the inpatient practice patterns of physicians in Florida and Oregon, using 1991 data from Medicare's National Claims History File. Their study used the following methodological recommendations proposed by the Physician Payment Review Commission: (1) Profiles must be analyzed for a well-defined population; (2) A sufficient number of observations should be included in the profile to ensure that differences are not due to chance and; (3) Adjustment should be made for differences in case mix. The physician was considered the unit of analysis in the study, and this was made possible by the unique physician-identification numbers (UPINs). For each attending physician, the total relative value of all services delivered by a physician during a given hospital stay was determined. Relative value was expressed in relative value units (RVUs), according to the resource-based relative-value scale used by Medicare for calculating reimbursement. The profiles of attending physicians were examined at two levels of aggregation: the medical staffs of hospitals and the population of physicians in each of the two states. Practice patterns for different specialties and for selected types of service were examined to analyze the relevant components of the total relative value of the physicians' services per admission for medical staffs and for each

state as a whole. Differences in RVUs were used for comparison purposes. It was found that Florida physicians used markedly more resources, on average, than their colleagues in Oregon. It was also noted that each staff had a different practice pattern and would require different efforts to improve efficiency.

A review of the literature revealed just one study that both evaluated physician profiling as a tool, and evaluated the efficacy of physician profiling and information sharing. This study by Welch, et al, in 1994 relates most closely to the proposed research.

Summary

The review of the literature indicates that one of the basic reasons for profiling was to assist with continuous quality improvement efforts, while at the same time controlling costs. The review has also shown that there is varied evidence on the value of physician profiling, which creates a need for further study.

Profiling will lose its value if physician behavior does not change as a result. Physicians in general do not like to be told how to practice medicine. However, they want to compare favorably with their peers. Showing providers how they rank compared to their peers appears to be one of the most effective methods of change. Frequently physicians need no further prompting; they are independent thinkers and often will change their own behavior to bring their practice patterns in conformance to the norm. The health plan, however, needs to avoid any adversarial relationships and be willing to work in partnership with physicians in bringing about change. If profiling techniques can enhance this partnership, then physician profiling can be useful.

The information presented in this chapter was synthesized to develop the methodology used in this study, by selecting the appropriate profiling techniques, and choosing pertinent methods of analysis. Both the profiling techniques and methods of analysis were also chosen based on consistency with the constructs of the Deming Cycle.

CHAPTER III

METHODS AND PROCEDURES

Introduction

This chapter summarizes the methods and procedures used to collect, tabulate, and analyze research data. The primary purpose of the study was to assess the impact of physician profiling and information sharing in modifying practice patterns among The Plan HMO primary care physicians in the state of Tennessee.

In order to address the purpose of the study the following research questions were formulated:

1. How do physician utilization practice patterns compare between HMO primary care physicians who receive physician profiles and those who do not?
2. How do physician utilization practice patterns compare between rural and urban HMO primary care physicians who receive physician profiles?
3. How do physician utilization practice patterns compare between rural and urban HMO primary care physicians who do not receive physician profiles?

In the early 1990s the Physician Payment Review Commission identified important methodological criteria for physician profiling:

1. Profiles must be analyzed for a well-defined population.
2. An adequate number of observations should be included to ensure that differences are not due to chance.
3. Adjustments should be made for differences in case mix.

4. Profiles must be analyzed for a small enough organizational unit so that members can feel responsible for their results and can work as a team to identify any necessary courses of action.

The methods of physician profiling in this study have taken these issues into account.

The Deming Cycle (theoretical framework for the study) was used in the study at two levels. First, it was used as a formative evaluation framework to evaluate The Plan's physician profiling program. Second, it was proposed that individual physicians could use the Deming Cycle to improve preventive screening rates (Naidoo and McSharry, 1999).

Study Population

The target population for the study included all commercial HMO primary care physicians from The Plan, who practiced medicine within the state of Tennessee during the study period (1 January 1998 to 31 December 1999). An HMO primary care physician was defined as a provider of care for patients at a primary level, not necessarily requiring tertiary (hospital or emergent) or specialty (disease- or body-specific) level of care. Primary care physicians generally fall into the following four medical specialties:

1. Pediatrics
2. Internal Medicine
3. Family Practice
4. General Practice.

Physicians are credentialed as Family Practitioners if they have completed a Family Practice residency, i.e. they do not have to be board certified to get credentialed

for this, they only need to have completed their residency, which lasts 3 years. The focus of the residency is on all aspects of primary care, so the designation Family Practitioner means that the physician has significant formal training in primary care.

Physicians are credentialed as General Practitioners if they are practicing primary care but have not completed their residency. For instance, if a physician completes only an internship (one year past medical school), they may obtain a Tennessee medical license to practice primary care designated as a General Practitioner. Another example is that if a surgeon decides that he/she does not want to operate anymore but wants to practice primary care, they can be credentialed as a General Practitioner.

Criteria for inclusion in the study

All commercial HMO primary care physicians from The Plan were included in the study, either as part of the experimental group or the control group (to be discussed below). It was discovered that some physicians practiced medicine under multiple specialties. For the purposes of this study, only those physicians whose primary specialty was one of the four specialties listed above were selected.

Assignment to groups

To ensure that profile measurement differences were not due to chance, only those physicians who had 50 or more members assigned to them during the study period were profiled. This was also the criterion used to assign physicians into the intervention (experimental) and control groups:

Experimental group: those physicians who had 50 or more members assigned to them and received a physician profile during the study period.

Control group: those physicians who had less than 50 members assigned to them during the study period and did not receive a physician profile.

Physicians who intermittently had more than 50 members (received profiles) and less than 50 members (did not receive profiles) assigned to them during the study period were excluded from the study.

Rural and Urban groups: assignment to the rural and urban groups was based on Federal Information Processing Standards (FIPS) county codes issued by the National Institute of Technology (NIST), and metropolitan areas (MAs) as defined by the Office of Management and Budget (OMB) in the Federal Register. The FIPS county codes designated as rural and urban were matched to the physician's county code to obtain final determinations of rural and urban classifications.

A list of all the commercial HMO primary care physicians was obtained by downloading the HMO provider file from The Plan's provider database. The file was sorted in a spreadsheet program to extract only those physicians who were active during the study period, and practiced medicine in the state of Tennessee under the medical specialties listed above. Member and provider assignment tables were also downloaded from The Plan's data warehouse into a separate spreadsheet file. This file was joined to the HMO provider file mentioned earlier to obtain a list of membership assignments for those PCPs who practiced in the state of Tennessee. This final list was sorted and used to separate the physicians into the experimental and control groups.

Sampling

No formal sampling technique was used to select those physicians who fell into either the control or experimental groups, or the rural or urban groups. The experimental

group (those physicians who received a profile) consisted of 463 physicians, whereas the control group (those physicians not receiving a profile) consisted of 1,213 physicians. A power analysis was conducted to ensure that the number of physicians in each group was large enough to detect significant differences at an alpha level of 0.05. The power was determined to be 0.94 (power > 0.80 is generally considered to be good).

Rural physicians who received a profile consisted of 122 physicians, and urban physicians who received a profile consisted of 341 physicians. The power was determined to be 0.83 for comparisons between these two groups.

Rural physicians who did not receive a profile consisted of 403 physicians, and urban physicians who did not receive a profile consisted of 810 physicians. The power was determined to be 0.88 for comparisons between these two groups.

Instrumentation

Development of the Intervention Instrument

Physician Profiling

A physician profile was defined as a physician-specific report with patient-level and procedure-specific detail, outlining precisely how physicians vary from their peers in the way they provide care based on specialty, practice, region, or numerous other relevant metrics. The utilization measures that were used in the profile being described for this study were chosen by a team of clinicians to cover as much of the broad spectrum of services for which the primary care physician was responsible. The team of clinicians consisted of three medical directors, with support from a nurse and two medical data analysts. Measures were chosen based on utilization characteristics that were directly

controlled by physicians, and included screening rates as well. The specific measures included:

1. Number of primary care visits per member per year.
2. Number of specialist visits per member per year.
3. Number of prescriptions written per thousand members per year.
4. Proportion of total prescriptions written as generic.
5. Number of hospital admissions per thousand members per year.
6. Number of emergency room visits per thousand members per year.
7. Number of outpatient surgeries per thousand members per year.
8. Mammography screening rate for women between the ages of 52 and 69. To be eligible a woman must have been in the plan for the last two years without a break in membership exceeding 45 days in either year. (Measure calculated for internal medicine, family practice, and general practice only.)
9. Hemoglobin A1C screening rate for all diabetic members in the past year. (Measure calculated for internal medicine, family practice, and general practice only.)
10. Cholesterol screening rate for men between the ages of 35 and 65, and women between 45 and 65 years of age. To be included in the measure, a member must have been eligible for benefits from the plan for at least one year, with a break of no more than 45 days. (Measure calculated for internal medicine, family practice, and general practice only.)

A quarterly profile was developed for each physician in the experimental group for 8 quarters. For measures 1–7 above quarterly rates were first calculated and then extrapolated to annualized measures, to match standardized healthcare industry metrics.

Human Subjects Factors

The study protocol was approved by both The Plan and the University of Tennessee Institutional Review Boards. Permission to use The Plan claims data was approved by the agency. In addition, The Plan ensured that the data were appropriately aggregated to prevent identification of individual subjects (physicians and patients).

Extracting Data for the Profiles

The first step in producing the quarterly profiles involved extracting one quarter's worth of claims information, for the measures mentioned above, from the claims database. The data were downloaded into a file using the SAS software tool. This file included records on physicians' services that contained the different data fields (e.g. unique physician identifier, physician address, physician zip code, procedure code, ICD-9 code, dates of service, unique member identification, member age, member sex, and so on) required to produce each of the measures on the profile.

Adjustment for Case Mix

One of the hallmarks of competent profiling was to effectively adjust for the illness burden of the provider's patients. Physicians who were confident that they were practicing high-quality care were not likely to accept data to the contrary unless it took into account their population's underlying illness burden; otherwise, these physicians were likely to excuse their performance with the assertion that "my patients are sicker, older, and different". A number of case-mix adjustment techniques were available that

dealt with the problem of differing illness burdens among physicians to level the playing field. Adjusted Clinical Groups (ACGs), which are a population-based case-mix adjuster (developed by the Johns Hopkins University) that uses ICD-9 codes, age, and gender for classification of patients were used in this study.

The ACG Grouper, for each physician, using just the following data elements: unique member identification, age, sex, and ICD-9 code were applied to the claims data file. Case-mix characteristics of patients assigned to primary care physicians were used as adjustment factors for average per-patient resource use during the quarter being profiled.

Calculation of Measures

The researcher chose to use a statistic called the percentile to represent the measures in the profile. In statistical terms, the *p*th percentile has *p*% of the data below it, with the median being the 50^{th} percentile. For example, if a physician's cholesterol screening rate was in the 76^{th} percentile, this meant that 76% of that physician's peers had a cholesterol screening rate lower than or equal to the physician's rate.

Using the case-mix adjusted (ACG) data from the claims file, percentiles for each measure were calculated in SAS for each individual physician.

Producing the Final Profile

In producing the physician profiling instrument for the intervention, the following important suggestions (Donald E. Lighter, 1998) were adhered to:

- To allow easy reading for physicians, a consistent report format was used to allow easier interpretation of the results.
- The profile was presented in the form of a graph since that type of presentation was much more effective than numeric reports.

- Physicians also appreciate concise reports that convey a great deal of information for rapid analysis. Hence, each physician's profile consisted of a one-page bar graph that contained the percentiles for the various measures.

The final profile for the quarter was produced in a software package called Microsoft Excel, for each physician who had 50 or more members assigned to them in that quarter.

Intervention

A complete packet of information that included the profile, a cover letter (explaining the purposes of the profile and it's intended use) from The Plan's chief medical officer, a user's guide (that explained the methodology, the measures, and statistics), and a short survey (to elicit feedback about the profile) were mailed to the individual physicians. In the survey, physicians were asked to evaluate the profiles based on usefulness, accuracy, and the ability to provide information to help evaluate the physician's patient panel. A sample of the complete physician profiling packet can be found in the Appendix.

At the end of 1998 physicians received profiles for the first three quarters of 1998 for informative purposes. All quarterly profiles mailed thereafter formed part of the intervention. Profiles were mailed out with one-quarter lag. For example, the first quarter 1999 profile was mailed at the beginning of July 1999. This was done to allow at least three months claims run-out to ensure completion of unaccounted claims liability.

Preparation of Data for Analysis

Using the case-mix adjusted claims file (described above), the data were examined at three levels of aggregation: physicians who received a profile (experimental group), and physicians who did not receive a profile (control group); rural physicians who received a

profile and urban physicians who received a profile; and rural physicians who did not receive a profile and urban physicians who did not receive a profile. For each group, a pre-intervention measure (spanning 1998), and a post-intervention measure (spanning 1999) was taken. For each measurement period, and each group as a whole, utilization rates were calculated for the ten individual utilization measures described above.

The aggregate data were stored in a file for statistical analysis in the SAS program. SAS results were reviewed for accuracy by a senior statistician prior to use.

Statistical Design

The study was classified as a quasi-experimental design since physicians were not randomly assigned into either the experimental or control groups, or rural or urban groups. To determine whether any improvements in utilization practice patterns among primary care physicians were attributable to the physician profiling intervention, the data for all ten measures on the profile were subjected to statistical analysis. For the purposes of the study, improvement for each measure was defined as follows:

1. Increase in number of primary care visits per member per year.
2. Decrease in number of specialist visits per member per year.
3. Decrease in number of prescriptions written per thousand members per year.
4. Increase in proportion of total prescriptions written as generic.
5. Decrease in number of hospital admissions per thousand members per year.
6. Decrease in number of emergency room visits per thousand members per year.
7. Decrease in number of outpatient surgeries per thousand members per year.

8. Increase in the Mammography screening rate for women between the ages of 52 and 69.
9. Increase in the Hemoglobin A1C screening rate for all diabetic members in the past year.
10. Increase in the Cholesterol screening rate for men between the ages of 35 and 65, and women between 45 and 65 years of age.

The first step was to describe the data for each measure with descriptive statistics using measures of central tendency (mean) and measures of variability (standard deviation). The Wilks-Shapiro test-statistic was used to test the underlying data for normality. Since the underlying data were found to be non-normal, non-parametric statistics (Wilcoxon Rank Sum tests, Wilcoxon Matched-Pairs Signed-Ranks test, and Chi-square tests) were used to facilitate inferences about the impact of physician profiling in modifying practice patterns. Selection of statistics were based on the recommendations of Evans (1995) and the practices of other researchers (e.g., Winickoff, 1984; Marton, 1985).

Testing was conducted at the 0.05 level of significance. This reference probability represented the alpha-level or the probability of making a Type I error. For example, at the 0.05 level, if 100 samples had been selected, and 100 hypotheses tested, the wrong conclusion would have been made about five times.

Independent Two-Sample Wilcoxon Rank Sum Test

For non-normal data from two independent samples, one of the most common non-parametric methods to compare means was the Wilcoxon Rank Sum test. This test was

used to compare whether there were significant differences in means for six of the profile measures (number of primary care visits per member per year, number of specialist visits per member per year, number of prescriptions written per thousand members per year, number of emergency room visits per thousand members per year, number of outpatient surgeries per thousand members per year, number of inpatient admissions per thousand members per year) between the following groups:

1. Pre-intervention Control Group and Pre-intervention Experimental Group.
2. Post-intervention Control Group and Post-intervention Experimental Group.
3. Pre-intervention Rural Experimental Group and Pre-intervention Urban Experimental Group.
4. Post-intervention Rural Experimental Group and Post-intervention Urban Experimental Group.
5. Pre-intervention Rural Control Group and Pre-intervention Urban Control Group.
6. Post-intervention Rural Control Group and Post-intervention Urban Control Group.

The general form of the Wilcoxon Rank Sum test-statistic can be expressed as follows:

$T = \Sigma R_j$ = Sum of Ranks

$$W = \frac{T - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}$$

Where R_j is the rank of the jth observation, and T is the rank score. W is the Wilcoxon test statistic and n is the sample size.

The probability of a greater W value under the null hypothesis (population difference for the two groups = 0) was given by a P value. The greater the difference in the rank sum for the two groups, the smaller the P value.

One of the reasons why the Wilcoxon Rank Sum test was chosen for this study is that it has the ability to handle non-normal data. Wilcoxon performs an analysis of the ranks of the data or the Wilcoxon scores. A continuity correction is also done whenever the two-sample test is used (as in this case).

Chi-Square Test

For discrete data from two independent samples, the simplest and most common method to determine whether observed differences in proportions between study groups were statistically significant was the use of the chi-square test. This test was used to compare whether there were significant differences in four of the profile measures (proportion of total prescriptions written as generic, mammography screening rate, Hemoglobin A1c screening rate, and cholesterol screening rate) between the following groups:

1. Pre-intervention Control Group and Pre-intervention Experimental Group.
2. Post-intervention Control Group and Post-intervention Experimental Group.
3. Pre-intervention Rural Experimental Group and Pre-intervention Urban Experimental Group.
4. Post-intervention Rural Experimental Group and Post-intervention Urban Experimental Group.
5. Pre-intervention Rural Control Group and Pre-intervention Urban Control Group.

6. Post-intervention Rural Control Group and Post-intervention Urban Control Group.

The general form of the chi-square statistic can be expressed as follows:

$$\chi^2_{(df)} = \Sigma \frac{(O - E)^2}{E}$$

Where O = the observed count in a category, E = expected count in that category under the null hypothesis, Σ = the sum of all categories, and *df* = the degrees of freedom associated with the test statistic.

The expected values for each category were determined based on the assumption that the null hypothesis of no difference between the study groups in the proportions of the factor of interest is true. This assumption can be applied to data that are presented in the form of a two-by-two table as shown in Table 3-1 below. The table of expected values was as shown in Table 3-2 below. The chi-square statistic for a particular measure on the profile could then easily be calculated using values from Table 3-1 and Table 3-2.

The probability of a greater χ^2 value under the null hypothesis (equality of measures between the two groups) was given by a *P* value. The greater the difference in proportions, the smaller the *P* value. To convert the chi-square statistic into a *P* value involved determining the relevant degrees of freedom and searching for the appropriate value in a table of the chi-square distribution. In general, the degrees of freedom for an r-by-c contingency table was equal to the product of the number of rows in the table minus one times the number of columns minus one, which is expressed as (r-1)(c-1). Thus, for any two-by-two table (as was the case for all the profile measures), the degrees of

Table 3-1. Presentation of data in a two-by-two table: observed values

Observed Values	Experimental Group	Control Group	**Total Patients**
Test Done	a	b	**a + b**
Test Not Done	c	d	**c + d**
Total Patients	**a + c**	**b + d**	**T = a + b + c + d**

Table 3-2. Calculation of expected values in a two-by-two table

Expected Values	Experimental Group	Control Group
Test Done	(a + b)(a + c)/T	(a + b)(b + d)/T
Test Not Done	(c + d)(a + c)/T	(c + d)(b + d)/T

freedom was equal to (2-1)(2-1) or 1. The *P* value could then be determined from the chi-square probability table.

Wilcoxon Matched-Pairs Signed-Ranks Test

To test for significant differences between the following groups:

1. Pre-intervention experimental group and post-intervention experimental group
2. Pre-intervention control group and post-intervention control group
3. Pre-intervention Rural Experimental Group and Post-intervention Rural Experimental Group
4. Pre-intervention Urban Experimental Group and Post-intervention Urban Experimental Group

5. Pre-intervention Rural Control Group and Post-intervention Rural Control Group
6. Pre-intervention Urban Control Group and Post-intervention Urban Control Group.

the Wilcoxon Matched-Pairs Signed-Ranks test was used. The Wilcoxon Matched-Pairs Signed-Ranks test is a non-parametric test, and was used in this case for all ten measures on the profile. One of the main reasons this test was chosen was the fact that both the pre- and post-intervention sample for both the experimental group and control groups, and urban and rural groups, respectively was dependent.

The Wilcoxon Matched-Pairs Signed-Ranks test is a nonparametric procedure employed in a hypothesis testing situation involving a design with two dependent samples. It is based on the following assumptions:

1. Each of n pairs of matched subjects has two interval/ratio scores (each score being obtained under one of the two experimental conditions).
2. The sample of n subjects is randomly selected from the population it represents.
3. The distribution of the difference scores in the populations represented by the two samples is symmetric about the median of the population of difference scores.

The procedure for the test is as follows: A difference score is computed for each pair of matched subjects by subtracting a subject's score in Condition 2 from their score in Condition 1. The interval/ratio difference scores of matched pairs of subjects are then ranked. The hypothesis evaluated with this test is whether or not in the underlying populations represented by the experimental condition, the median of the difference scores equals zero. If a significant difference is obtained ($P < 0.05$), it indicates there is a high likelihood the two conditions represent two different populations.

Data Analyses

Test statistics and corresponding *P* values were calculated for each measure in all four comparisons. For example, the format of results for the comparison between the pre-intervention experimental group and the pre-intervention control group, and the post-intervention experimental group and the post-intervention control group was as shown in Table 3-3 below (similar for respective rural vs urban comparison).

The format of results for the comparison between the pre-intervention

Table 3-3. Sample format for presentation of results between groups.

Comparison between the pre-intervention experimental group and the pre-intervention control group	**Pre-intervention Experimental Group Mean**	**Pre-intervention Control Group Mean**	**Test Statistic**	***P* value**
1. Primary care visits per member per year			Wilcoxon Rank Sum	
2. Specialist visits per member per year			Wilcoxon Rank Sum	
3. Number of prescriptions per member per year			Wilcoxon Rank Sum	
4. Proportion of generic prescriptions written			χ^2 test (DF=1)	
5. Admissions per one thousand members per year			Wilcoxon Rank Sum	
6. Emergency Room visits per one thousand members per year			Wilcoxon Rank Sum	
7. Outpatient surgeries per one thousand members per year			Wilcoxon Rank Sum	
8. Cholesterol screening rate for eligible members			χ^2 test (DF=1)	
9. Mammography screening rate for eligible members			χ^2 test (DF=1)	
10. Hemoglobin A1C screening rate for eligible members			χ^2 test (DF=1)	

Table 3-4. Sample format for presentation of results within groups.

Comparison between the pre-intervention control group and the post-intervention control group	Pre-intervention Control Group Mean	Post-intervention Control Group Mean	Mean Difference	*P* value
1. Primary care visits per member per year			Wilcoxon Signed Rank	
2. Specialist visits per member per year			Wilcoxon Signed Rank	
3. Number of prescriptions per member per year			Wilcoxon Signed Rank	
4. Proportion of generic prescriptions written			Wilcoxon Signed Rank	
5. Admissions per one thousand members per year			Wilcoxon Signed Rank	
6. Emergency Room visits per one thousand members per year			Wilcoxon Signed Rank	
7. Outpatient surgeries per one thousand members per year			Wilcoxon Signed Rank	
8. Cholesterol screening rate for eligible members			Wilcoxon Signed Rank	
9. Mammography screening rate for eligible members			Wilcoxon Signed Rank	
10. Hemoglobin A1C screening rate for eligible members			Wilcoxon Signed Rank	

experimental group and the post-intervention experimental group, and the pre-intervention control group and the post-intervention control group was as shown in Table 3-4 above (similar for respective rural vs urban comparison).

Research Question Analyses

Research Question No. 1: **How do physician utilization practice patterns compare between primary care physicians who receive physician profiles and those who do not?**

χ^2 tests and Wilcoxon Rank Sum tests were used to test for significance ($P < 0.05$) of difference between the rates of measures on the profiles for the following two comparisons:

1. Pre-intervention Control Group and Pre-intervention Experimental Group. (Baseline.)
2. Post-intervention Control Group and Post-intervention Experimental Group. (Post-intervention measure.)

The Wilcoxon Matched-Pairs Signed-Ranks test was used ($P < 0.05$) to test for significant differences between the rates of measures on the profiles for the following two comparisons:

1. Pre-intervention Experimental Group and Post-intervention Experimental Group.
2. Pre-intervention Control Group and Post-intervention Control Group.

Research Question No. 2: **How do physician utilization practice patterns compare between rural and urban HMO primary care physicians who receive physician profiles?**

χ^2 tests and Wilcoxon Rank Sum tests were used to test for significance ($P < 0.05$) of difference between the rates of measures on the profiles for the following two comparisons:

1. Pre-intervention Rural Experimental Group and Pre-intervention Urban Experimental Group. (Baseline.)
2. Post-intervention Rural Experimental Group and Post-intervention Urban Experimental Group. (Post-intervention measure.)

The Wilcoxon Matched-Pairs Signed-Ranks test was used ($P < 0.05$) to test for significant differences between the rates of measures on the profiles for the following two comparisons:

1. Pre-intervention Rural Experimental Group and Post-intervention Rural Experimental Group.
2. Pre-intervention Urban Experimental Group and Post-intervention Urban Experimental Group.

Research Question No. 3: **How do physician utilization practice patterns compare between rural and urban HMO primary care physicians who do not receive physician profiles?**

χ^2 tests and Wilcoxon Rank Sum tests were used to test for significance ($P < 0.05$) of difference between the rates of measures on the profiles for the following two comparisons:

1. Pre-intervention Rural Control Group and Pre-intervention Urban Control Group. (Baseline.)
2. Post-intervention Rural Control Group and Post-intervention Urban Control Group. (Post-intervention measure.)

The Wilcoxon Matched-Pairs Signed-Ranks test was used ($P < 0.05$) to test for significant differences between the rates of measures on the profiles for the following two comparisons:

1. Pre-intervention Rural Control Group and Post-intervention Rural Control Group.

2. Pre-intervention Urban Control Group and Post-intervention Urban Control Group.

Summary

This chapter presented the research methodology and procedures, including issues related to the study population, instrumentation, and methods of statistical analysis. Claims data from The Plan's claims databases were used. A quasi-experimental method, applying non-parametric statistics was used to determine the impact of physician profiling and information sharing in modifying practice patterns among HMO primary care physicians in the state of Tennessee. The next chapter, Chapter IV, contains the analyses and findings of the data collected using the methodology and procedures described in Chapter III.

CHAPTER IV

ANALYSIS AND INTERPRETATION OF DATA

Introduction

This chapter presents the analysis and interpretation of the data from this study. The primary purpose of the study was to assess the impact of physician profiling and information sharing in modifying practice patterns among HMO primary care physicians in the state of Tennessee.

In order to address the purpose of the study the following research questions were formulated:

1. How do physician utilization practice patterns compare between HMO primary care physicians who receive physician profiles and those who do not?
2. How do physician utilization practice patterns compare between rural and urban HMO primary care physicians who receive physician profiles?
3. How do physician utilization practice patterns compare between rural and urban HMO primary care physicians who do not receive physician profiles?

In reviewing the results, descriptive statistics will be discussed first which describe the characteristics of the total sample of physicians, followed by the characteristics of the rural and urban physician subsamples. This will be followed by tests of significance contrasting profile measures that are significantly different between the subsamples.

Description of the Groups Studied

In total, 1,676 primary care physicians were included in the study. Figure 4-1 shows that the total group consisted of 24% (406) internal medicine physicians, 8% (135)

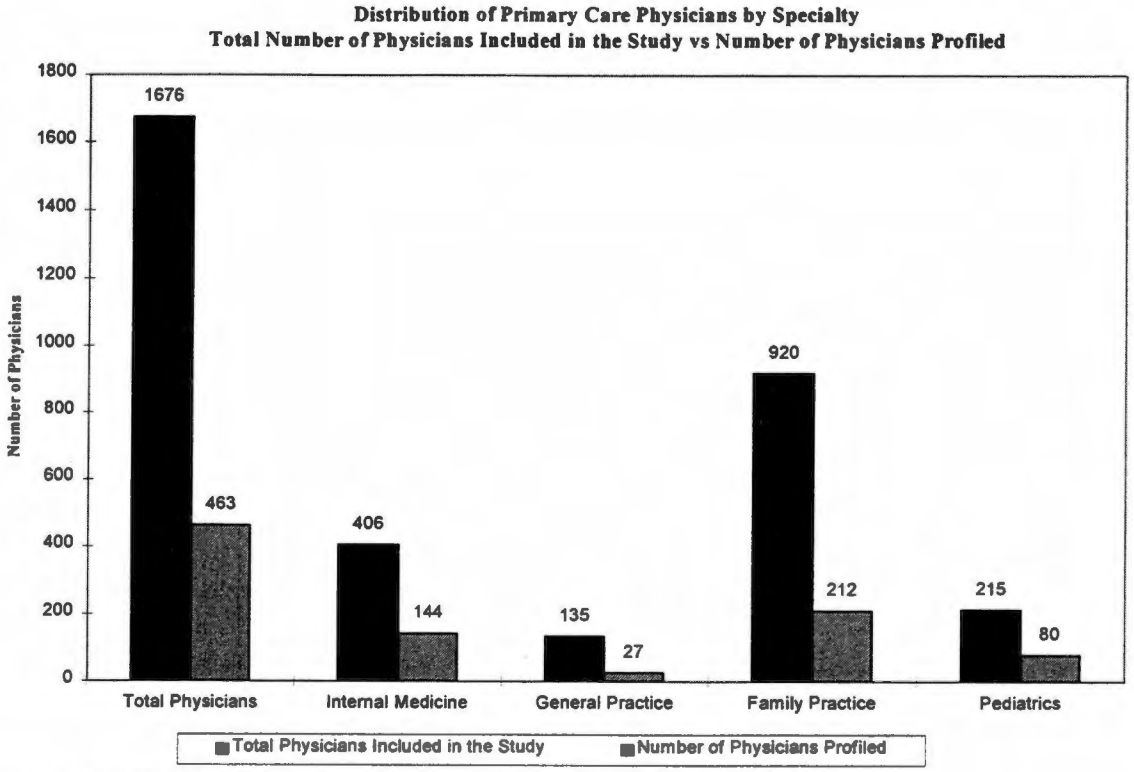


Figure 4-1. Distribution of primary care physicians by specialty.

general practice physicians, 55% (920) family practice physicians, and 13% (215) pediatricians. These 1,676 primary care physicians were responsible for providing services to, and coordinating the care of, 80,502 HMO members (see Figure 4-2 below).

The experimental group (group of physicians receiving profiles) consisted of 463 primary care physicians, and accounted for 28% of the total physicians in the study. Table 4-1 shows that this group consisted of 31% internal medicine physicians, 6% general practice physicians, 46% family practice physicians, and 17% pediatricians. From Figure 4-2 we can see that this group of physicians were responsible for coordinating the care of 61,427 HMO members.

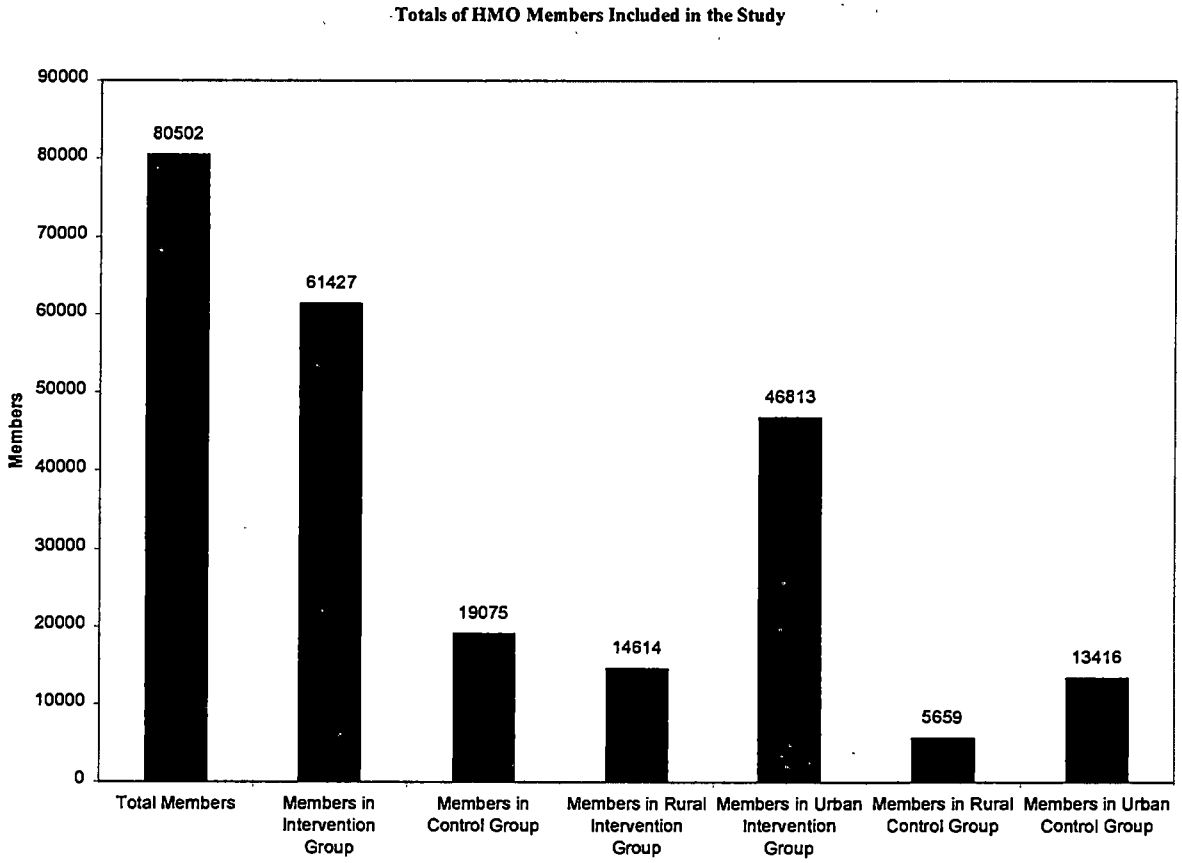


Figure 4-2. Number of members included in each study group.

Table 4-1. Distribution of physicians by specialty for each group studied.

DISTRIBUTION OF PHYSICIANS BY SPECIALTY FOR EACH GROUP STUDIED		
STUDY GROUP	**NUMBER OF PHYSICIANS**	**PERCENT OF TOTAL FOR GROUP**
Experimental Group		
Internal Medicine	144	31.1%
General Practice	27	5.8%
Family Practice	212	45.8%
Pediatrics	80	17.3%
Control Group		
Internal Medicine	262	21.6%
General Practice	108	8.9%
Family Practice	708	58.4%
Pediatrics	135	11.1%
Rural Experimental Group		
Internal Medicine	26	21.3%
General Practice	14	11.5%
Family Practice	71	58.2%
Pediatrics	11	9.0%
Urban Experimental Group		
Internal Medicine	118	34.6%
General Practice	13	3.8%
Family Practice	141	41.3%
Pediatrics	69	20.2%
Rural Control Group		
Internal Medicine	63	15.6%
General Practice	52	12.9%
Family Practice	264	65.5%
Pediatrics	24	6.0%
Urban Control Group		
Internal Medicine	199	24.6%
General Practice	56	6.9%
Family Practice	444	54.8%
Pediatrics	111	13.7%

Experimental Group = Physicians receiving profiles
Control Group = Physicians not receiving profiles

The control group (group of physicians not receiving profiles) consisted of 1,213 primary care physicians. Table 4-1 shows that this group consisted of 22% internal medicine physicians, 9% general practice physicians, 58% family practice physicians, and 11% pediatricians. Figure 4-2 shows that this group of physicians was responsible for coordinating the care of 19,075 HMO members.

The rural experimental group (rural physicians receiving profiles) consisted of 122 primary care physicians, and accounted for 26% of the total physicians in the experimental group. From Table 4-1 we see that this group consisted of 21% internal medicine physicians, 12% general practice physicians, 58% family practice physicians, and 9% pediatricians. This group of physicians was responsible for coordinating the care of 14,614 HMO members.

The urban experimental group (urban physicians receiving profiles) consisted of 341 primary care physicians, and accounted for 74% of the total physicians in the experimental group. Table 4-1 shows that this group consisted of 35% internal medicine physicians, 4% general practice physicians, 40% family practice physicians, and 20% pediatricians. Figure 4-2 shows that this group of physicians was responsible for coordinating the care of 46,813 members.

The rural control group (rural physicians not receiving profiles) consisted of 403 physicians, and accounted for 33% of the total physicians in the control group. From Table 4-1 we see that this group consisted of 16% internal medicine physicians, 13% general practice physicians, 65% family practice physicians, and 6% pediatricians. This group of physicians was responsible for coordinating the care of 5,659 members.

The urban control group (urban physicians not receiving profiles) consisted of 810 physicians, and accounted for 67% of the total physicians in the control group. Table 4-1 shows that this group of physicians consisted of 24% internal medicine physicians, 7% general practice physicians, 55% family practice physicians, and 14% pediatricians. Figure 4-2 shows that this group of physicians was responsible for coordinating the care of 13,416 members.

For each group a total of 10 measures were analyzed. These consisted of 7 utilization measures: number of primary care visits per member per year, number of specialist visits per member per year, number of prescriptions per member per year, proportion of generic prescriptions written by the physicians, number of inpatient admissions per 1,000 members per year, number of emergency room visits per 1,000 members per year, number of outpatient surgeries per 1,000 members per year; and 3 quality measures: physician cholesterol screening rate, physician mammography rate, and physician hemoglobin A1c screening rate.

Significance Tests for Research Question 1

How do physician utilization practice patterns compare between HMO primary care physicians who receive physician profiles and those who do not?

The analysis for this research question comprised four different comparisons: pre-intervention experimental group versus pre-intervention control group; post-intervention experimental group versus post-intervention control group; pre-intervention experimental group versus post-intervention experimental group; pre-intervention control group versus post-intervention control group. For between-group comparisons, Wilcoxon Rank Sum

tests and chi-square tests were employed to determine statistically significant differences between the groups. For within-group comparisons, Wilcoxon Signed Rank tests were used. Differences were considered to be statistically significant if the p-value for the test was less than 0.05.

Utilization Measures

The pre-intervention experimental group and pre-intervention control group comparison showed significant differences across all 7 measures. Significantly lower utilization ($p < 0.05$) for two of the measures - % generic prescriptions written and ER visits/1000 members PY - was observed for the pre-intervention experimental group. The pre-intervention control group had significantly lower utilization ($p < 0.05$) for primary care visits PMPY, specialty visits PMPY, number of prescriptions PMPY, admits/1000 members PY, and outpatient surgeries/1000 members PY (see Table 4-2).

Comparison between the post-intervention experimental group and post-intervention control group also showed significant difference across all 7 utilization measures. The post-intervention experimental group showed significantly lower utilization ($p < 0.05$) only for the % generic prescriptions written. All other measures were significantly lower ($p < 0.05$) in the post-intervention control group (see Table 4-3).

The pre-intervention and post-intervention experimental group comparison showed significant differences for all utilization measures except outpatient surgeries/1000 members PY. Significantly higher utilization ($p < 0.05$) for primary care visits PMPY, number of prescriptions PMPY, and % generic prescriptions written were seen in the post-intervention experimental group. Significantly lower utilization ($p < 0.05$) in the

Table 4-2. Comparison between the pre-intervention experimental group and the pre-intervention control group

Comparison between the pre-intervention experimental group and the pre-intervention control group	**Pre-intervention Experimental Group Mean**	**Pre-intervention Control Group Mean**	**Test Statistic**	***P* value**
1. Primary care visits per member per year	0.94	0.57	11.80 Wilcoxon Rank Sum	0.000
2. Specialist visits per member per year	0.56	0.41	11.23 Wilcoxon Rank Sum	0.000
3. Number of prescriptions per member per year	8.37	5.20	13.54 Wilcoxon Rank Sum	0.000
4. Proportion of generic prescriptions written	39.47%	42.00%	253.32 χ^2 test (DF=1)	0.001
5. Admissions per one thousand members per year	29.36	15.30	19.92 Wilcoxon Rank Sum	0.000
6. Emergency Room visits per one thousand members per year	80.39	85.60	7.92 Wilcoxon Rank Sum	0.000
7. Outpatient surgeries per one thousand members per year	417.89	282.62	11.39 Wilcoxon Rank Sum	0.000
8. Cholesterol screening rate for eligible members	0.24	0.27	24.34 χ^2 test (DF=1)	0.001
9. Mammography screening rate for eligible members	0.67	0.67	0.06 χ^2 test (DF=1)	0.810
10. Hemoglobin A1C screening rate for eligible members	0.57	0.64	13.59 χ^2 test (DF=1)	0.001

Table 4-3. Comparison between the post-intervention experimental group and the post-intervention control group

Comparison between the post-intervention experimental group and the post-intervention control group	**Post-intervention Experimental Group Mean**	**Post-intervention Control Group Mean**	**Test Statistic**	***P* value**
1. Primary care visits per member per year	1.19	1.05	5.13 Wilcoxon Rank Sum	0.000
2. Specialist visits per member per year	0.52	0.39	10.20 Wilcoxon Rank Sum	0.000
3. Number of prescriptions per member per year	8.79	6.33	12.43 Wilcoxon Rank Sum	0.000
4. Proportion of generic prescriptions written	41.22%	44.83%	587.79 χ^2 test (DF=1)	0.001
5. Admissions per one thousand members per year	21.80	18.32	16.42 Wilcoxon Rank Sum	0.000
6. Emergency Room visits per one thousand members per year	66.79	62.43	12.02 Wilcoxon Rank Sum	0.000
7. Outpatient surgeries per one thousand members per year	411.62	285.60	11.00 Wilcoxon Rank Sum	0.000
8. Cholesterol screening rate for eligible members	0.25	0.29	80.81 χ^2 test (DF=1)	0.001
9. Mammography screening rate for eligible members	0.66	0.71	12.14 χ^2 test (DF=1)	0.001
10. Hemoglobin A1C screening rate for eligible members	0.57	0.69	57.44 χ^2 test (DF=1)	0.001

post-intervention experimental group was seen for specialist visits PMPY, admits/1000 members PY, and ER visits/1000 members PY (see Table 4-4).

Comparison between the pre-intervention control group and post-intervention control group showed significant difference ($p < 0.05$) in 4 of the 7 utilization measures. The post-intervention control group showed significantly higher utilization for primary care visits PMPY, number of prescriptions PMPY, and % generic prescriptions written. The same group showed significantly lower utilization ($p < 0.05$) for ER visits/1000 members PY (see Table 4-5).

Quality Measures

Comparison between the pre-intervention experimental group and pre-intervention control group showed significantly lower rates ($p < 0.05$) for cholesterol screening and Hemoglobin A1c screening in the pre-intervention experimental group. From table 4-2 we see that there was no difference in rates for mammography screening between the groups.

The post-intervention experimental group and post-intervention control group comparison showed significant differences ($p < 0.05$) across all 3 quality measures. Significantly lower rates were seen in the post-intervention experimental group (see Table 4-3).

Comparison between the pre-intervention and post-intervention experimental group showed no significant differences between the groups for all 3 quality measures (see Table 4-4).

Table 4-4. Comparison between the pre-intervention experimental group and the post-intervention experimental group

Comparison between the pre-intervention experimental group and the post-intervention experimental group	Pre-intervention Experimental Group Mean	Post-intervention Experimental Group Mean	Mean Difference	*P* value
1. Primary care visits per member per year	0.94	1.19	0.24 Wilcoxon Signed Rank	0.000
2. Specialist visits per member per year	0.56	0.52	-0.04 Wilcoxon Signed Rank	0.000
3. Number of prescriptions per member per year	8.37	8.79	0.46 Wilcoxon Signed Rank	0.000
4. Proportion of generic prescriptions written	39.47%	41.22%	1.75% Wilcoxon Signed Rank	0.000
5. Admissions per one thousand members per year	29.36	21.80	-7.70 Wilcoxon Signed Rank	0.000
6. Emergency Room visits per one thousand members per year	80.39	66.79	-13.72 Wilcoxon Signed Rank	0.000
7. Outpatient surgeries per one thousand members per year	417.89	411.62	-3.26 Wilcoxon Signed Rank	0.843
8. Cholesterol screening rate for eligible members	0.24	0.25	0.01 Wilcoxon Signed Rank	0.093
9. Mammography screening rate for eligible members	0.67	0.66	-0.00 Wilcoxon Signed Rank	0.712
10. Hemoglobin A1C screening rate for eligible members	0.57	0.57	0.00 Wilcoxon Signed Rank	0.825

Table 4-5. Comparison between the pre-intervention control group and the post-intervention control group

Comparison between the pre-intervention control group and the post-intervention control group	Pre-intervention Control Group Mean	Post-intervention Control Group Mean	Mean Difference	*P* value
1. Primary care visits per member per year	0.57	1.05	0.47 Wilcoxon Signed Rank	0.000
2. Specialist visits per member per year	0.41	0.39	-0.02 Wilcoxon Signed Rank	0.076
3. Number of prescriptions per member per year	5.20	6.33	1.19 Wilcoxon Signed Rank	0.000
4. Proportion of generic prescriptions written	42.00%	44.83%	2.83% Wilcoxon Signed Rank	0.012
5. Admissions per one thousand members per year	15.30	18.32	3.14 Wilcoxon Signed Rank	0.062
6. Emergency Room visits per one thousand members per year	85.60	62.43	-22.19 Wilcoxon Signed Rank	0.000
7. Outpatient surgeries per one thousand members per year	282.62	285.60	3.42 Wilcoxon Signed Rank	0.406
8. Cholesterol screening rate for eligible members	0.27	0.29	0.03 Wilcoxon Signed Rank	0.000
9. Mammography screening rate for eligible members	0.67	0.71	0.025 Wilcoxon Signed Rank	0.207
10. Hemoglobin A1C screening rate for eligible members	0.64	0.69	0.00 Wilcoxon Signed Rank	0.7917

The pre-intervention and post-intervention control group comparison showed a significant increase ($p < 0.05$) in the cholesterol screening rate for the post-intervention group. From Table 4-5 we see that there was no significant difference in the rates for mammography screening and hemoglobin A1c screening.

Significance Tests for Research Question 2

How do physician utilization practice patterns compare between rural and urban HMO primary care physicians who receive physician profiles?

The analysis for this research question comprised four different comparisons: pre-intervention rural experimental group versus pre-intervention urban experimental group; post-intervention rural experimental group versus post-intervention urban experimental group; pre-intervention rural experimental group versus post-intervention rural experimental group; pre-intervention urban experimental group versus post-intervention urban experimental group. For between-group comparisons, Wilcoxon Rank Sum tests and chi-square tests were employed to determine statistically significant differences between the groups. For within-group comparisons, Wilcoxon Signed Rank tests were used. Differences were considered to be statistically significant if the p-value for the test was less than 0.05.

Utilization Measures

The pre-intervention rural experimental group and pre-intervention urban experimental group comparison showed significant differences for 4 of the 7 measures. Significantly lower utilization ($p < 0.05$) for three of the measures - specialist visits

PMPY, ER visits/1000 members PY, and outpatient surgeries/1000 members PY - was observed for the pre-intervention rural experimental group. The pre-intervention rural experimental group had significantly higher utilization ($p < 0.05$) for % generic prescriptions written (see Table 4-6)

Table 4-7 shows significant differences for 3 of the 7 measures when comparing the post-intervention rural experimental group and the post-intervention urban experimental group. Specialist visits PMPY and outpatient surgeries/1000 members PY were significantly lower in the post-intervention rural experimental group. Proportion of generic prescriptions written were significantly higher ($p < 0.05$) in the post-intervention rural experimental group.

The pre-intervention rural experimental group and post-intervention rural experimental group comparison showed significant differences for 4 of the 7 utilization measures. Significantly higher utilization ($p < 0.05$) for two of the measures – primary care visits PMPY and number of prescriptions PMPY – was seen in the post-intervention rural experimental group. Specialist visits PMPY and admits/1000 members PY were significantly lower ($p < 0.05$) in the post-intervention rural experimental group (see Table 4-8).

Table 4-9 shows significant differences for 6 of the 7 utilization measures when comparing the pre-intervention urban experimental group and post-intervention urban experimental group. Primary care visits PMPY, number of prescriptions written PMPY, and % generic prescriptions written were all significantly higher ($p < 0.05$) in the post-intervention urban experimental group. Specialist visits PMPY, admits/1000 members

Table 4-6. Comparison between the pre-intervention rural experimental group and the pre-intervention urban experimental group

Comparison between the pre-intervention rural experimental group and the pre-intervention urban experimental group	Pre-intervention Rural Experimental Group Mean	Pre-intervention Urban Experimental Group Mean	Test Statistic	*P* value
1. Primary care visits per member per year	0.92	0.95	-0.39 Wilcoxon Rank Sum	0.692
2. Specialist visits per member per year	0.48	0.59	-3.14 Wilcoxon Rank Sum	0.001
3. Number of prescriptions per member per year	7.81	8.57	-1.26 Wilcoxon Rank Sum	0.207
4. Proportion of generic prescriptions written	42.65%	38.62%	625.92 χ^2 test (DF=1)	0.001
5. Admissions per one thousand members per year	29.41	29.34	0.24 Wilcoxon Rank Sum	0.809
6. Emergency Room visits per one thousand members per year	69.13	84.42	-3.69 Wilcoxon Rank Sum	0.000
7. Outpatient surgeries per one thousand members per year	294.33	462.10	-4.39 Wilcoxon Rank Sum	0.000
8. Cholesterol screening rate for eligible members	0.21	0.25	102.16 χ^2 test (DF=1)	0.001
9. Mammography screening rate for eligible members	0.64	0.68	4.106 χ^2 test (DF=1)	0.043
10. Hemoglobin A1C screening rate for eligible members	0.45	0.60	51.83 χ^2 test (DF=1)	0.001

Table 4-7. Comparison between the post-intervention rural experimental group and the post-intervention urban experimental group

Comparison between the post-intervention rural experimental group and the post-intervention urban experimental group	Post-intervention Rural Experimental Group Mean	Post-intervention Urban Experimental Group Mean	Test Statistic	*P* value
1. Primary care visits per member per year	1.07	1.23	-1.74 Wilcoxon Rank Sum	0.080
2. Specialist visits per member per year	0.44	0.55	-3.09 Wilcoxon Rank Sum	0.002
3. Number of prescriptions per member per year	9.19	8.65	1.05 Wilcoxon Rank Sum	0.294
4. Proportion of generic prescriptions written	42.30%	40.87%	82.80 χ^2 test (DF=1)	0.001
5. Admissions per one thousand members per year	28.14	19.50	1.31 Wilcoxon Rank Sum	0.187
6. Emergency Room visits per one thousand members per year	65.64	67.21	-0.65 Wilcoxon Rank Sum	0.510
7. Outpatient surgeries per one thousand members per year	308.11	449.34	-4.24 Wilcoxon Rank Sum	0.000
8. Cholesterol screening rate for eligible members	0.21	0.27	190.89 χ^2 test (DF=1)	0.001
9. Mammography screening rate for eligible members	0.66	0.66	0.037 χ^2 test (DF=1)	0.848
10. Hemoglobin A1C screening rate for eligible members	0.45	0.61	70.11 χ^2 test (DF=1)	0.001

Table 4-8. Comparison between the pre-intervention rural experimental group and the post-intervention rural experimental group

Comparison between the pre-intervention rural experimental group and the post-intervention rural experimental group	Pre-intervention Rural Experimental Group Mean	Post-intervention Rural Experimental Group Mean	Mean Difference	*P* value
1. Primary care visits per member per year	0.92	1.07	0.16 Wilcoxon Signed Rank	0.018
2. Specialist visits per member per year	0.48	0.44	-0.04 Wilcoxon Signed Rank	0.003
3. Number of prescriptions per member per year	7.81	9.19	1.49 Wilcoxon Signed Rank	0.000
4. Proportion of generic prescriptions written	42.65%	42.30%	-0.35% Wilcoxon Signed Rank	0.254
5. Admissions per one thousand members per year	29.41	28.14	-1.45 Wilcoxon Signed Rank	0.016
6. Emergency Room visits per one thousand members per year	69.13	65.64	-3.51 Wilcoxon Signed Rank	0.144
7. Outpatient surgeries per one thousand members per year	294.33	308.11	18.52 Wilcoxon Signed Rank	0.599
8. Cholesterol screening rate for eligible members	0.21	0.21	0.01 Wilcoxon Signed Rank	0.528
9. Mammography screening rate for eligible members	0.64	0.66	0.04 Wilcoxon Signed Rank	0.304
10. Hemoglobin A1C screening rate for eligible members	0.45	0.45	0.03 Wilcoxon Signed Rank	0.368

Table 4-9. Comparison between the pre-intervention urban experimental group and the post-intervention urban experimental group

Comparison between the pre-intervention urban experimental group and the post-intervention urban experimental group	Pre-intervention Urban Experimental Group Mean	Post-intervention Urban Experimental Group Mean	Mean Difference	*P* value
1. Primary care visits per member per year	0.95	1.23	0.27 Wilcoxon Signed Rank	0.000
2. Specialist visits per member per year	0.59	0.55	-0.04 Wilcoxon Signed Rank	0.000
3. Number of prescriptions per member per year	8.57	8.65	0.09 Wilcoxon Signed Rank	0.002
4. Proportion of generic prescriptions written	38.62%	40.87%	2.25% Wilcoxon Signed Rank	0.000
5. Admissions per one thousand members per year	29.34	19.50	-9.97 Wilcoxon Signed Rank	0.000
6. Emergency Room visits per one thousand members per year	84.42	67.21	-17.44 Wilcoxon Signed Rank	0.000
7. Outpatient surgeries per one thousand members per year	462.10	449.34	-11.20 Wilcoxon Signed Rank	0.602
8. Cholesterol screening rate for eligible members	0.25	0.27	0.01 Wilcoxon Signed Rank	0.100
9. Mammography screening rate for eligible members	0.68	0.66	-0.02 Wilcoxon Signed Rank	0.228
10. Hemoglobin A1C screening rate for eligible members	0.60	0.61	-0.02 Wilcoxon Signed Rank	0.319

PY, and ER visits/1000 members PY were significantly lower in the post-intervention urban experimental group.

Quality Measures

The pre-intervention rural experimental group and the pre-intervention urban experimental group comparison showed significant differences ($p < 0.05$) in rates across all three quality measures. Significantly higher rates were seen in the post-intervention urban experimental group (see table 4-6).

Table 4-7 shows that the cholesterol screening rate and hemoglobin A1c screening rate in the post-intervention urban experimental group were significantly higher ($p < 0.05$) than the post-intervention rural experimental group. There was no significant difference in the mammography screening rate between the two groups.

On comparing the pre-intervention rural experimental group with the post-intervention rural experimental group no significant differences were seen in any of the three quality measures (see table 4-8).

The pre-intervention urban experimental group and post-intervention urban experimental group comparison also showed no significant differences in the three quality measures (see table 4-9).

Significance Tests for Research Question 3

How do physician utilization practice patterns compare between rural and urban HMO primary care physicians who do not receive physician profiles?

The analysis for this research question comprised four different comparisons: pre-intervention rural control group versus pre-intervention urban control group; post-intervention rural control group versus post-intervention urban control group; pre-intervention rural control group versus post-intervention rural control group; pre-intervention urban control group versus post-intervention urban control group. For between-group comparisons, Wilcoxon Rank Sum tests and chi-square tests were employed to determine statistically significant differences between the groups. For within-group comparisons, Wilcoxon Signed Rank tests were used. Differences were considered to be statistically significant if the p-value for the test was less than 0.05.

Utilization Measures

The pre-intervention rural control group and pre-intervention urban control group comparison showed significant differences for 4 of the 7 utilization measures (see Table 4-10). Specialist visits PMPY and outpatient surgeries/1000 members PY were significantly lower ($p < 0.05$) in the pre-intervention rural control group. Proportion of generic prescriptions written and admits/1000 members PY were significantly higher in the pre-intervention rural control group.

Table 4-11 shows significant differences for 3 of the 7 measures when comparing the post-intervention rural control group with the post-intervention urban control group. The post-intervention rural control group had significantly lower utilization for specialist visits PMPY and outpatient surgeries/1000 members PY, and significantly higher utilization for % generic prescriptions written ($p < 0.05$).

Table 4-10. Comparison between the pre-intervention rural control group and the pre-intervention urban control group

Comparison between the pre-intervention rural control group and the pre-intervention urban control group	Pre-intervention Rural Control Group Mean	Pre-intervention Urban Control Group Mean	Test Statistic	*P* value
1. Primary care visits per member per year	0.61	0.55	0.64 Wilcoxon Rank Sum	0.522
2. Specialist visits per member per year	0.38	0.43	-2.36 Wilcoxon Rank Sum	0.018
3. Number of prescriptions per member per year	5.09	5.26	-0.57 Wilcoxon Rank Sum	0.568
4. Proportion of generic prescriptions written	44.19%	41.06%	96.30 χ^2 test (DF=1)	0.001
5. Admissions per one thousand members per year	15.76	15.07	-1.97 Wilcoxon Rank Sum	0.049
6. Emergency Room visits per one thousand members per year	91.85	82.49	-1.53 Wilcoxon Rank Sum	0.125
7. Outpatient surgeries per one thousand members per year	233.45	307.08	-3.30 Wilcoxon Rank Sum	0.001
8. Cholesterol screening rate for eligible members	0.24	0.28	8.50 χ^2 test (DF=1)	0.004
9. Mammography screening rate for eligible members	0.59	0.68	5.40 χ^2 test (DF=1)	0.020
10. Hemoglobin A1C screening rate for eligible members	0.59	0.66	3.76 χ^2 test (DF=1)	0.053

Table 4-11. Comparison between the post-intervention rural control group and the post-intervention urban control group

Comparison between the post-intervention rural control group and the post-intervention urban control group	Post-intervention Rural Control Group Mean	Post-intervention Urban Control Group Mean	Test Statistic	*P* value
1. Primary care visits per member per year	1.01	1.08	-1.24 Wilcoxon Rank Sum	0.215
2. Specialist visits per member per year	0.31	0.43	-4.84 Wilcoxon Rank Sum	0.000
3. Number of prescriptions per member per year	6.45	6.27	0.72 Wilcoxon Rank Sum	0.473
4. Proportion of generic prescriptions written	47.02%	43.83%	122.88 χ^2 test (DF=1)	0.001
5. Admissions per one thousand members per year	21.63	16.72	-0.21 Wilcoxon Rank Sum	0.831
6. Emergency Room visits per one thousand members per year	71.55	57.99	-0.65 Wilcoxon Rank Sum	0.514
7. Outpatient surgeries per one thousand members per year	254.38	300.80	-2.24 Wilcoxon Rank Sum	0.025
8. Cholesterol screening rate for eligible members	0.24	0.31	44.00 χ^2 test (DF=1)	0.001
9. Mammography screening rate for eligible members	0.66	0.72	5.456 χ^2 test (DF=1)	0.019
10. Hemoglobin A1C screening rate for eligible members	0.59	0.73	17.597 χ^2 test (DF=1)	0.001

The pre-intervention rural control group and post-intervention rural control group comparison showed significant differences for 4 of the utilization measures (see Table 4-12). Primary care visits PMPY and number of prescriptions written PMPY were significantly higher ($p < 0.05$) in the post-intervention rural control group. Specialist visits PMPY and ER visits/1000 members PY were significantly lower in the post-intervention rural control group.

Table 4-13 shows significant differences for 5 of the 7 measures when comparing the pre-intervention urban control group with the post-intervention urban control group. The post-intervention urban control group had significantly higher utilization for primary care visits PMPY, number of prescriptions written PMPY, % generic prescriptions written, and admits/1000 members per year. This group also had significantly lower utilization for ER visits/1000 members PY ($p < 0.05$).

Quality Measures

On comparing the pre-intervention rural control group and pre-intervention urban control group, significantly higher rates were found in the pre-intervention urban control group for cholesterol screening and mammography screening. No significant difference was noted between the groups for the hemoglobin A1c measure (see table 4-10).

The post-intervention urban control group had significantly higher utilization ($p < 0.05$) across all 3 quality measures when compared to the post-intervention rural control group (see Table 4-11).

Table 4-12. Comparison between the pre-intervention rural control group and the post-intervention rural control group

Comparison between the pre-intervention rural control group and the post-intervention rural control group	Pre-intervention Rural Control Group Mean	Post-intervention Rural Control Group Mean	Mean Difference	*P* value
1. Primary care visits per member per year	0.61	1.01	0.39 Wilcoxon Signed Rank	0.001
2. Specialist visits per member per year	0.38	0.31	-0.07 Wilcoxon Signed Rank	0.008
3. Number of prescriptions per member per year	5.09	6.45	1.39 Wilcoxon Signed Rank	0.000
4. Proportion of generic prescriptions written	44.19%	47.02%	2.83% Wilcoxon Signed Rank	0.137
5. Admissions per one thousand members per year	15.76	21.63	5.77 Wilcoxon Signed Rank	0.870
6. Emergency Room visits per one thousand members per year	91.85	71.55	-18.94 Wilcoxon Signed Rank	0.000
7. Outpatient surgeries per one thousand members per year	233.45	254.38	19.20 Wilcoxon Signed Rank	0.649
8. Cholesterol screening rate for eligible members	0.24	0.24	0.04 Wilcoxon Signed Rank	0.214
9. Mammography screening rate for eligible members	0.59	0.66	0.02 Wilcoxon Signed Rank	0.441
10. Hemoglobin A1C screening rate for eligible members	0.59	0.59	0.00 Wilcoxon Signed Rank	0.734

Table 4-13. Comparison between the pre-intervention urban control group and the post-intervention urban control group

Comparison between the pre-intervention urban control group and the post-intervention urban control group	Pre-intervention Urban Control Group Mean	Post-intervention Urban Control Group Mean	Mean Difference	*P* value
1. Primary care visits per member per year	0.55	1.08	0.52 Wilcoxon Signed Rank	0.000
2. Specialist visits per member per year	0.43	0.43	-0.00 Wilcoxon Signed Rank	0.711
3. Number of prescriptions per member per year	5.26	6.27	1.09 Wilcoxon Signed Rank	0.000
4. Proportion of generic prescriptions written	41.06%	43.83%	2.77% Wilcoxon Signed Rank	0.023
5. Admissions per one thousand members per year	15.07	16.72	1.86 Wilcoxon Signed Rank	0.016
6. Emergency Room visits per one thousand members per year	82.49	57.99	-23.78 Wilcoxon Signed Rank	0.000
7. Outpatient surgeries per one thousand members per year	307.08	300.80	-4.27 Wilcoxon Signed Rank	0.505
8. Cholesterol screening rate for eligible members	0.28	0.31	0.02 Wilcoxon Signed Rank	0.220
9. Mammography screening rate for eligible members	0.68	0.72	0.05 Wilcoxon Signed Rank	0.260
10. Hemoglobin A1C screening rate for eligible members	0.66	0.73	-0.02 Wilcoxon Signed Rank	0.940

When comparing the pre-intervention rural control group and the post-intervention rural control, no significant differences were found in the quality measures (see Table 4-12).

Similarly, when comparing the pre-intervention urban control group and the post-intervention urban control group, no significant differences were found across all 3 quality measures (see Table 4-13).

Summary

This chapter presented the results of the analysis of the data and the interpretation of these results. Overall, the group of physicians receiving profiles formed approximately one third of the total physicians. There were more family practice and internal medicine physicians in the study than there were general practice physicians and pediatricians. Although the experimental group physicians in the study numbered fewer than the control group physicians, they coordinated the care for more than three quarters of the members in the HMO. As expected, there were more urban physicians in the study than there were rural physicians. The urban physicians also reported coordinating the care for more members than the rural physicians.

In total, 12 comparisons were analyzed for each of 7 utilization measures and 3 quality measures to test for significant differences between and within various groups (experimental, control, rural, urban, pre-intervention, post-intervention). This resulted in 120 significance tests.

Comparisons between the pre-intervention experimental and control groups, and comparisons between the post-intervention experimental and control groups showed

significant differences across all 7 utilization measures. Most of the measures were lower in the control group than in the experimental group. Over time (i.e. pre-intervention versus post-intervention) the experimental group had significant improvement in 5 of the utilization measures, whereas the control group had significant improvement in only 3 of the measures. Both the experimental and control groups showed significant reductions in emergency room utilization.

For the quality measures there were significant differences between the pre-intervention experimental and control groups, and between the post-intervention experimental and control groups, with the control group having better rates in both comparisons. However, there was almost no significant improvement over time for the quality measures within both the experimental and control groups.

Except for specialist visits PMPY and % generic prescriptions written, there were no significant differences between the pre-intervention rural and urban experimental group comparison, and the post-intervention rural and urban experimental group comparison. For both these measures, the rural group had significantly better utilization than the urban group. Over time (i.e. pre-intervention versus post-intervention) the rural experimental group showed significant improvement in 3 of the utilization measures, whereas the urban experimental group showed significant improvement in 5 of the utilization measures.

For the quality measures there were significant differences between the pre-intervention rural and urban experimental groups, and between the post-intervention rural and urban experimental groups, with the urban group having better rates in both

comparisons. However, there was no significant improvement over time for the quality measures within both the rural experimental and urban experimental groups.

Except for specialist visits PMPY and % generic prescriptions written, there were no significant differences between the pre-intervention rural and urban control group comparison, and the post-intervention rural and urban control group comparison. For both these measures, the rural group had significantly better utilization than the urban group. Over time (i.e. pre-intervention versus post-intervention) the rural control group, as well as the urban control group, showed significant improvement in 3 of the utilization measures. Both groups showed significant reductions in emergency room utilization over time.

For the quality measures there were significant differences between the pre-intervention rural and urban control groups, and between the post-intervention rural and urban control groups, with the urban group having better rates in both comparisons. However, there was no significant improvement over time for the quality measures within both the rural control and urban control groups.

The next chapter, Chapter V, will address the summary, findings, conclusions, and recommendations based on the analysis and interpretation of the results.

CHAPTER V

SUMMARY OF THE FINDINGS, CONCLUSIONS, AND RECOMMENDATIONS

Introduction

This chapter presents the study findings, conclusions, and recommendations. The primary purpose of the study was to assess the impact of physician profiling and information sharing in modifying practice patterns among HMO primary care physicians in the state of Tennessee.

To fulfill this purpose, the utilization patterns of 1,676 HMO primary care physicians, who practiced medicine within the state of Tennessee in 1998 and 1999, were analyzed in aggregate based on various groupings. Physicians were analyzed based on their assignment to either an experimental or control group. The experimental group (physicians receiving profiles) consisted of 463 physicians, while the control group (physicians not receiving profiles) consisted of 1,213 physicians. More analysis was conducted based on the time of intervention (start of the physician profiling program in December 1998) – the pre-intervention period was calendar year 1998 and the post-intervention period was calendar year 1999. A further analysis was conducted between rural and urban physician groupings within both the experimental and control groups. Comparisons between pairs of groups and within groups were made to determine if there were any significant differences in ten measures that appeared on the physician profile. These ten measures, listed below, determined a physician's practice pattern:

1. Number of primary care visits per member per year.
2. Number of specialist visits per member per year.
3. Number of prescriptions written per thousand members per year.

4. Proportion of total prescriptions written as generic.
5. Number of emergency room visits per thousand members per year.
6. Number of outpatient surgeries per thousand members per year.
7. Number of hospital admissions per thousand members per year.
8. Mammography screening rate for women between the ages of 52 and 69.
9. Hemoglobin A1C screening rate for all diabetic members in the past year.
10. Cholesterol screening rate for men between the ages of 35 and 65, and women between 45 and 65.

Data obtained from The Plan's claims database system were used to test for significant differences in the ten measures between and within the various groupings. Test statistics included Wilcoxon Rank Sum tests, Chi-square tests, and Wilcoxon Matched-Pairs Signed Rank tests.

Findings

Overall descriptive findings of the study are discussed first, followed by the findings specific to each research question.

There were more control group (physicians not receiving profiles) than experimental group (physicians receiving profiles) physicians in the study, with the experimental group physicians coordinating the care for more than three quarters of the members in the HMO. There were more urban physicians in the study than there were rural physicians. The urban physicians coordinated the care for approximately three quarters of the members in the HMO. There were almost four times as many family

practice and internal medicine physicians in the study than there were general practice physicians and pediatricians.

Research Question-1: How do physician utilization practice patterns compare between HMO primary care physicians who receive physician profiles and those who do not?

1. When compared to the pre-intervention experimental group, significantly better utilization was seen for 5 of the 7 utilization measures in the pre-intervention control group. Primary care visits PMPY and ER visits/1000 members PY were better in the pre-intervention experimental group. The post-intervention control group showed significantly better utilization than the post-intervention experimental group for 6 of the 7 utilization measures. Although there were significant differences in both groups at baseline (pre-intervention) and post-intervention, the difference in utilization between the two groups had narrowed in the post-intervention period. The control group appeared to have a better practice pattern with respect to 5 of the 7 utilization measures than the experimental group, both at baseline and post-intervention. These were specialist visits PMPY, number of prescriptions PMPY, % generic prescriptions, admits/1000 members PY, and outpatient surgeries/1000 members PY.
2. A comparison of the pre-intervention and post-intervention experimental group showed significant improvement in 5 of the 7 utilization measures in the post-intervention experimental group. Emergency room utilization in the experimental group was significantly reduced by 17% from pre-intervention to post-intervention. When comparing the pre-intervention and post-intervention control group, significant improvement was seen in 3 of the 7 utilization measures. Also, emergency room

utilization in the control group was significantly reduced by about 27% from pre-intervention to post-intervention. A significantly greater proportion (61.2%) of physicians in the experimental group showed improvement in 4 or more metrics than the physicians in the control group (39.8%). Figure 5-1 below gives an example where the experimental group does much better than the control group.

3. The cholesterol screening rate and hemoglobin A1c screening rate were much higher in the control group than the experimental group at baseline. There was no difference

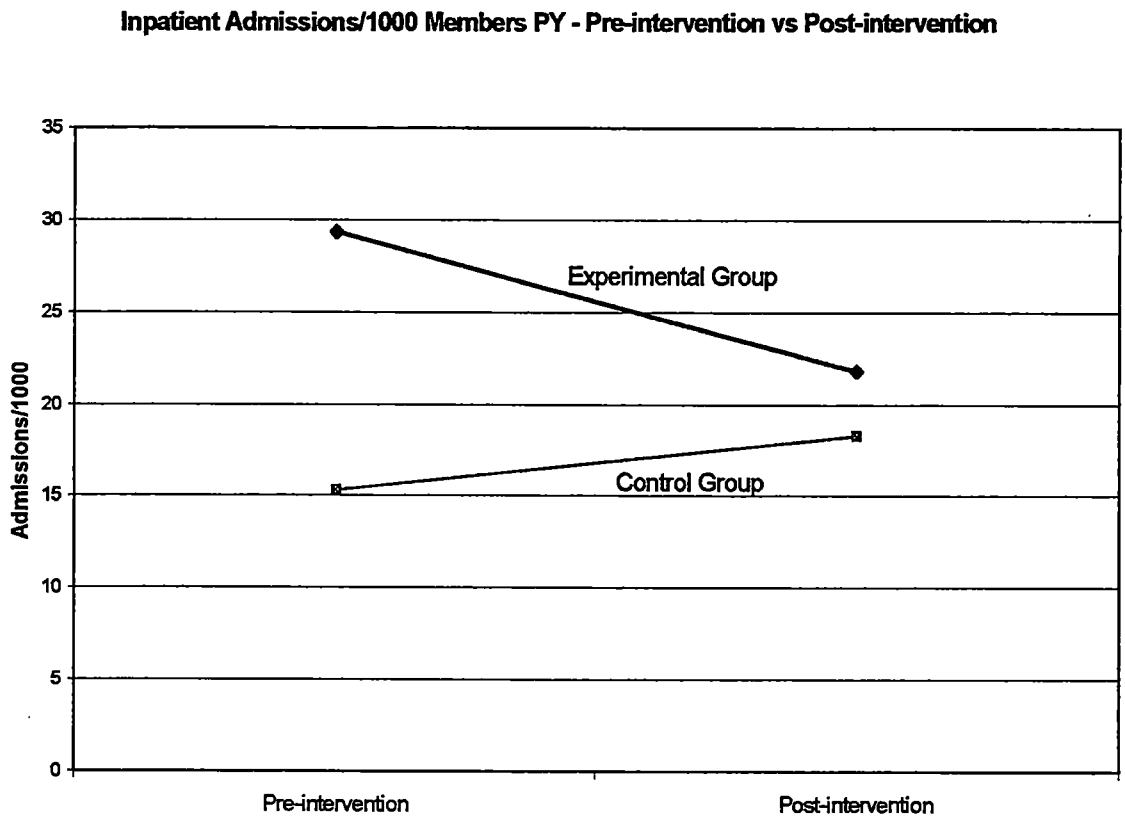


Figure 5-1. Inpatient Admissions/1000 members per year – Pre-intervention versus Post-intervention.

in the mammography screening rate between the groups at baseline. Significantly higher screening rates were noticed for all three quality measures in the post-intervention control group, when compared to the post-intervention experimental group. The control group appeared to have a better practice pattern, with respect to the three quality measures, than the experimental group both at baseline and post-intervention.

4. Comparing the pre-intervention and post-intervention experimental group showed no significant improvement in any of the three quality measures. Cholesterol screening showed the only significant improvement over time when comparing the pre-intervention and post-intervention control group. Except for cholesterol screening in the control group, both the control and experimental groups showed no improvement from baseline to post-intervention.

Research Question-2: How do physician utilization practice patterns compare between rural and urban HMO primary care physicians who receive physician profiles?

1. When compared to the pre-intervention urban experimental group, significantly better utilization was seen for 4 of the 7 utilization measures in the pre-intervention rural experimental group. The post-intervention rural experimental group showed significantly better utilization than the post-intervention urban experimental group for 3 of the 7 utilization measures. The rural experimental group appeared to have a better practice pattern with respect to 3 of the 7 utilization measures than the urban experimental group, both at baseline and post-intervention. These were specialist

visits PMPY, % generic prescriptions written, and outpatient surgeries/1000 members PY.

2. A comparison of the pre-intervention and post-intervention rural experimental group showed significant improvement in 3 of the 7 utilization measures in the post-intervention rural experimental group. When comparing the pre-intervention and post-intervention urban experimental group, significant improvement was seen in 5 of the 7 utilization measures. Also, emergency room utilization in the urban experimental group was significantly reduced by about 20% from pre-intervention to post-intervention. The urban experimental group had shown improvement in more of the utilization metrics than the rural experimental group, from baseline to post-intervention.
3. All three quality measures were significantly higher in the urban experimental group than the rural experimental group at baseline. Significantly higher screening rates were noticed for cholesterol screening and hemoglobin A1c screening in the post-intervention urban experimental group, when compared to the post-intervention rural experimental group. The urban experimental group appeared to have a better practice pattern, with respect to the three quality measures, than the rural experimental group both at baseline and post-intervention.
4. Comparison of the pre-intervention and post-intervention rural experimental group, and comparison of the pre-intervention and post-intervention urban experimental group showed no significant improvement in any of the three quality measures over time.

Research Question-3: How do physician utilization practice patterns compare between rural and urban HMO primary care physicians who do not receive physician profiles?

1. When compared to the pre-intervention urban control group, significantly better utilization was seen for 3 of the 7 utilization measures in the pre-intervention rural control group – number of specialist visits PMPY and % generic prescriptions written. The pre-intervention urban control group had significantly better utilization than the pre-intervention rural control group for admits/1000 members PY. The post-intervention rural control group showed significantly better utilization than the post-intervention urban control group for 3 of the 7 utilization measures. The rural control group appeared to have a better practice pattern with respect to 3 of the 7 utilization measures than the urban control group, both at baseline and post-intervention. These were specialist visits PMPY, % generic prescriptions written, and outpatient surgeries/1000 members PY.
2. A comparison of the pre-intervention and post-intervention rural control group showed significant improvement in 3 of the 7 utilization measures over time. Emergency room utilization in the rural control group was significantly reduced by 22% from baseline to post-intervention. When comparing the pre-intervention and post-intervention urban control group, significant improvement was seen in 3 of the 7 utilization measures. Emergency room utilization in the urban control group was significantly reduced by about 30% from pre-intervention to post-intervention. Both the urban control group and rural control group had shown improvement in three of the utilization metrics, from baseline to post-intervention.

3. Two of the three quality measures – cholesterol screening and mammography screening - were significantly higher in the urban control group than the rural control group at baseline. Significantly higher screening rates were noticed for all three quality measures in the post-intervention urban control group, when compared to the post-intervention rural control group. The urban control group appeared to have a better practice pattern, with respect to the three quality measures, than the rural control group, both at baseline and post-intervention.
4. Comparison of the pre-intervention and post-intervention rural control group, and comparison of the pre-intervention and post-intervention urban control group showed no significant improvement in any of the three quality measures over time.

Conclusions

Based on the analysis of the data and subsequent findings the following conclusions have been reached.

1. Physicians not receiving profiles appear to have better practice patterns than physicians receiving profiles.
2. Physician profiling appears to be effective in modifying practice patterns for utilization measures, but not for the quality measures.
3. Physician profiling has a greater effect on urban physicians than rural physicians.

Recommendations

The following recommendations are driven by the purpose of this study and are based on the aforementioned findings and conclusions.

1. Expand the physician profiling program to include those physicians who do not receive physician profiles.
2. Institute interventions aimed at increasing the rates for cholesterol screening, mammography screening, and hemoglobin A1c screening.
3. Educate rural physicians as to how they may use the physician profiles to improve practice patterns over time.
4. Increase office visits by plan medical directors to physicians who have aberrant utilization practice patterns.
5. Initiate studies directed at determining interventions that are most effective in improving physician utilization patterns.
6. Conduct future studies that would include random sampling, and analysis by specialty, when assessing the impact of physician profiling on utilization practice patterns.

Summary

This chapter discussed the findings, conclusions, and recommendations generated by this study. The next chapter, Chapter VI, will include the discussion, retrospectively observe the study, and discuss specific future directions for recommendations generated by this study.

CHAPTER VI

EPILOGUE

Introduction

The purpose of the study was to assess the impact of physician profiling and information sharing in modifying practice patterns among HMO primary care physicians in the state of Tennessee. To address this purpose comparisons were made to determine differences in practice patterns between primary care physicians who received profiles and primary care physicians who did not receive profiles, between urban physicians who received profiles and rural physicians who received profiles, and between urban physicians who did not receive profiles and rural physicians who did not receive profiles.

Discussion

Changing inappropriate utilization practice patterns is one way to control costs and improve the quality of health care (Balas, et al., 1996). Physician profiling has been used to this end in the study. It was shown that physician profiling is a helpful, but not sufficient, methodology as an incentive to change physician behavior toward improving the quality measures. However, it does appear to be effective in controlling certain utilization measures, thereby reducing the cost of care.

Direct cost measures were intentionally excluded from the profile, since the primary purpose of the profile was to improve quality through the reduction of inappropriate utilization. With the inclusion of cost measures physicians may have perceived the profile as an economic tool rather than a quality improvement tool.

Unlike the typical HMO that uses capitation to reimburse its primary care physicians, this Plan's HMO paid most primary care physicians on a fee for service basis. This limited any potential for under-utilization being reflected on measures in the study.

Although much maligned, claims data can be an important ingredient in a physician profile. While using claims data for quality improvement purposes is still controversial, a considerable amount of experience in using them for this purpose has accumulated (Coulam & Gaumer, 1997). This study has successfully used claims data in physician profiles to improve utilization patterns.

Observations of the Study

The study has shown that physician profiling can be effective in improving utilization patterns, but is not as effective in improving preventive screening rates. This was largely due to the fact that services like specialist visits and inpatient admissions were within the control of physicians, whereas preventive screening services in general were within the control of the patient. In order to effectively improve preventive screenings multiple interventions are warranted, not just physician profiling alone.

Although measurement is necessary to monitor improvement, physician profile (or report card) quality and utilization measurement alone will not improve care. What seems to really improve outcomes for specific patients is the incorporation of a productive provider and patient interaction into daily work (Von Korff, et al., 1997). Better aggregate care may result when there is ongoing measurement, monitoring processes of care, and a process for designing and testing changes in the care.

In the study it was noted that there was a significant reduction in emergency room utilization from 1998 to 1999. This was partly due to The Plan implementing an emergency room utilization reduction initiative in January 1999, after noticing an increasing trend over the previous years. The initiative targeted physicians whose members made frequent visits to the emergency room. The physicians were asked to increase their office hours, and educate members about the appropriate use of the emergency room in emergent versus non-emergent situations. A concern of this study was that there was little that could be done to separate the effects of the initiative and the effects of physician profiling within the scope of the study.

Another concern of the study was that the researcher could not determine whether the secondary diagnosis reported in the claims databases were comorbidities or complications. Comorbidities that are present at admission to the hospital are a necessary element of the risk adjustment process in profiling. However, complications should not be used for patient severity because they may be reflective of the quality of inpatient care. For some diagnoses, there should be no doubt about their genesis. A chronic illness such as hypertension, diabetes, or asthma is clearly a comorbidity. However, some of the diagnoses in the International Classification of Diseases, Ninth Edition, Clinical Modification (ICD-9-CM) range 996-999 range are clearly complications of care (Goldfield, 1999). In addition, the diagnoses used in this study rely heavily on the accurate reporting from physician's offices. This is a problem experienced by most researchers conducting physician profiling studies.

Fortunately, the Adjusted Clinical Group risk adjustment methodology (methodology used in this study) allows the option for researchers to not apply secondary diagnoses,

with the preface that some information may be lost. In this study the researcher chose to use only the primary diagnosis in risk adjustment to avoid some of the potential problems mentioned above. The fact that this study has used a risk adjustment methodology lends validity to the study.

A strength of this study was the ability to track physicians under both an experimental group and control group condition. Also lending to the ease of analysis for the study was the fact that the intervention occurred at the end of 1998, allowing full calendar years for classification of pre-intervention and post-intervention groups.

Future Directions

Better aggregate care and individual care is likely to result when there is also a balanced assessment of processes and outcomes that matter to patients, and education of and feedback to patients and physicians in real time about the quality of their interaction (Von Korff, 1997). The Plan in this study currently mails out physician profiles on a quarterly basis. Physicians in general cannot act on data that is more than three months old, and have expressed a need for more current or recent data that is actionable. A recommendation for the plan would be to move towards producing profiles on a monthly basis, using data from their authorization and precertification databases (more current) to complement the claims data which are associated with long lag times.

Rural physicians appeared to utilize far fewer services (seen from both utilization and quality measures) than urban physicians. This has the following implications: (i) there are possible access to care problems for patients living in rural areas, (ii) some rural patients may not believe in the utility of certain preventive services. Some ways for the

plan to address these issues are, for example, to encourage physicians to offer the patient education about the benefits about mammograms, and to use mobile mammography units to increase access. Holding day clinics for one day a week in remote areas, or encouraging physicians to offer home visits, can increase the number of primary care visits, for example.

Physician profiles are essentially comparative in nature, as very few norms or standards have been identified for any of the measurables on the profile (Goldfield, 1999). The profile analyzed in this study compares the performance of physicians against each other in a defined data set. However, to raise the standard of healthcare in the United States, physicians should be aware of benchmark data that would enable them to become best practices. The following are suggested benchmarks for some of the measures on the physician profile, and some recommendations for improving these measures.

Primary care visits :

In this study there are on average 1.2 primary care visits per member per year for the group of physicians receiving profiles. Kongstvedt (1996) suggests that in a well-managed HMO primary care visits should average around 2.5 visits per member per year. This increases interaction and communication between the physician and patient, with improved chances for prevention. Birthday card reminders to members or a call from the physician office to set up an appointment for annual checkups may be an effective measure in improving primary care visits.

Specialty care visits:

Kongstvedt (1996) also suggests that there should be on average 1.1 specialist visits per member per year. However, in an HMO it is encouraged to keep specialist visits to a

minimum and for the primary care physician to try and provide as much of the care as is needed. The group of physicians receiving profiles has on average 0.52 specialist visits per member per year. This was largely due to the strong referral management process that The Plan maintains.

Number of prescriptions per member per year:

The 1999 Novartis Pharmacy Benefit Report has shown that nationally the average number of prescriptions written per patient per year is around 8.2. For the group of physicians who received profiles, the average number of prescriptions written per patient per year is around 8.8. To control pharmacy utilization The Plan could develop a pharmacy report card for physicians to provide information about their prescription patterns. Other measures the plan could take is developing a pharmacy network, sending educative mailings to physicians informing them of the efficacy of certain drugs, and implementing pharmacy case management with physicians for patients with a high volume of pharmacy usage.

Proportion of generic prescriptions written:

Containing rising pharmacy costs is a challenge for every managed care organization in the country. For The Plan used in this study, pharmacy costs have now exceeded inpatient costs as a percentage of total cost of care. Encouraging physicians to prescribe generic products instead of brand name is one method that organizations use to try and control spiraling costs. Some managed care companies provide feedback to physicians about their prescribing patterns. There are no known benchmarks for proportion of generic prescriptions written.

Inpatient admissions:

Managed care organizations generally rely on pre-admission certification for medically necessary services, and good physician judgement, to control inpatient admissions. The Milliman & Robertson Health Care Cost Guidelines (1997) for well-managed health care plans suggest an average of 48 admissions per 1,000 members per year. The group of physician receiving profiles has an admissions per 1,000 members per year rate of 21.8, which is much below the Milliman & Robertson average. This is largely due to the strong pre-admission certification process of The Plan.

Emergency room visits:

Milliman & Robertson (1997) have also suggested that for well-managed plans ER visits/1000 members per year should be between 60 to 80. Any rate below 60 would be bordering on under-utilization. The group of physicians receiving profiles has an ER visits/1000 members per year rate of 66.8. Due to the ER utilization reduction initiative that was started by The Plan in 1999, the rate was drastically reduced from previous years.

Outpatient surgeries:

The Plan has a strong procedure Appropriateness Review process to determine medical necessity for high volume procedures. There are no known benchmarks for outpatient surgery visits.

Cholesterol screening, mammography screening, and hemoglobin A1c screening:

Based on Healthy People 2010 benchmarks, this is an area for most improvement. Healthy People 2010 states that: (i) the mammography screening rate should be at 70% (the physicians receiving profiles have a rate of 66%), (ii) the hemoglobin A1c screening

rate should be at 50% (the physicians receiving profiles have a rate of 57%), and (iii) the cholesterol screening rate should be at 80% (the physicians receiving profiles have a rate of 25%). Health plans have access to lists of members who are not compliant with certain preventive screenings. These lists need to be shared with physicians, and supplement the physician profiling process, in order for any effective change in preventive measures is to take place. Physicians may use these lists to send reminders to patients who are non-compliant, or use the list to identify non-compliant patients when they schedule annual checkups (Rimer, 1994). A physician recommendation is one of the most well documented facilitators to a woman obtaining a mammogram. One study found that "94% of women whose physicians had recommended mammograms had had one in the last two years, while only 36% of women whose physicians had not made the recommendation had had a mammogram" (Wong, 1997).

Whilst most measures on the profile address issues related to cost, the preventive screening measures relate directly to quality. As described above, there is a known value to improving the screening rates. Early prevention can avoid the incurral of costs, and utilization of other medical services, in the future.

Utility of the Study

The study shows that physician profiling does indeed aid in controlling physician utilization practice patterns. The profile can be used as a tool to encourage discussions between The Plan medical directors and physicians who appear to have inappropriate utilization patterns even after case-mix adjustment.

The Plan may utilize the information to form part of a physician's credentialing file, along with other information. This information can be used to select physicians who have best practices for special network privileges, e.g. elimination of the prior inpatient authorization and specialist referral approval processes for these physicians.

The results may also be used as part of the National Committee for Quality Accreditation (NCQA) criteria as an evaluation of the process for monitoring both over- and under-utilization of services. Measuring over- and under-utilization forms a critical part of the NCQA Utilization Management Standard 11 (UM-11), when plans go through the accreditation process.

In addition, the study has shown areas for improvement: access to care for rural members, improvement of rates for quality screening measures, and special areas of utilization on which to focus.

Future Studies

To eliminate the Hawthorne effect as an explanation for the improvement in utilization patterns for the group of physicians that received profiles, the study should be replicated and trended using data for an additional year, post-intervention.

Another area for future study is the development of episodes of care to measure consistency of treatment by physicians. Development of these episodes in the HMO environment can then be transferred for specialty specific use in the Preferred Provider Organization (PPO) environment.

To generalize the study further, it may be prudent to use physicians from other HMO plans within the state, who have more than 50 members assigned to them, and then

randomly select physicians into groups of physicians that receive profiles and those that do not.

Summary

It is the researcher's belief that physicians must be responsible for improving the quality of care that they provide, because they are the ones in the position to do so. The researcher also believes that employer organizations, government, and the public should attempt to convince physicians that altering processes of care will actually lead to better outcomes. This can be done by striving to make sure that process measures are part of the data system used for case-mix adjustment and physician profiling whenever possible. The other effective mechanism is to advocate for the public release of physician data.

We are still in the infancy of the development of physician profiles and the dissemination of these to the public. There is much room for improvement with respect to the statistical methods, the documentation of methods and data restrictions in public reports, and the careful statement of caveats associated with public releases.

This chapter discussed the observations of the study. It also reviewed some benchmarks and process applications based on the findings, conclusions, and recommendations of the study. Future directions in the way of further research were also discussed.

REFERENCES

Adjusted Clinical Groups. The Johns Hopkins University.

Avorn J, Soumerai SB, et al. A randomized trial of a program to reduce the use of psychoactive drugs in nursing homes. N Engl J Med. 1992; 327: 168-73.

Balas EA, Boren SA, Brown DB, Ewigman, GB, Mitchell, JA, Perkoff GT. Effect of Physician Profiling on Utilization: Meta-analysis of Randomized Clinical Trials. JGIM. 1996; 11: 584-90.

Bernstein, AB. Ready or not, here it comes: medical practice in the new millenium. Seminars in Medical Practice, 1998; 1(1): 2-6.

Berwick DL, Coltin KL. Feedback reduces test use in a health maintenance organization. JAMA, 1986; 255: 1450-4.

Carey, RG, Lloyd, RC. Measuring Quality Improvement in Healthcare. New York, NY: Quality Resources, 1995.

Clinical Resource Management System. Patterns Profiler. HPR Inc., 1997.

Clinical Resource Management System. Quality Profiler. HPR Inc., 1997.

Coleman, MT, Endsley, S. Quality Improvement: First Steps. Fam Pract Mgmt, 1999; 6(3): 23-26.

Coulam, RF, Gaumer, GL. Medicare's Prospective Payment System: A Critical Appraisal. Health Care Financing Review annual suppl. 13 (1997): 45-76.

Emmons DW, Wozniak GD. Profiles and feedback: who measures physician performance. Socioeconomic Characteristics of Medical Practice, 1994. Chicago, Ill: American Medical Association; 1994.

Enthoven, AC, Kronick, R. Universal Health Insurance Through Incentives Reform. The Nation's Health. Sudbury, MA: Jones and Bartlett Publishers, 1997.

Epstein, AM. Changing Physician Behavior, Increasing Challenges for the 1990's. Archives of Internal Medicine, 1991; 151:2147.

Evans JR, Lindsay WM. The Management and Control of Quality. 3rd ed. St. Paul, MN: West Publishing Company, 1996.

Evans, JH, Hwang, Y, Nagarajan, N. Physicians' Response to Length of Stay Profiling. Medical Care, 1995; 33(11): 1106-1119.

Frazier LM, Brown JT, Divine GW, et al. Can physician education lower the cost of prescription drugs? A prospective controlled trial. Ann Intern Med, 1991; 115: 116-21.

Gerard, JC, Arnold, FL. Performance Improvement with a Hybrid FOCUS-PDCA Methodology. Journal on Quality Improvement, 1995; 22(10): 660-672.

Goldfield, N. Physician Profiling and Risk Adjustment. Gaithersburg, MD: Aspen Publishers, 1999.

Hennekens, CH, Buring, JE. Principles of Epidemiology. Boston, MA: Little, Brown and Company, 1987.

http://www.health.gov/healthypeople

Hurley, RE, et al. Gatekeeping effects on patterns of physician use. Journal of Family Practice, 1991; 32: 167-174.

Johnson CC, Martin M, Epstein SM, Lee JD. The effect of a physician education program on hospital length of stay and total patient charges. JSC Med Assoc, 1993; 89(6): 293-301.

Johnson CC, Martin M. Effectiveness of a physician education program in reducing consumption of hospital resources in elective total hip replacement. South Med J, 1996; 89(3):282-9.

Joint Commision on Accreditation of Healthcare Organizations: Designing and implementing an outcomes-based performance improvement project: beahvioral health care. Jt Comm J Qual, 1998; 24:435-453.

Joint Commission on Accreditation of Healthcare Organizations. Designing and Implementing an Outcomes-Based Performance Improvement Project: Behavioral Health Care. Journal on Quality Improvement, 1996; 24(8): 435-453.

Keller RB, Chapin AM, Soule DN. Informed inquiry into practice variations: The Maine Medical Assessment Foundation. Qual Assur Health Care. 1990; 2: 69-75.

Kongstvedt, PR. The Managed Health Care Handbook. 3rd ed. Gaithersburg, MD: Aspen Publications, 1996.

Lassen LC, Kristensen FB. Peer comparison feedback to achieve rational and economic drug therapy in general practice: a controlled intervention study. Scan J Prim Health Care, 1992; 10: 76-80.

Lighter, DE. Physician Profiling Evaluation. 1998.

Marton KI, Tul V, Sox HC. Modifying test-ordering behavior in the outpatient medical clinic: a controlled trial of two educational interventions. Arch Intern Med, 1985; 145: 816-25.

McPhee SJ, Bird JA, Jenkins C, Fordham D. Promoting cancer screening: a randomized control trial of three interventions. Arch Intern Med, 1989; 149: 1866-72.

Metfessel BA. Specialist profiling using claims and administrative databases. American Journal of Integrated Health Care, 1998; 1(4): 168-74.

Milliman & Robertson. Healthcare Management Guidelines, 1997.

Monfils, MK. Needs Assessment and Implementation of an Employee Assistance Program. AAOHN Journal, 1995; 43(5): 263-269.

Naidoo, A, McSharry, PF. Quality Improvement: A Practical Approach for Physicians Using the PDSA Cycle and Physician Profiling. Tennessee Medicine, 1999; 92(8): 293-296.

National Health Information, LLC. Data, MD collaboration yield elusive support for disease management. Physician Profiling and Behavior Change Report, 1999; 2(1): 1-7.

National Health Information, LLC. Educate physicians, don't coerce them to implement successful guideline program. Physician Profiling and Behavior Change Report, 1999; 2(12): 177-192.

National Health Information, LLC. Physician Compensation & Incentives Under Capitation. Atlanta, GA: National Health Information, LLC; 1998.

National Health Information, LLC. Physicians give clinical profiling program 'thumbs up'. Physician Profiling and Behavior Change Report, 1998; 1(1): 1-4.

Novartis. Pharmacy Benefit Report: Facts & Figures, 1999.

Office of Management and Budget. Federal Register, 1997. 55 FR 12154-12160.

Paradise, JL. Controversies: Treatment of Acute Otitis Media. JAMA, 1998; 279: 1783-1784.

Polit, DF, Hungler, BP. Nursing Research: Principles and Methods. 6th ed. Philadelphia, PA: Lippincott Publications, 1999.

Pugh JA, Frazier LM, et al. Effect of a daily charge feedback on inpatient charges and physician knowledge and behavior. Arch Intern Med, 1989; 149: 426-9.

Rimer, BK. Mammography use in the US: trends and the impact of interventions. Ann Behav Med, 1994; 16: 317-326.

SAS/STAT User's Guide V6. SAS Institute Inc., 1994.

Schwartz, M, Landis, SE, Rowe, JE. A team approach to quality improvement. Fam Pract Mgmt, 1999; 6(4):25-30.

Seidman JJ, Greenfield S. Quality for good measure. Seminars in Medical Practice, 1998; 1(1): 7-10.

Selker, HP. Criteria for Adoption in Practice of Medical Practice Guidelines. American Journal of Cardiology, 1993; 71: 339-341.

Spoeri, RK, Ullman, R. Measuring and Reporting Managed Care Performance: Lessons Learned and New Initiatives. Annals of Internal Medicine, 1997; 127(8): 726-32.

Steele MA, Bess DT, Franse VL, Graber SE. Cost effectiveness of two interventions for reducing outpatient prescribing costs. DICP Ann Pharmacother, 1989; 23: 497-500.

Tierney WM, Miller ME, McDonald CJ. The effect on test ordering of informing physicians of the charges for outpatient diagnostic tests. N Engl J Med, 1990; 322(21): 1499-504.

US Department of Health and Human Services. Healthy People 2010 Objectives, 1999. Washington, DC: US Department of Health and Human Services, Public Health Service.

Von Korff, M, et al. Collaborative Management of Chronic Illness. Annals of Internal Medicine, 1997; 127: 1097-1102.

Weiner, JP, et al. Development and application of a population oriented measure of ambulatory care case mix. Medical Care, 1991; 29: 452-472.

Welch GH, Miller ME, Welch WP. Physician Profiling: An analysis of Inpatient Practice Patterns in Florida and Oregon. N Engl J Med, 1994; 330(9): 607-12.

Winickoff, RN, Coltin, KL, et al. Improving physician performance through peer comparison feedback. Med Care, 1984; 22(6): 527-34.

Wong, FL. Manual of Intervention Strategies to Increase Mammography Rates. Centers for Disease Control and Prevention, 1997.

APPENDICES

APPENDIX A

SAMPLE PHYSICIAN PROFILE PACKET

February 7, 2000

«FName» «LName», «Degree»
«ADDLN1»
«AddLn2»
«City», «State» «ZipCode»

Dear Dr. «LName»:

Enclosed please find your Primary Care Physician (PCP) Profile for our Commercial HMO Products. The information reflects data from July 1999 through September 1999 (Third Quarter).

The Plan's physician profiles are intended to show physicians where they stand relative to their peers on certain key indicators of performance. The key indicators listed in our Managed Care Performance profile are areas of common interest to The Plan and participating physicians and are critical to our mutual success. For an accurate reflection of your patient population, the Managed Care Performance profile is separated into two categories: Adults and Children.

The data have been case mix adjusted using Adjusted Clinical Groupings (ACGs). ACGs were developed by Johns Hopkins University and is widely used throughout the United States.

Also, a User's Guide has been included to explain the report in greater detail. Profiles are distributed for PCPs every quarter, with a three-month claims lag (allows three months after the quarter ends for claims to be submitted for services provided during the quarter) applied to the reporting period. For example, if the reporting period is for the first quarter, the profile will be distributed in the third quarter.

It is very important to us that we provide you with a profile that is easy to understand, relevant and educational. We have enclosed a brief survey for any comments you may have about the profiling package. Please take a few minutes to complete the survey and return to the specified address or fax
us at (111) 555-5555 if you prefer.

We hope you find this information helpful and look forward to receiving your comments. If you have questions or comments, please contact your [REGIONAL MEDICAL DIRECTOR] by calling (333) 555-5555 between the hours of 8:00 a.m. and 5:00 p.m. Thank you for your time and feedback!

Sincerely,

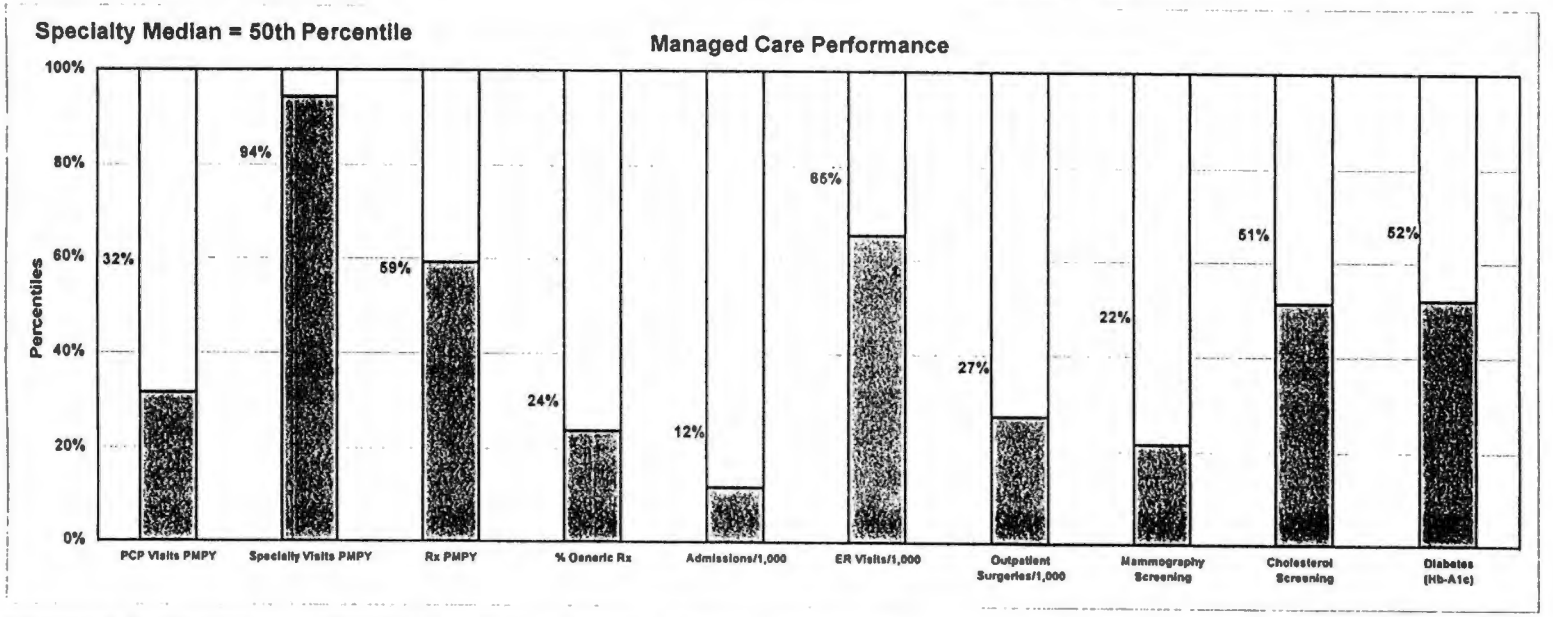
Chief Medical Officer
Managed Care

CONFIDENTIAL

Corporate Profile - Commercial HMO
Quarterly Snapshot

July 1999 - September 1999

Provider Name:	JOHN SMITH MD
Provider Number:	123456
Specialty:	Family Practice
Average Panel Size Per Month:	109
Specialty Average Panel Size Per Month:	55



Services/Member	**Primary Care Visits PMPY	**Specialty Visits PMPY	**Rx/PMPY	**% Generic Rx	**Admissions /1,000	**ER Visits /1,000	**Outpatient Surgeries /1,000	Mammography Screening	Cholesterol Screening	Diabetes (Hb-A1c)
Your Members	0.68	1.36	8.88	34.2%	4.78	112.85	218.04	33.3%	20.0%	66.7%
Specialty Median	0.88	0.42	8.15	42.2%	18.59	68.35	361.40	66.7%	20.0%	66.7%

**The measure is Case-Mix adjusted
N/A - No members or events available for the measure

Medical Information Management
February 7, 2000

36563B654321

USER'S GUIDE

for

PRIMARY CARE PHYSICIAN PROFILE

Produced By:

The Plan

Primary Care Physician (PCP) Profile

The purpose of this profile is to provide physicians with information about their utilization and quality of care as it compares to the physician's peer group within The Plan.

Methodology

Based on physician feedback, The Plan designed a corporate physician profiling format. The layout of the profile is intended to describe the physician's member panel and illustrate the physician's practice in terms of access, utilization, and quality of care.

The profile measures are based on data obtained from claim, membership and provider information systems.

To make comparisons between PCPs more meaningful, the Utilization Services measures for the commercial HMO products are age, gender, and Case-Mix adjusted using Adjusted Clinical Groupings (ACG). The ACGs were developed by Johns Hopkins University and is widely used throughout the healthcare industry.

The Plan treats PCPs as health care managers. Therefore, the profiles reflect ALL medical services incurred by the physician's member panel.

The PCP's data is presented in a spreadsheet format. Relative standing of the PCP within his/her specialty is then graphically displayed in a bar chart format using percentiles. A percentile indicates the percentage of population at or below a specific score. For example, if a PCP's percentile rank for Admissions/1,000 is at 80%, this means that 80% of PCPs within the same specialty have the same or lower Admissions/1000 rate.

Provider Information

This portion of the profile describes the provider's specialty and the average number of members enrolled during the reporting period.

- **Average Panel Size Per Month** – average number of members enrolled with the PCP during the reporting period.
- **Specialty Average Panel Size Per Month** – average number of members enrolled with PCPs within the specialty during the reporting period.

Managed Care Performance

Adults/Children

The Adults and Children profiles describe the physician's performance as it pertains to the adult and children members of his/her panel, respectively. All age criteria are based upon the member's age at the end of the reporting period.

Distribution

A profile for Adults and Children will be produced quarterly and distributed to PCPs according to his/her specialty.

- **Family Practice**
- **General Practice**
- **Internal Medicine**
- **Pediatrics**

Measures

The measures indicated in the PCP profile reflect three categories: Access (to care), Utilization, and Quality of Care.

Access

Data from members' claims were included in these measures if the service occurred during the reporting period and member was assigned to that PCP at the end of the reporting period.

- **Primary Care Visits PMPY** – average number of office visits to the PCP per member per year.
- **Specialty Visits PMPY** – average number of office visits to a specialist per member per year.

Utilization Services

Data from members' claims were included in these measures if the service occurred while the member was enrolled with that PCP within the reporting period.

- **RX PMPY** – average number of prescriptions per member per year.
- **% Generic Rx** - percentage of prescriptions that are generic.
- **Admissions/1,000** – number of acute inpatient admissions per 1,000 members per year.
- **ER Visits/1,000** – number of emergency room visits per 1,000 members per year.
- **Outpatient Surgeries/1,000** – number of outpatient surgical procedures per 1,000 members per year.

Quality of Care

Data from members' claims were included in these measures if the member was enrolled with that PCP at the end of the reporting period.

Adults

- **Mammography Screening** – ([1] HEDIS guidelines are applied.) percentage of women age 52 through 69 years who had a mammogram during the 24 months preceding the end of the reporting period.
- **Cholesterol Screening** – percentage of members that should have had a cholesterol screening that have actually been screened.
- **Diabetes (Hb-A_1c)** – [2]percentage of diabetic members identified as having had their Hb-A1c measured during the past twelve months.

1 Health Employer Data Information Set (HEDIS® 3.0) is a set of standardized measures designed to provide purchasers and consumers with information needed to reliably compare the performance of managed health care plans. It is sponsored, supported and maintained by the National Committee for Quality Assurance (NCQA) – a non-profit organization committed to evaluating and publicly reporting on the quality of managed care plans. HEDIS guidelines for enrollment and eligibility will be followed in order to maintain comparability among providers being profiled.

2 For Commercial HMO products only.

Let Us Know What You Think...

We are always looking for ways to improve our service to you. Please take a moment to answer the following questions about this informational packet.

Title of person responding: Primary Care Physician Specialist Office Personnel

On a scale of one to five, please evaluate your satisfaction of your *Primary Care Physician profile*:

	Very Dissatisfied	Dissatisfied	Neutral	Satisfied	Very Satisfied
1.Usefulness of information	1	2	3	4	5

2. Please explain anything in your profile that you think is not accurate.

_____
_____
_____

3. Will you use the information provided on the profile? Yes No

If so, how?_____
_____
_____

If not, why?_____
_____
_____

4. What additional information would be helpful for you to evaluate your patient panel?

_____
_____
_____

On a scale of one to five, please evaluate your satisfaction of the *Primary Care Physician (PCP) Profile User's Guide:*

	Very Dissatisfied	Dissatisfied	Neutral	Satisfied	Very Satisfied
5.Usefulness	1	2	3	4	5
6.Clarity	1	2	3	4	5
7.Content	1	2	3	4	5

8. What suggestions do you have to make the User's Guide more beneficial to you?

_____
_____
_____
_____

APPENDIX B

EXAMPLES OF STATISTICAL ANALYSIS OUTPUT

N P A R 1 W A Y P R O C E D U R E

Wilcoxon Scores (Rank Sums) for Variable Z98PCPVI
Classified by Variable PROFILE

PROFILE	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
N	1213	912824.0	1017100.50	8836.03084	752.53421
Y	463	492502.0	388225.50	8836.03084	1063.71922

Average Scores Were Used for Ties

Wilcoxon 2-Sample Test (Normal Approximation)
(with Continuity Correction of .5)

S = 492502 Z = 11.8012 Prob > |Z| = 0.0001

T-Test Approx. Significance = 0.0001

Kruskal-Wallis Test (Chi-Square Approximation)
CHISQ = 139.27 DF = 1 Prob > CHISQ = 0.0001

N P A R 1 W A Y P R O C E D U R E

Wilcoxon Scores (Rank Sums) for Variable Z98SPCVI
Classified by Variable PROFILE

PROFILE	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
N	1213	917842.0	1017100.50	8840.91338	756.67106
Y	463	487484.0	388225.50	8840.91338	1052.88121

Average Scores Were Used for Ties

Wilcoxon 2-Sample Test (Normal Approximation)
(with Continuity Correction of .5)

S = 487484 Z = 11.2271 Prob > |Z| = 0.0001

T-Test Approx. Significance = 0.0001

Kruskal-Wallis Test (Chi-Square Approximation)
CHISQ = 126.05 DF = 1 Prob > CHISQ = 0.0001

N P A R 1 W A Y P R O C E D U R E

Wilcoxon Scores (Rank Sums) for Variable Z98RX
Classified by Variable PROFILE

PROFILE	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
N	1213	897112.0	1017100.50	8859.09039	739.58120
Y	463	508214.0	388225.50	8859.09039	1097.65443

Average Scores Were Used for Ties

Wilcoxon 2-Sample Test (Normal Approximation)
(with Continuity Correction of .5)

S = 508214 Z = 13.5441 Prob > |Z| = 0.0001

T-Test Approx. Significance = 0.0001

Kruskal-Wallis Test (Chi-Square Approximation)
CHISQ = 183.44 DF = 1 Prob > CHISQ = 0.0001

N P A R 1 W A Y P R O C E D U R E

Wilcoxon Scores (Rank Sums) for Variable Z98ADMIT
Classified by Variable PROFILE

PROFILE	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
N	1213	852477.500	1017100.50	8265.59464	702.78442
Y	463	552848.500	388225.50	8265.59464	1194.05724

Average Scores Were Used for Ties

Wilcoxon 2-Sample Test (Normal Approximation)
(with Continuity Correction of .5)

S = 552849 Z = 19.9166 Prob > |Z| = 0.0001

T-Test Approx. Significance = 0.0001

Kruskal-Wallis Test (Chi-Square Approximation)
CHISQ = 396.67 DF = 1 Prob > CHISQ = 0.0001

N P A R 1 W A Y P R O C E D U R E

Wilcoxon Scores (Rank Sums) for Variable Z98ERVIS
Classified by Variable PROFILE

PROFILE	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
N	1213	947565.0	1017100.50	8782.53003	781.174773
Y	463	457761.0	388225.50	8782.53003	988.684665

Average Scores Were Used for Ties

Wilcoxon 2-Sample Test (Normal Approximation)
(with Continuity Correction of .5)

S = 457761 Z = 7.91742 Prob > |Z| = 0.0001

T-Test Approx. Significance = 0.0001

Kruskal-Wallis Test (Chi-Square Approximation)
CHISQ = 62.686 DF = 1 Prob > CHISQ = 0.0001

N P A R 1 W A Y P R O C E D U R E

Wilcoxon Scores (Rank Sums) for Variable Z98OSURG
Classified by Variable PROFILE

PROFILE	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
N	1213	916721.0	1017100.50	8815.60576	755.74691
Y	463	488605.0	388225.50	8815.60576	1055.30238

Average Scores Were Used for Ties

Wilcoxon 2-Sample Test (Normal Approximation)
(with Continuity Correction of .5)

S = 488605 Z = 11.3865 Prob > |Z| = 0.0001

T-Test Approx. Significance = 0.0001

Kruskal-Wallis Test (Chi-Square Approximation)
CHISQ = 129.65 DF = 1 Prob > CHISQ = 0.0001

TABLE OF PRECHOL1 BY CHOL1

PRECHOL1(PRECHOL1) CHOL1(CHOL1)

Frequency Percent Row Pct Col Pct	N	Y	Total
C98CHOL	6431	2343	8774
	12.35	4.50	16.85
	73.30	26.70	
	16.39	18.27	
E98Chol	32808	10483	43291
	63.01	20.13	83.15
	75.78	24.22	
	83.61	81.73	
Total	39239	12826	52065
	75.37	24.63	100.00

Frequency Missing = 2211

STATISTICS FOR TABLE OF PRECHOL1 BY CHOL1

Statistic	DF	Value	Prob
Chi-Square	1	24.338	0.001
Likelihood Ratio Chi-Square	1	23.990	0.001
Continuity Adj. Chi-Square	1	24.204	0.001
Mantel-Haenszel Chi-Square	1	24.337	0.001
Fisher's Exact Test (Left)			5.15E-07
(Right)			1.000
(2-Tail)			9.98E-07
Phi Coefficient		-0.022	
Contingency Coefficient		0.022	
Cramer's V		-0.022	

Effective Sample Size = 52065
Frequency Missing = 2211

TABLE OF PREMAM1 BY MAM1

PREMAM1(PREMAM1) MAM1(MAM1)

Frequency Percent Row Pct Col Pct	N	Y	Total
C98MAM	367	742	1109
	6.25	12.64	18.89
	33.09	66.91	
	19.06	18.80	
E98MAM	1558	3204	4762
	26.54	54.57	81.11
	32.72	67.28	
	80.94	81.20	
Total	1925	3946	5871
	32.79	67.21	100.00

Frequency Missing = 48405

STATISTICS FOR TABLE OF PREMAM1 BY MAM1

Statistic	DF	Value	Prob
Chi-Square	1	0.058	0.810
Likelihood Ratio Chi-Square	1	0.057	0.811
Continuity Adj. Chi-Square	1	0.042	0.838
Mantel-Haenszel Chi-Square	1	0.058	0.810
Fisher's Exact Test (Left)			0.609
(Right)			0.418
(2-Tail)			0.831
Phi Coefficient		0.003	
Contingency Coefficient		0.003	
Cramer's V		0.003	

Effective Sample Size = 5871
Frequency Missing = 48405
WARNING: 89% of the data are missing.

Pre-intervention: Experimental Group vs Control Group - Diabetic Screening
15:41 Monday, April 3, 2000

TABLE OF PREDIAB1 BY DIAB1

PREDIAB1(PREDIAB1) DIAB1(DIAB1)

Frequency Percent Row Pct Col Pct	N	Y	Total
C98DIAB	274	497	771
	7.01	12.72	19.74
	35.54	64.46	
	16.94	21.71	
E98DIAB	1343	1792	3135
	34.38	45.88	80.26
	42.84	57.16	
	83.06	78.29	
Total	1617	2289	3906
	41.40	58.60	100.00

Frequency Missing = 50370

STATISTICS FOR TABLE OF PREDIAB1 BY DIAB1

Statistic	DF	Value	Prob
Chi-Square	1	13.595	0.001
Likelihood Ratio Chi-Square	1	13.768	0.001
Continuity Adj. Chi-Square	1	13.296	0.001
Mantel-Haenszel Chi-Square	1	13.592	0.001
Fisher's Exact Test (Left)			1.22E-04
(Right)			1.000
(2-Tail)			2.37E-04
Phi Coefficient		-0.059	
Contingency Coefficient		0.059	
Cramer's V		-0.059	

Effective Sample Size = 3906
Frequency Missing = 50370
WARNING: 93% of the data are missing.

Univariate Procedure

Variable=PCPDIFF

Moments

N	453	Sum Wgts	453
Mean	0.243783	Sum	110.4337
Std Dev	0.719421	Variance	0.517567
Skewness	-0.08521	Kurtosis	0.259446
USS	260.8622	CSS	233.9403
CV	295.1073	Std Mean	0.033801
T:Mean=0	7.212222	Pr>|T|	0.0001
Num ^= 0	453	Num > 0	286
M(Sign)	59.5	Pr>=|M|	0.0001
Sgn Rank	19596.5	Pr>=|S|	0.0001

Quantiles(Def=5)

100% Max	2.626005	99%	1.974141
75% Q3	0.688703	95%	1.383999
50% Med	0.261806	90%	1.06627
25% Q1	-0.19953	10%	-0.74643
0% Min	-1.96976	5%	-1.02003
		1%	-1.53696
Range	4.595764		
Q3-Q1	0.888229		
Mode	-1.96976		

Extremes

Lowest	Obs	Highest	Obs
-1.96976(	400)	1.974141(	217)
-1.80389(	227)	1.99875(	15)
-1.64803(	226)	2.024096(	72)
-1.54453(	403)	2.110948(	165)
-1.53696(	448)	2.626005(	100)

Missing Value	.
Count	10
% Count/Nobs	2.16

Univariate Procedure

Variable=SPCDIFF

Moments

N	453	Sum Wgts	453
Mean	-0.04131	Sum	-18.7135
Std Dev	0.239299	Variance	0.057264
Skewness	0.167648	Kurtosis	4.649882
USS	26.6563	CSS	25.88324
CV	-579.274	Std Mean	0.011243
T:Mean=0	-3.67422	Pr>|T|	0.0003
Num ^= 0	453	Num > 0	176
M(Sign)	-50.5	Pr>=|M|	0.0001
Sgn Rank	-12770.5	Pr>=|S|	0.0001

Quantiles(Def=5)

100% Max	1.309122	99%	0.667848
75% Q3	0.078704	95%	0.359
50% Med	-0.04209	90%	0.215137
25% Q1	-0.15672	10%	-0.29824
0% Min	-1.29813	5%	-0.37686
		1%	-0.67586
Range	2.607253		
Q3-Q1	0.235421		
Mode	-1.29813		

Extremes

Lowest	Obs	Highest	Obs
-1.29813(	424)	0.667848(	158)
-0.84662(	425)	0.678072(	242)
-0.80082(	303)	0.786642(	144)
-0.70389(	299)	0.819455(	232)
-0.67586(	19)	1.309122(	219)

Missing Value	.
Count	10
% Count/Nobs	2.16

Univariate Procedure

Variable=RXDIFF

Moments

N	453	Sum Wgts	453
Mean	0.466436	Sum	211.2955
Std Dev	4.135617	Variance	17.10333
Skewness	-1.69408	Kurtosis	6.003168
USS	7829.259	CSS	7730.704
CV	886.6418	Std Mean	0.194308
T:Mean=0	2.400495	Pr>|T|	0.0168
Num ^= 0	453	Num > 0	322
M(Sign)	95.5	Pr>=|M|	0.0001
Sgn Rank	16056.5	Pr>=|S|	0.0001

Quantiles(Def=5)

100% Max	11.12075	99%	7.804466
75% Q3	2.68788	95%	5.7171
50% Med	1.170899	90%	4.522729
25% Q1	-0.60276	10%	-4.92324
0% Min	-25.0335	5%	-7.30734
		1%	-13.8408
Range	36.15422		
Q3-Q1	3.29064		
Mode	-25.0335		

Extremes

Lowest	Obs	Highest	Obs
-25.0335(	187)	7.804466(	430)
-22.8819(	5)	7.812804(	404)
-14.9924(	197)	8.78368(	1)
-13.8699(	2)	10.6776(	456)
-13.8408(	298)	11.12075(	131)

Missing Value	.
Count	10
% Count/Nobs	2.16

Univariate Procedure

Variable=ADMDIFF

Moments

N	453	Sum Wgts	453
Mean	-7.69887	Sum	-3487.59
Std Dev	66.15767	Variance	4376.837
Skewness	2.258767	Kurtosis	84.3476
USS	2005181	CSS	1978330
CV	-859.317	Std Mean	3.108359
T:Mean=0	-2.47683	Pr>|T|	0.0136
Num ^= 0	434	Num > 0	158
M(Sign)	-59	Pr>=|M|	0.0001
Sgn Rank	-14177.5	Pr>=|S|	0.0001

Quantiles(Def=5)

100% Max	844.4637	99%	129.543
75% Q3	4.627804	95%	46.74884
50% Med	-3.03016	90%	26.13616
25% Q1	-20.7155	10%	-52.3691
0% Min	-675.558	5%	-79.5028
		1%	-136.826
Range	1520.022		
Q3-Q1	25.34334		
Mode	0		

Extremes

Lowest	Obs	Highest	Obs
-675.558(	458)	129.543(	346)
-331.959(	444)	141.6138(	271)
-180.329(	212)	153.6299(	87)
-144.364(	246)	219.2427(	158)
-136.826(	43)	844.4637(	86)

Missing Value	.
Count	10
% Count/Nobs	2.16

Univariate Procedure

Variable=ERDIFF

Moments

N	453	Sum Wgts	453
Mean	-13.7222	Sum	-6216.14
Std Dev	60.07109	Variance	3608.536
Skewness	0.62677	Kurtosis	4.722561
USS	1716357	CSS	1631058
CV	-437.767	Std Mean	2.822386
T:Mean=0	-4.8619	Pr>|T|	0.0001
Num ^= 0	453	Num > 0	164
M(Sign)	-62.5	Pr>=|M|	0.0001
Sgn Rank	-16214.5	Pr>=|S|	0.0001

Quantiles(Def=5)

100% Max	371.444	99%	159.7151
75% Q3	20.79252	95%	86.78652
50% Med	-15.6818	90%	53.64242
25% Q1	-48.0351	10%	-80.2305
0% Min	-284.122	5%	-98.3459
		1%	-150.519
Range	655.5659		
Q3-Q1	68.82761		
Mode	-284.122		

Extremes

Lowest	Obs	Highest	Obs
-284.122(	274)	159.7151(	449)
-196.976(	375)	174.3625(	165)
-164.213(	201)	184.654(	167)
-161.153(	272)	195.5452(	331)
-150.519(	268)	371.444(	23)

Missing Value	.
Count	10
% Count/Nobs	2.16

Univariate Procedure

Variable=OSGDIFF

Moments

N	453	Sum Wgts	453
Mean	-3.2603	Sum	-1476.91
Std Dev	213.8788	Variance	45744.13
Skewness	0.171364	Kurtosis	4.686966
USS	20681161	CSS	20676346
CV	-6560.1	Std Mean	10.0489
T:Mean=0	-0.32444	Pr>|T|	0.7458
Num ^= 0	453	Num > 0	224
M(Sign)	-2.5	Pr>=|M|	0.8510
Sgn Rank	-552.5	Pr>=|S|	0.8432

Quantiles(Def=5)

100% Max	1162.925	99%	492.2278
75% Q3	106.801	95%	299.8956
50% Med	-1.92968	90%	219.5444
25% Q1	-95.5754	10%	-244.054
0% Min	-946.539	5%	-336.657
		1%	-613.34
Range	2109.464		
Q3-Q1	202.3764		
Mode	-946.539		

Extremes

Lowest	Obs	Highest	Obs
-946.539(	12)	492.2278(	251)
-814.763(	344)	597.963(	18)
-746.249(	276)	945.4354(	15)
-685.142(	10)	969.1185(	33)
-613.34(	229)	1162.925(	463)

Missing Value	.
Count	10
% Count/Nobs	2.16

VITA

Umagasen (Allen) Naidoo was born May 17, 1965 in Durban, South Africa. He attended public schools there and graduated from the Shallcross High School in 1982. After working as a kardex clerk and statistical analyst for a few years he completed a Bachelor of Science Degree in Mathematics at the University of Natal in 1989. He moved to Cape Town (South Africa) in 1990 to work as an actuarial analyst, and to start taking the Society of Actuaries associateship examinations. He passed five of the associateship examinations in total. In 1992 he came to the United States to study for a Master of Science degree in Statistics at Miami University (Ohio), and graduated in 1994. He has held various positions in the areas of statistical research, actuarial analysis, survey research, statistics education, and managed care. He is currently a Senior Information Project Manager with a managed care company in Tennessee. He received academic honors from the national honorary mathematics society of Pi Mu Epsilon and is currently program chair-nominate for the Health Policy Statistics Section of the American Statistical Association. Mr. Naidoo expects to receive the Doctor of Philosophy degree in Community Health in August 2000.