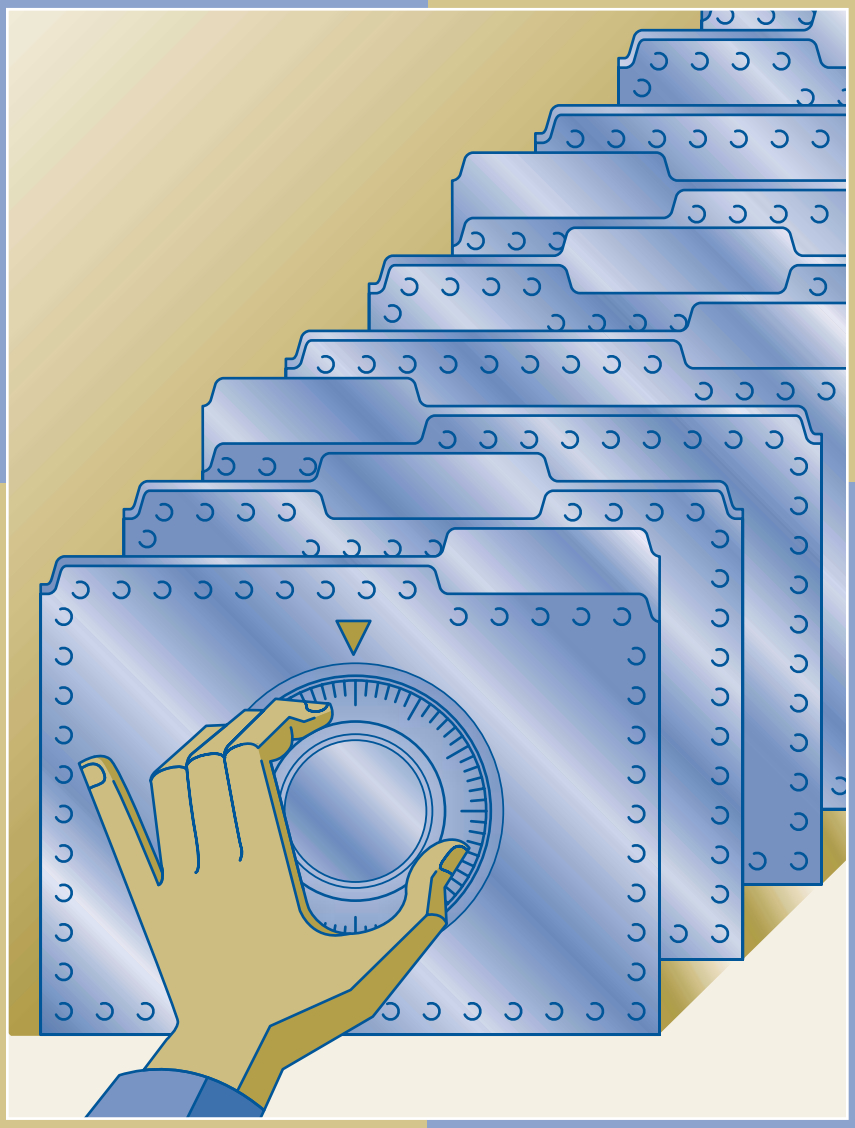
HEALTH PLAN DATA

A Rich Resource



Ripe with Challenges

Produced for
the Robert Wood Johnson Foundation's
Changes in Health Care Financing
and Organization (HCFO) Program
Written and Published by The Academy for
Health Services Research and Health Policy

HEALTH PLAN DATA: A RICH RESOURCE RIPE WITH CHALLENGES

TABLE OF CONTENTS

Executive Summary 2

Introduction 3

Making the Best Use of Available Data 4

Recognizing Wide Range of Information Plan Data Sets Offer 4

Learning Multi-step Approach Required to Work Well with Plan Data 5

Understanding Plan Data Limitations and Potential: Appropriate Models Critical 6

Motivating Plan Cooperation 7

Dealing with Inevitable Technical Challenges Built Into Plan Data 10

Creatively Addressing Generalizability Problem 11

Final Thoughts 11

FOREWORD

With growing numbers of people receiving their health coverage and care through entities known as health plans, understanding the way these plans work and how they effect the quality and availability of care has also grown in importance. At the same time, the plans themselves remain the sole sources of data for answering many of the questions researchers and policymakers need and want to explore. Researchers are, therefore, having to learn to work with plans, which operate differently as data sources than do such administrative data sources as the Health Care Financing Administration (HCFA). In working with plan data, researchers have discovered an array of challenges, ranging from convincing the plans to cooperate to addressing the difficulties of generalizing findings based on a single plan's data. Despite the limitations and complications, most researchers remain convinced that plan data is an irreplaceable resource. The need, therefore, is to work through the challenges of developing research topics and structuring projects to meet the needs of both researchers and plans.

To help researchers sort through the variety of issues involved, the Changes in Health Care Financing and Organization (HCFO) program held a small invitational meeting in July 1999. This report seeks to follow up on the issues and ideas discussed at the meeting in a way that supports researchers' efforts to make the fullest possible use of this resource. We hope that it will provide valuable perspective and practical advice for handling the inevitable challenges that accompany a commitment to using health plan data to gain better understanding of our ever-evolving health care system.

Anne K. Gauthier
Program Director

EXECUTIVE SUMMARY

Recent growth in managed care has meant that health plans have become the primary financial and administrative structures through which most patients, providers, and payers interact. Many health policy researchers have, therefore, found it necessary to turn to the various data sets that are developed, maintained, and used by the plans themselves in order to examine and understand health plans and their impact on the system. Although such data — referred to as plan data — offer rich research potential, they are usually collected for the plans' business purposes and do not exactly match the needs and issues of researchers. In addition, plans are not always comparable in the benefits they provide, the patients they treat, the way they treat them, or the costs they incur, making the use of plan data for broad-based policy research and analysis difficult.

Overcoming the challenges inherent in using plan data and drawing on this potential resource require that researchers fully understand plan data's limitations and build project plans and proposals that recognize those limitations. Researchers must also take into account the plans' interests, which will affect everything from the types of questions addressed to the timing of the results. Specifically, researchers need to avoid approaching plans with proposals that assume and seek to show that health plans contribute to problems in the current health care system. In addition, researchers need to address the reality that for research to have value for the plans, researchers may need to turn reports around for internal, plan-only use more quickly than it is possible to release reports for scholarly publications.

In the long-term, researchers will also always have to deal with both the ever-growing concerns about privacy for patient records and the privacy regulations those concerns lead to, but such complications do not change the fundamental principles involved in developing sound research projects. Nor do new regulations diminish the value of and need for the answers that research using plan data can provide. As a result, researchers who proceed to study health plans with a bit of caution, a moderate dose of recognition of data strengths and weaknesses, and a large amount of advance planning and communication with the plans can well contribute to policy leaders' understanding of our health care system.

HEALTH PLAN DATA: A RICH RESOURCE RIPE WITH CHALLENGES

Introduction

Over the past decade, an increasingly competitive health care marketplace has led to dramatic growth in managed care and resulted in health plans becoming the primary financial and administrative structures through which most patients, providers, and payers interact. The ongoing changes in the health care system have intensified the need for public and private policymakers to understand the effects of managed care plans on the availability, use, and quality of care. For many years, however, conducting rigorous studies of health plans has been difficult at best, with poor data and the inability to generalize from one plan to another standing out among the challenges researchers face.

To gain any headway in understanding plans and their impact on the system, many health policy researchers have found it necessary to turn to the various data sets, which are developed, maintained, and used by the plans themselves. Overall, most researchers' experience has convinced them that the data plans have to offer — referred to as plan data — include an array of indispensable though problematic sources of valuable information.

Exactly What is a Health Plan Anyway?

With new types of plans emerging almost daily, no single definition prevails. A health plan is generally agreed to be an organization that is responsible for providing or arranging the delivery of a defined set of health care benefits for an enrolled population.

"Because of how managed care is evolving, plan data are the only way to answer many questions," says Barry Saver, M.D., assistant professor at the University of Washington, "but the data are not generally collected for answering research questions. They are usually collected for the plans' business purposes."

Having been collected for business purposes means "that plan data do not exactly match the needs and issues of researchers," says Jon Christianson, professor of health policy and management at the University of Minnesota. "For example, researchers may be pursuing "questions that cannot be answered without both contracts and claims, which are generally maintained for different purposes in different parts of a given organization." And even when researchers can accumulate the data they need, issues of timeliness, historical completeness, and consistency of provider coding emerge as just a few of the issues that lead to questions about the quality of plan data for research purposes.

A further difficulty involved in using plan data for broad-based policy research and analysis lies in the fact that plans are often not comparable in the benefits they provide, the patients they treat, the way they treat them, or the costs they incur. Sometimes the data contain variables that allow for adjustment so that meaningful comparisons across plans can be made. Still, studies of individual plan populations or organizational strategies often are not generalizable to the overall population of a state or locality.

Finally, selection bias is often mentioned as a possible confounding factor in plan data analyses. Selection bias refers to the possibility that patients or enrollees with certain characteristics may select one type of plan rather than another, so analyzing how enrollees respond to changes in organization or payment within a given plan may not be representative of how other (different) enrollees or patients might respond in another plan.

Still, "there is probably nowhere else to go," says Christianson. "It would be very hard to do research on process of care or quality indicators for an employed population without using plan data. It allows us to study the largest segment of the population that is insured. We simply cannot run away from plan data usage."

Making the Best Use of Available Data

In the interest of helping researchers sort through both the potential and the problems related to using plan data, the *Changes in Health Care Financing and Organization* (HCFO) program held a small invitational meeting of experienced researchers in July 1999, called "Using Plan Data for Analysis: Methods and Issues." This report seeks to follow up on the issues and ideas discussed at the meeting in a way that facilitates the use of plan data to its maximum potential, given its inherent limitations.

The report begins by defining the data sets most commonly referred to as plan data. It then outlines how researchers might go about using such data, considering both the limitations of the data itself and the challenges of motivating plan cooperation in projects that may require significant time and other resources. This report also considers ways to manage the technical pitfalls built into the data sets — even when the research projects are built on models appropriate for the data.

Recognizing Wide Range of Information Plan Data Sets Offer

Among the types of data most commonly referred to as plan data are administrative data sets, such as *claims data* gathered from payment forms that detail when services were provided, what type of service was provided, diagnosis, provider type, a unit price for the service, and co-insurance amounts.

There are also *enrollment data*, which include information on insurance product and benefit levels. Enrollment data can also include some socioeconomic characteristics used primarily for rate-setting as well as other pieces of demographic information. In addition, plans maintain provider contracts, which detail how providers are paid and the conditions of their employment. Separate medical departments also maintain a variety of forms and computerized data bases, such as laboratory and referral information, and plans have *medical records data*[1], which are generated by health care providers on health status, symptoms, diagnosis, detailed clinical findings, and treatment of enrollees. Generally, medical records take a variety of forms and are not as standardized as other types of data generated from forms. Finally, plans have *marketing data*, which are typically not made available in any form to outside sources.

Clearly, a wealth of detailed information exists inside the individual health plans, and the use of plan data can offer some advantages over designing new surveys, which can be prohibitively expensive and time consuming. For example, while there are exceptions[2], working with plan data sets means that benefit structures, premiums, and plan characteristics can be held constant or accounted for; data are consistent across enrollees; and the data have a defined population and defined service use over time. Such data are difficult to obtain elsewhere, with the exception of the Medicare program and some Medicaid programs; therefore, resolving the challenges of working with plan data can help make rich resources available to policy researchers and policymakers.

1 In some cases, plans have access to medical records through agreements with providers, as opposed to having medical record data routinely available.

2 For plans that subcontract out care of different patient groups or have different organizational structures with different groups of patients (such as staff/group-model HMO for some patients and IPA-model for others), data can be quite inconsistent. In addition, benefit structures can differ by the contract under which an enrollee is covered.

Examples of Projects that Use Plan Data

Despite the challenges involved in using health plan data, researchers have successfully sought funding from the HCFO program for such projects in recent years. From examining the effects of resource management strategies on patient outcomes to better understanding the reason why employees choose to enroll in managed care versus fee-for-service, these studies have both advanced the field in understanding the myriad interrelationships among consumers, purchasers, and plans, and improved understanding of the problems and advantages that come with using these rich data sources.

Claims data, with their information on service and referral utilization, physician reimbursement practices and patterns, and consumer and purchaser costs, can help answer questions on how aspects of a health plan's design affect cost, quality, access, and satisfaction. Claims data have been, therefore, the primary type used in most of the projects funded. Studies that have used claims data include the following:

- *The Effect of a Physician Gatekeeper on the Cost of, Access to, and Quality of Care in an Employed Population (David Blumenthal)*
- *Impact of Physician Compensation Mechanisms on the Process of Care (Terry Fields)*
- *Investigation of Physician Organization Under Managed Care (Barry Saver)*
- *Managed Care and Contractual Arrangements with Physicians and Implications for Pediatric Health Care Use (Elizabeth Shenkman)*
- *The Effects of PPOs on Health Care Use and Costs (Dean Smith)*

Researchers have also turned to other types of plan data to research more complex organizational issues. In some cases, researchers have found that health plans do not have data systems available and have thus turned to the case study interview method to learn more about health plan behavior. Studies in this realm might look at questions ranging from how managed care organizations and physician groups implement and use health information technology to how physician practice management organizations are attempting to create efficient and quality care provider systems. Such recent HCFO studies include:

- *Evolution of Physician Organization Under Managed Care (Jamie Robinson)*
- *Economics of Health Information Technology in Physician Organizations (Robert Miller)*
- *Selective Contracting for Tertiary Services by Managed Care Organizations (Jack Hadley)*

Learning Multi-step Approach Required to Work Well With Plan Data

Effectively using plan data requires that researchers go through a multi-step process. They must begin by understanding the limitations and potential of the plan data and then developing sound projects that both appropriately use what the data have to offer and include incentives for plan cooperation. To get projects underway, researchers must develop relationships with plans and convince them to participate. Researchers must also deal with the inevitable technical challenges inherent to the data itself and address the

intrinsic generalizability problems related to studies conducted on one or only a few plans.

As researchers begin this multi-stage process, Saver offers one further, critical insight to help minimize the frustrations that can accompany uninformed expectations. "Working with plan data always takes more time that anyone expects," he says. "If, early on, I had just had someone to tell me to expect two-to-three years on this or that project instead of one year to 18 months, it would have been of great value."

Understanding Plan Data Limitations and Potential: Appropriate Models Critical

Christianson notes that researchers need to recognize going in that working with plan data is likely to be more complicated than working with other types of secondary data. "An awful lot of researchers view plan data as comfortably as they view other secondary sources," he says. "They buy tapes from the government [from the Health Care Financing Administration, for example] that comes with documentation manuals, and the researchers are off and running. Plan data does not work that way. There is a lot more involved."

First, plan data are not usually collected by the external researchers seeking to use them. Because the plan data are already collected, researchers cannot — unlike in clinical trials where researchers can control for confounding factors in their experimental design — rely on research design to make up for data limitations. It is critical, given these limitations, that analysts understand what the data really represent and how closely what is measured reflects the research model being tested.

"Modeling essentially means developing a theory about what the world is like and then seeing if the data bears out our thinking," explains Michael Finch, director of research programs for United HealthCare. "A sophisticated model seeks to take into account what happened between the physician and the patient and compensate for the fact that we are so far away from being able to manipulate the design. We cannot tweak here and there and see what happens. Plan data are a Tower of Babel — they cannot speak for themselves. We have to have an idea of what we are looking for and then see if the data speaks to specific questions."

Where and How to Start

A possible initial approach for using plan data is to start with simple models and add variables to see how results change and how robust the analyses are. Another possibility would be to take subsamples of data or different datasets and replicate analyses to evaluate how sensitive results are to the data being used.

Another reality researchers face in dealing with plan data is the potential for endogenous relationships, which result from ever-present factors that contribute to the random variation inherent in processes and systems. For example, if certain types of plans locate in areas that tend to pay providers by a certain methodology, models that correlate provider payment and plan type must account for this, meaning these relationships must be considered and modeled to the extent possible, especially when making policy recommendations or causal conclusions.

Finally, researchers must also be careful not to push the data too far. "It is important not to try to read more into plan data than one ought," says Saver. "There are simply some questions plan data cannot answer."

For example, it cannot, by definition, be used to conduct prevalence estimates of diseases. It also cannot be used to examine trends in services that are not traditionally provided by private insurance (e.g., long-term care). In addition, any causal relationships that are inferred from using one or even several plans' data must be regarded skeptically. Because

plans vary on so many characteristics, sorting out causality is an extremely difficult proposition.

"Still," says Saver, "plan data can be very useful for policy-relevant studies," and once researchers recognize the data's inherent limitations, the next challenge may be avoiding being totally stymied by the fact that they are not perfect resources. Again, matching expectations to the given reality becomes important. "These analyses cannot and do not need to be totally accurate the first time as long as they provide a stimulus for more thought and additional research," says Mark Pauly, Benhein Professor at Universtity of Pennsylvania, Wharton School.

Motivating Plan Cooperation

Once researchers decide on the project they want to pursue using plan data, negotiating with the designated plan or plans to gain access to data becomes an essential next step. Ultimately, each project needs to become an effective collaboration between researcher and plan in order to succeed. To get through the door of the selected plan or plans, researchers need to convince plans not only that the proposed research has value for the health system broadly, but also that the plan's cooperation with the researchers will benefit the plan itself in some tangible way. Essentially, researchers will need to show that their work will contribute in some way to the plan's business interests of growth and financial health. For example, the reports resulting from research on using physician gatekeepers could help inform a specific plans' efforts to determine which direction to go with new disease management approaches. Other projects might assess the value of certain types of monitoring and reporting and so help plans determine if they could make a case for eliminating certain types of costly paperwork required by outside organizations. Although thinking through potential projects from the plans' perspectives may lead to some solid ideas being set aside due to obvious incompatibility with plan interests, learning to view research options as plans would will eventually mean researchers can approach plans with sound research projects that provide sufficient benefit to the plans to encourage their supportive participation.

DO *structure research proposals so that they benefit plans enough to motivate plan researchers to participate — which will require work above and beyond their already full-time jobs.*

"We want to collaborate," says Eileen Peterson, vice president of the Center for Health Care Evaluation at United HealthCare. "We know what you can and cannot do with health plan data, and we can provide researchers with the opportunity to access data systems and decision makers. Researchers just need to structure proposals so that they are sufficiently mutually beneficial to be worth our being distracted from our already full-time work."

DON'T *approach plans with proposals based on the underlying assumption that a plan's impact is negative.*

Time Frames

Peterson comments that acceptable timetables often differ significantly for researchers and plans. "Researchers often get the money, proceed with their studies and some day maybe publish something out of it. Plans meanwhile operate in the real world all the time, and if they participate in a study and never see anything come of it, it is not satisfactory for them. They can be convinced to a certain extent that participating in research is important, but for the plans, the risk/reward ratio is different from that for the researchers."

To accommodate this difference, Peterson suggests that researchers add into their projects some quicker-turnaround reports that might not meet the classic requirement of rigorous research, but will help inform plan decision makers who want clear "yes's" and "no's" about options and who often operate on their "gut instinct." "It is hard because of researchers'

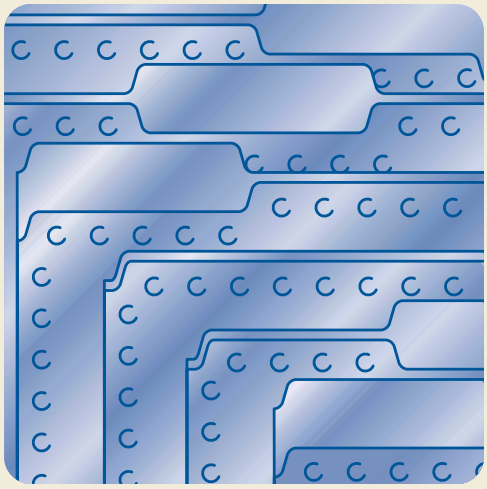
FACT: *Plan and researcher cultures differ, which means their timetables for being able to see concrete results from studies differ as well.*

discipline against extending their findings beyond the strictest recognition of the limitations of the numbers," but it may be necessary to report — just to plan decision makers — information that meets less demanding standards to make the collaboration worthwhile for the plans.

Supporting Peterson's point about accelerating the release of some form of results to help serve plans' interests, Jim Hester, regional administrator for the MVP Plan in Vermont, suggests that such shorter-turnaround reports could be the product of exercises researchers need to go through to become more familiar with the data before they proceed with their more careful publishable work.

Working With Large Plans

The interests and concerns that researchers will need to understand and accommodate will vary depending on the size and resources of the plan or plans being approached. Most published research and research currently underway uses data from the 10 largest plans, such as Kaiser Permanente and United HealthCare, which have 750,000 or more enrollees[3]. These plans typically have internal research divisions where outside researchers will find knowledgeable colleagues who can often enhance their projects. At the same time, however, academicians and other outside researchers need to realize that plan researchers will be operating to a greater extent in the interest of their own plans.



Another reality external researchers need to recognize is that they will be depending on inside researchers and/or data managers to learn what the data has to offer and how to use it.

3 In the HMO market enrollees in firms with enrollment of more than 750,000 total approximately 20,000,000 which represented approximately 26 percent of the 78,795,740 HMO enrollees in 1998. (Source: The InterStudy Competitive Edge, Part II: HMO Industry Report, using data as of July 1, 1998)

IDEA: *Producing "quick and dirty" reports for plan decision makers may actually serve researchers' needs to become familiar and learn how to work with the data.*

Research projects require plan personnel to invest time and expertise in training outside researchers. The plans, therefore, want to work with researchers "who are making a commitment to this type of research [beyond a single project]," says Peterson. "Researchers need to show a track record in using this type of data, and they need to show that they understand managed care plans' imperatives."

Regarding specific project proposals, Peterson reminds researchers that they are approaching "the people who are actually responsible for the care delivery system." This means that researchers seeking to look inside the business of managed care should not submit proposals based "on underlying assumptions that the plan's impact is negative. For example, a research question such as 'How do managed care organizations limit access to specialty care?' is not going to be very appealing," she says. "Researchers need to be objective in how they address questions. They should at least state the questions in terms of seeking to understand the differences managed care plans and policies make.

"It is surprising how often proposals come to us that are less than rigorously academic," she continues. "The anti-managed care backlash may be affecting researchers in ways they do not realize. Whatever the reason, a lack of objectivity makes plans question why they would work with researchers. Plans are looking for indications of how a given study might help them."

In addition to being sensitive to the perspective of the plans, researchers need to be knowledgeable about the types of research and data collection that the plans they approach have done. For example, researchers need to know if the plan has already invested money and time to participate in the Health Plan Employer Data Information System (HEDIS) of the National Committee for Quality Assurance (NCQA).

Because the 10 largest plans have sizeable resources and represent a significant share of the market, studying their best practices and procedures has been valuable. But the organization of these plans does not necessarily represent what is going on in the rest of the market, nor does it necessarily represent the direction the market is heading. Opinions vary. This makes broadening the pool of plans that can be studied a worthwhile exercise.

Working With Mid-sized and Small Plans

Data from mid-size plans are less often used because these plans tend not to have sophisticated data systems or the resources needed to organize data for research purposes. This makes it difficult both for researchers to use data and for the plans to make it available to them. Small plans face even greater challenges.

Reluctance on the part of plans with between 250,000 and 750,000 members to work with researchers does not lie in a lack of appreciation of analytical research[4]. Rather it lies in the realities of an increasingly competitive marketplace. "These are all first class organizations —- most of which have very analytically driven managements," says Hester. "But they are caught between a rock and a hard place. Their competition has evolved from largely local firms. They are now competing with at least one national player with deeper pockets for product development, geographic expansion, updating information systems, etc.

4 In the HMO market enrollees in firms sized 250,000-750,000 total approximately 21,000, 000, which represented approximately 27 percent of the 78,795,740 HMO enrollees in 1998. (Source: The InterStudy Competitive Edge, Part II: HMO Industry Report, using data as of July 1, 1998)

DO *expect to start the learning process anew with each plan.*

"In the more aggressive competitive environment, medium-sized locally controlled firms are having to put all their energies on how to survive," he explains. "With such pressures demanding the attention of senior management, just getting their attention to consider working on a research study will require a greater effort."

Arguing that these medium-sized firms represent enough of the managed care market to be worth the extra effort, Hester asserts that "wonderful data is available," and that researchers can structure research so that it truly serves the host plans' interests. Hester also reiterates the need for studies to focus on issues that also have value for firms whose "primary reason for existence lies not in providing information to the field. They are focused on surviving and improving performance, so questions such as 'What drives provider performance as they divide into all different types of risk-bearing units?' and 'How

What do funders/proposal reviewers want to see in a proposal using plan data?

- *A letter of support from someone at the vice- president–or–above level from the plan.*
- *Any available evidence of a working relationship with the plan, such as a review by plan representatives who know what the plan data really measures.*
- *Tables showing researchers have actually used the plan data and understand it.*
- *Description of researchers' approach to obtaining the data that makes sense from a plan perspective and shows awareness of plan and funder timeliness issues.*
- *Discussion of the potential generalizability of the study outside of the plan.*

effective are different types of provider incentives?' reflect the sort of questions these firms will value." Hester adds that providing some of the "quick and dirty reports" discussed above may be even more important for medium-sized plans to find value in participating in studies.

Some researchers suggest that developing data repositories, or warehouses, may be one way to make the data of plans without internal research resources available and useful. While many plans do not have the resources to analyze their own data, they may be willing to contribute some data to a centralized organization that will analyze it, along with data from other plans. The Agency for Healthcare Research and Quality (AHRQ, formerly the Agency for Health Care Policy and Research) has served this function for plan data on their clinical guidelines, as has NCQA for selected quality indicators. Plans must, however, have some reason for contributing their data to such a repository. That means they must either obtain analyses of their own data useful for their internal functions, or they must obtain analyses that compare their functioning to other similar plans. The organizations creating these repositories must address issues such as data confidentiality and standardization. Funding is also required for these repositories.

Dealing with Inevitable Technical Challenges Built into Plan Data

Beyond obtaining the data, researchers have to realize what they obtain from plans will likely come packaged with pitfalls and limitations, so researchers need to adapt their strategies and expectations accordingly. "Lots of things may differ by plan that have nothing to do with plans," explains Saver. "For example, satisfaction measures do not necessarily relate to the plans because the same providers may well be available in multiple plans. We need to realize plan data can be a useful tool, but we have to be aware that there will be limitations that we will not know up front."

When working with any plan data, researchers suggest specifically investigating and evaluating the following:

- Units of analysis. Plans differ in their definition of an encounter, a visit, or a claim (when they have them). These units can differ within a given plan's data system, and almost certainly differ across plans.
- Face validity. When a data item does not appear to make sense, be sure to check it out with people who know the data well. The problem could be one of documentation or coding.
- Definitions. Some variables are extremely difficult to measure well (e.g., physician incentives, which can vary from plan to plan and can be ambiguous and difficult to delineate within a single plan). Again, be sure to check exactly how a data element is defined.
- Time periods/censoring/persistence. Researchers need to watch for records being truncated in some time periods. Researchers must also determine how continuous data need to be for analysis, and whether such time periods are available.
- Item non-response for relevant items. Researchers must determine how much missing data will affect the validity of their estimates and define acceptable levels of non-response. If there is substantial non-response, researchers must investigate for potential non-response bias and if possible, correct for it.
- Confidentiality issues. Any data containing specific medical information is becoming increasingly sensitive.[5] Researchers must also be careful to address not only legal requirements, but "good faith" agreements. The data belong to the plan, and any breaches of trust will damage a researcher's relationship with the plan for future studies.

5 The Department of Health and Human Services issued draft privacy regulations in November 1999, with final regulation release expected sometime in 2000. The new regulations could affect researchers' ability to collect confidential patient data from health care providers without patients' informed consent.

The process of learning the strengths and weaknesses of a specific plan database essentially begins at ground zero with each new plan approached. Detailed knowledge of a database is especially important when the data are to be used in studies for which the data were not originally collected —- which is almost always the case in research projects. A working relationship with someone within the plan who knows the data intimately is almost a necessity.

Another necessity, which researchers find frustrating, is that they must often invest in learning the data before obtaining project funding. Some plans also require resources to provide data and support to researchers. One potential solution is for funders to provide planning-grant money to build relationships between external and internal plan researchers or data staff.

Creatively Addressing Generalizability Problem

Using one, or even several, plans to make generalizations to the entire health care market is obviously problematic. As many researchers now joke, "When you've seen one plan, you've seen one plan." Still, extensive knowledge of one plan (especially a very large plan, which may differ by site), can be informative when examining the impact of medical, financial, and organizational interventions on a variety of outcome variables, including cost, access, quality and effectiveness of care, and provider and organizational efficiency of care. For example, within United HealthCare there are capitated and non-capitated arrangements with providers, and Harvard-Pilgrim Health Plan is really four plans with an umbrella administration. This means "you can get needed variation with the simplicity of a single data source," notes Hester. These studies can help plans operate more effectively and feed into continuous quality improvement initiatives as well as inform public decision makers.

Yet because individual health plans differ in so many ways — in their organizational structure, their financial and managerial relationships with providers, and their benefits, to name a few — policy analysts must often evaluate whether results of studies using one (or even several) plans' data are generalizable to broader federal and/or state policy issues. Researchers must be extremely careful in how they frame their findings and their generalizability to the broader marketplace.

On the other hand, some argue that it is the complete body of research, on many different plans and from many complementary surveys and other data sources, that policymakers must consider, not the isolated results of an individual study. Each study is a piece of the puzzle, not the whole picture. The researcher must justify the contribution of their project to the overall body of knowledge. At the same time, the researcher must deal with the paradox that while it may be harder to obtain funding for innovative and groundbreaking studies because the quality of the data (or methods) is unknown and their policy or research relevance is unclear, such studies are more highly valued by prestigious journals.

Final Thoughts

Balancing the array of interests and potential audiences involved is a challenge intrinsic to almost any type of research. The incentive in working through these issues as they relate to using plan data lies in the growing number of Americans obtaining care through health plans. With each new plan enrollee, the importance of understanding what does and does not work in terms of system structure and care delivery increases.

Moving ahead, as the number of people enrolled in health plans grows, so does the concern about privacy of health-related information. Such concern will likely generate additional regulations and restrictions on data availability, making the use of plan data ever more complicated. Such developments, however, will neither diminish the need to study how well health plans serve enrollees nor change the fundamental principles involved in developing sound research projects using plan data.

HEALTH PLAN DATA: A RICH RESOURCE RIPE WITH CHALLENGES

Produced for The Robert Wood Johnson Foundation

This report was prepared by the Academy for Health Services Research and Health Policy under The Robert Wood Johnson Foundation's *Changes in Health Care Financing and Organization* (HCFO) program. The HCFO Program supports research, demonstration and evaluation projects that examine major changes in health care financing and organization issues with implications for current public policy. The Academy for Health Services Research and Health Policy serves as the Foundation's program office for this initiative.

Written by:

Carole Lee and Amy Bernstein

Program Director:

Anne K. Gauthier

Designed by:

Bremmer & Goris Communications, Inc.

MANY thanks to:

Jon B. Christianson, David C. Colby, R. Adams Dudley, Terry S. Fields, Michael D. Finch, Jim A. Hester, Christopher Hogan, Mark C. Hornbrook, Haiden A. Huskamp, Robert J. Jeddeloh, David J. Knutson, Helene Levens Lipton, Alan Lyles, Robert H. Miller, Stephen T. Parente, Mark V. Pauly, Eileen H. Peterson, Barry G. Saver, Kevin A. Schulman, Elizabeth A. Shenkman, Dean G. Smith, Kenneth E. Thorpe, Jonathan P. Weiner, Judith Wooldridge

1801 K Street NW

Suite 701-L

Washington, DC 20006

Tel: 202.292.6700

Fax: 202.292.6800

E-mail: hcfo@ahsrhp.org

Internet: www.ahsrhp.org

July 2000