

# **THE METHODOLOGY USED TO MEASURE HEALTH CARE CONSUMPTION DURING THE FIRST YEAR OF THE HEALTH INSURANCE EXPERIMENT**

**PREPARED UNDER A GRANT FROM THE DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE**

**KENT H. MARQUIS**

**R-2126-HEW  
AUGUST 1977**



The research reported herein was performed pursuant to Grant No. 016B-7501/2-P2021 from the U.S. Department of Health, Education, and Welfare, Washington, D.C. The opinions and conclusions expressed herein are solely those of the author, and should not be construed as representing the opinions or policy of any agency of the United States Government.

# **THE METHODOLOGY USED TO MEASURE HEALTH CARE CONSUMPTION DURING THE FIRST YEAR OF THE HEALTH INSURANCE EXPERIMENT**

**PREPARED UNDER A GRANT FROM THE DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE**

**KENT H. MARQUIS**

**R-2126-HEW  
AUGUST 1977**





PREFACE

The Rand Health Insurance Study (HIS), funded by a grant from the U.S. Department of Health, Education, and Welfare, includes an experiment investigating the effects of different health care financing arrangements on the demand for medical services. Obtaining detailed, reliable, and accurate information about the consumption of health services by families in the experiment is essential to the goals of the study. This report discusses conceptual issues, pretest results, and the methodology used to measure consumption of health care goods and services in the first year of the experiment. It is the first in a projected series of reports discussing the measurement of consumption in the Health Insurance Study. Subsequent reports will contain the results of experiments designed to detect and estimate any remaining biases in the utilization data.

Additional Rand reports and related publications discuss other design features of the study. The experimental design for estimating effects of financing on demand for care is described in Joseph P. Newhouse, "A Design for a Health Insurance Experiment," *Inquiry*, Vol. 11, 1974, pp. 5-27. Features of the design that permit the estimation of effects on utilization behavior attributable solely to participation in the experiment are discussed in Joseph P. Newhouse, Kent H. Marquis, Charles E. Phelps, and William H. Rogers, "Measurement Issues in the Second Generation of Social Experiments: The Health Insurance Study," *Proceedings of the Social Statistical Section, American Statistical Association*, 1976. Carl N. Morris, *A Finite Selection Model for Experimental Design of the Health Insurance Study*, The Rand Corporation, R-1837-HEW, forthcoming, describes the logic and advanced techniques used to determine (1) optimum sample sizes for experimental treatments and (2) the selection and balanced allocation of individual families to experimental treatments. Issues related to the assessment of health are treated by Robert H. Brook and John E. Ware, *Conceptualization and Measurement of Health for Adults in the Health Insurance Study*: Dayton, The Rand Corporation, R-1987-HEW, forthcoming; and by John E. Ware,

*The Conceptualization and Measurement of Health for Policy Relevant Research in Medical Care Delivery*, The Rand Corporation, P-5599, February 1976. David H. Stewart, *A Design for Information Processing in the Health Insurance Experiment*, The Rand Corporation, P-5229, September 1974, describes the design of the data storage and retrieval system used by the study.

### SUMMARY

The research design of the Health Insurance Study addresses questions about the demand for health care. Constructs specified by economic theory are measured to observe relationships between consumption of medical services and the prices for those services faced by consumers. An important goal of the research is to generate unbiased results that apply to a variety of financing and coverage possibilities. These theoretical and applied objectives shape the measurement requirements for the research.

The demand estimation goals of the experiment require that measurement be designed to observe a broad range of health care services, to infer demand for specific services, to obtain details of each health service used, to scale each item consumed in ratio units, to assign each observed service to a point in time and an episode of illness, and to minimize measurement effects. Measurement biases correlated with the experimental treatments or "true" utilization can distort estimates of the price effects. Other kinds of biases, such as (additive) Hawthorne effects and random response error, will decrease the precision of estimated relationships or cause misestimation of a mean or the constant term in a linear regression model.

The most probable measurement bias would be a failure to detect all health care consumption because of reliance on only insurance claims to measure utilization. Therefore, a supplementary reporting mechanism was established. A small scale pretest was conducted on a pilot sample to compare the effect of weekly and monthly supplementary reporting on the amount of outpatient consumption detected within each of two extreme experimental financing treatments: one in which all outpatient care was free and another in which families paid the entire cost of care until a large deductible requirement was met. Results suggested that a measurement system using only insurance claims could result in a 20 percent underestimate of overall utilization. Measurement effects were correlated with the experimental insurance plan treatments. The group submitting weekly reports reported significantly more

consumption than the monthly report group within a deductible treatment. No significant effects of supplementary reporting were found in a plan in which service was free. The exact nature of the measurement bias could not be inferred from the data: Weekly reports could stimulate utilization (a Hawthorne effect), or induce more complete reporting of otherwise forgotten utilization (or both). The pretest also indicated that providers (doctors, pharmacies, etc.) did not always furnish complete information and that provider nonresponse was correlated with the experimental insurance treatments.

The utilization measurement procedures used in year one for the regular sample included both insurance claims and supplementary reports. A Health Report, filled out weekly or biweekly, was the central part of the supplementary reporting system designed to detect utilizations not reported on Medical Expense Reports. Families were paid for sending in completed health reports on time; form and item nonresponse were followed up. To minimize recall error, short reference periods (one to two weeks) were used and respondents were asked to keep daily records of utilization on forms furnished by the HIS. The forms included a wall calendar with probe questions and a small tear-off sheet on each outpatient Medical Expense Report that could be retained until it was time to enter the utilization on the Health Report.

The Medical Expense Reports (claim forms) were used to obtain details and costs of the separate health services used by participants. Separate forms were designed for the major kinds of services (hospital, physician, dentist, pharmacy or supplies). Each contained short answer questions, check boxes, and specific coding categories so the detailed treatment cost data required for analysis could be obtained. The expense reports were the basis of cost reimbursement and were edited for completeness (providers were recontacted as necessary for missing information). Attempts were made to obtain a Medical Expense Report for each utilization reported on a Health Report.

The evaluation of the year one system for measuring utilization is not yet complete. Preliminary findings indicate poor quality measurement for control group families, which contributed to the decision to drop this group from the HIS. The new Health Report and tracking



procedures may not have eliminated all of the measurement bias correlated with experimental insurance treatments. As a result, the HIS began paying 5 percent of the costs of each health service for families with deductible coverage and has designed record check experiments that will be more definitive about the nature and extent of any remaining measurement biases correlated with the experimental variables. Estimates of measurement biases (if any) derived from these experiments will be used to adjust parameter estimates.



ACKNOWLEDGMENTS

The design of the Health Insurance Study's utilization reporting system, described in this report, represents the work of many people. The initial design and pretest were directed by Charles Phelps working closely with Joseph Newhouse, Principal Investigator for the HIS, and with the Urban Opinion Surveys Division of Mathematica, Inc., which administered the pretest. Susan Marquis, Rand Economist, served as liaison between the economic researchers and the survey measurement group, translating the concepts, concerns, and methodologies across the disciplines so that a revised year one design could be formulated to meet the experimental objectives within realistic measurement and operational constraints. Several of the conceptual and quantitative formulations in the report are a direct result of her contributions. Preparation of the manuscript was aided by helpful comments from Ronald Andersen of the University of Chicago and Stephen Crocker of Rand. The conscientious efforts of Marilyn Hecox, Sarah Slaughter, and Rebecca Wier of Glen Slaughter and Associates (administrative services subcontractors to the Health Insurance Study) are gratefully acknowledged.

The research and administrative assistance of Marie Brown, Sally Hanes, and Anita Stewart of Rand is greatly appreciated, as is the patience and care exercised by Constance Smedley in typing the manuscript drafts.



CONTENTS

PREFACE .....	iii
SUMMARY .....	v
ACKNOWLEDGMENTS .....	ix
Section	
I. INTRODUCTION .....	1
II. MEASUREMENT CONSIDERATIONS .....	3
Measurement Implications of the Research Goals ...	3
Measurement Bias and Random Error .....	11
An Alternative .....	21
Unresolved Questions About Measurement Bias .....	22
III. PILOT STUDY TO ASSESS NONREPORTING OF UTILIZATION ..	23
Research Design .....	23
Results .....	29
Other Information from the Pilot Studies	
Relevant to Measurement Design .....	38
Implications of the Pilot Results .....	41
IV. FIRST YEAR DESIGN FOR MEASURING HEALTH CARE	
CONSUMPTION .....	43
Overview of the Measurement Design .....	43
Supplementary Reporting System .....	45
Data Followup .....	52
The Medical Expense Report .....	55
Alternatives to Medical Expense Reports .....	58
Measures of Utilization for Which MERs	
Are Not Required .....	66
Health Report Reference Period Experiment .....	66
V. PRELIMINARY EVALUATIONS .....	70
REFERENCES .....	75



## I. INTRODUCTION

In this report attention is given to the methodology used to measure the consumption of health care services within the context of the experimental part of the Health Insurance Study (HIS). The report has three major objectives: to present the goals of the experiment with their consequent measurement implications, to consider the data collection implications of some pilot study results, and to describe the utilization measurement procedures of the first year of the HIS demand experiment. Consideration is limited to measuring the type, quantity, and timing of health care goods and services consumed.

The Health Insurance Study includes a large scale, interdisciplinary social experiment that seeks to estimate the demand for health care as a function of prices faced by consumers and as a function of other variables. The prices are experimentally varied by assigning families to specially created health insurance plans. For the analysis of demand, the dependent variables to be measured are the types and quantities of health care services consumed.<sup>1</sup>

The experiment is designed to contribute basic knowledge in the area of consumer and physician economic behavior and also knowledge relevant to public policy issues of health care financing. The need to generate information about a broad range of policy options has resulted in creation of procedures for measuring health service consumption that go beyond the former state of the art.

There are a large number of proposals for what national health insurance should be. They differ in terms of what kinds of services would be covered (e.g., hospital, outpatient physician, dentistry, mental health, prescription medicine, chiropractic); what amounts various income groups would be expected to pay; whether there should be coinsurance or deductibles, or limits on total out-of-pocket expenses,

---

<sup>1</sup>Other major objectives of the experiment include assessing the effect of price on both the quality of health care received and changes in biological, psychological, and social health. The measurement of these classes of dependent variables is not discussed here.

and if so how much. The number of permutations of the various coverage and financing features that represent "viable" policy alternatives is probably too large to consider conducting demonstration studies of each.

Newhouse (1974a) points out the possibility of abstracting the various issues of the National Health Insurance (NHI) debate into the context of economic theory. The various NHI proposals can be represented as points along a single dimension of the price of care faced by the consumer. The effects of price on the consumption of total and specific kinds of health care can be investigated by experimentally testing a limited number of price conditions, none of which, by itself, necessarily represents an existing or future NHI proposal. This, indeed, is what the HIS experiment attempts to do. To be able to abstract the applied problems to the level of theory, collect data relevant to the theory, and then generalize results back to a large number of policy options necessarily imposes heavy demands on experimental and sample designs, measurement designs, and analysis strategies. These demands are much more comprehensive than those involved in the design and conduct of a demonstration study of a single financing option. For measurement, the most significant consequence is the need for precise, accurate primary data about each of a full range of components of the dependent variables. These data can then be pulled apart and reaggregated to make unbiased inferences about NHI options.



## II. MEASUREMENT CONSIDERATIONS

This section describes the major conceptual issues that guided the formulation of the methodology for measuring health care consumption in the first year of the Health Insurance Experiment. The research goals, which determine measurement content, units of analysis, and scaling requirements, are discussed first followed by a consideration of how measurement bias and random measurement error can affect the specific kinds of statistical estimates to be made from the data. From this, the major measurement design uncertainties are derived and the issues for pretesting are posed.

### MEASUREMENT IMPLICATIONS OF THE RESEARCH GOALS

The measurement requirements that derive from the demand estimation goals of the research are considered here. Attention is given to the choice of operational definitions, the scope of elements to be measured, the depth or detail of information needed about each health service use, scaling requirements, and the need to relate details of utilization to specific time periods and illness episodes.

### Operational Definitions of the Construct of Demand

Demand is a theoretical construct that is part of economic theory. For this study, demand refers to the quantity of services the individual *wishes* to obtain (or the physician, acting as his agent, wishes him to obtain) over a specified time period, given the price of services and the individual's purchasing power or "ability to pay." In the theory, demand combines with other variables (e.g., supply) to determine consumption. An operational definition or measurement rule is needed to translate the construct into observables for purposes of empirical research. There are two approaches to the operational definition; they may be referred to as hypothetical and observational. Each has potential strengths and weaknesses.

Hypothetical. The hypothetical approach to operationalizing demand involves obtaining responses to cognitively defined, hypothetical circumstances (how much health care would you choose to buy if it cost \$X

per unit, your health was at level Y, and your income was \$Z per year?). From one point of view, the hypothetical approach is ideal. It permits the measurement of responses to conditions that cannot be experienced by the respondent. The main problem is the questionable ability of the responses to hypothetical situations to generalize to real world conditions (and it is usually this kind of generalization that the researcher wants to make). The generalization can fail for at least two reasons: the research has misspecified (e.g., omitted) some relevant, real world conditions in his hypothetical questions, or the responses given do not reflect the same psychological variables in the right mix and quantity that determine economic decisions in the experiential (real) world.

How well do answers to hypothetical questions generalize to observable behavior? The most direct evidence (summarized by McNeil, 1974) comes from the large scale surveys of consumer buying expectations, intentions, and attitudes conducted by the U.S. Bureau of the Census from 1959-1973. Throughout this period a variety of questioning and analysis strategies unsuccessfully attempted to produce leading indicator estimates of durable goods purchases, especially new car sales. The failure of the program was due, in part, to the lack of predictive validity of the individual household data. According to McNeil (1974, p. 7), "households with no plans to buy [automobiles] accounted for a very large share of subsequent purchases." At a more general level, the social psychological literature contains numerous studies reporting a failure of measured prior "attitudes" to predict future behavior.<sup>1</sup> For the hypothetical approach to be valid, attitudes, intentions, knowledge, and other cognitive variables must be causally related to subsequent behavior.

This same literature, especially experiments following Festinger's (1957) publication of the theory of cognitive dissonance, indicates that cognitively expressed desired (valuations, satisfactions) for an item can be more a *result* than a cause of purchase or other choice decisions. At this point consumers do not appear to be good predictors of what

---

<sup>1</sup>See, for example, LaPiere (1934), Brehm and Cohen (1962), Festinger (1964), Deutscher (1966, 1973), Fishbein (1967), Wicker (1969, 1971).

they are going to do given the current state of the art of subjective measurement.

Observation of Consumption. Purchase decisions can be observed under a variety of conditions. An operational definition of demand can involve observations of actual consumption. The main problem with this approach, which is directly derivable from theory, is that actual consumption is not necessarily isomorphic with the construct of demand. What is consumed is a function both of demand (quantity wanted at a given price) and of other variables ("supply"). For example, it is possible that desired and available services do not coincide. Technology is not advanced enough to provide what is desired at the particular cost; providers of services are too overworked and must ration their time; laws or other regulations may prohibit delivery of the desired service. If demand is operationally defined as actual consumption, the other causal variables must be specified, measured, and taken into account at the analysis stage. The consumption observation approach has been chosen for the HIS and the problem of specifying, measuring, and analyzing non-cost constraints to utilization addressed in detail. See Newhouse (1974b) for a discussion of the research design and the steps taken to minimize threats to the external validity (generalizability) of the demand estimates based on observation of consumption behavior.

#### The Nature of Health Care Consumption

The need for the complex experimental and analytical designs described by Newhouse is partly the result of the heterogeneity of consumption and payment possibilities in the health care market. Many different kinds of purchases are possible (primary professional care, specialist services, medicines, appliances, custodial care, dentistry, hospitalization, and so forth), some of which are substitutes for each other (e.g., hospital vs. outpatient care, medical vs. surgical treatment). Within and across the services, quality of care and prices per unit vary. The consumer can choose whether or not to use services; if the choice is positive, he can decide in conjunction with his physician how much and what type to buy. The consumer can substitute one kind of advice or treatment for another kind, can decide to invest in his future health,

attempt to maintain his present status, or decide to postpone care. The decisions are influenced, in part, by relative costs of the services to the consumer, which are influenced by what his own health insurance will pay for. His insurance may pay more for one kind of service than another and reimburse at different rates depending on how much has already been spent. The consumer can also consider preferences for a particular style of delivery (e.g., a physician's bedside manner, common cultural heritage, attractiveness of the receptionist), how much he wants to search for low cost suppliers, whether he should take time off from other activities for treatment, and how far he wants to travel. For purposes of generalizing research results, these decisions, both as independent events and as they affect each other, need to be represented in the analysis models; and appropriate measurements need to be designed for estimating the complex models.

#### Scope of Services to be Measured

For the purposes of assessing consumption a choice must be made among a number of options for the range of health services to be covered and measured. Measurement could be confined to a narrow range of services such as the most costly (e.g., hospitalization), the most controversial in terms of NHI coverage (e.g., dentistry, preventive care, prescription medicine, mental health), or the most typical (e.g., outpatient physician services). Orr (1974) points out that excluding coverage of a service may lead to distorted utilization patterns, such as using covered hospital services instead of uncovered outpatient care. Failing to measure all covered services, regardless of whether there is any charge for them by the provider, will result in biased estimates of total demand. The threats to generalizing experimental results become even greater if the unobserved consumption of a service is correlated with one of the "predictor" variables and substitutes for services that can be detected.

The kinds of health care goods and services covered and measured in the HIS experiment represent a very broad scope. In addition to hospital, physician, and clinic services, they include services of licensed nonmedical practitioners (nurses, psychologists, optometrists,

chiropractors, Christian Science practitioners) as well as prescribed medicines, appliances, and supplies.<sup>2</sup> Special efforts are made to detect consumption of services for which there is no direct charge, such as telephone consultation; follow-up visits; and "free" treatments at company, school, and community clinics (see Sec. IV).

The kinds of health care not measured during the first year included advice from friends, relatives, pharmacists, religious consultants, etc.; nonprescription medicine and supplies (except those actually covered and reimbursed by the experimental insurance plans);<sup>3</sup> and care received while in jail, in the military, and in similar "total" institutions.

#### Measurement Detail

Given that a broad range of health care services are to be measured, the next question is in how much detail should the characteristics of consumption be detected.

An important issue in policy formulation will be which services to cover and with what cost sharing provisions. To contribute meaningfully to this debate, consumption measurement should be disaggregated at least to the level of service at which the policy debate is focusing. This requires separate estimates of the major kinds of health services: hospital room and board; inpatient medical and surgical services; inpatient ancillary services; outpatient physician care (preventive separated from therapeutic); dental services; mental health service (by type of practitioner); nursing and other paraprofessional care; services of licensed, nonmedical health care practitioners (separated by specialty); prescribed medicine; and other prescribed items (with separate estimates of eyeglasses and hearing aid consumption).

Disaggregation to a finer level is necessary to avoid potential estimation bias. Even if cross-price elasticities among medical services are zero,<sup>4</sup> the experimental estimates of the response of a set of

---

<sup>2</sup>For a detailed discussion of coverage, see Clasquin (1973).

<sup>3</sup>Nonprescribed medicines, appliances, and supplies were added to the measurement system at the beginning of the second year of the experiment.

<sup>4</sup>The experiment will attempt limited estimates of cross-price elasticities.

medical services to price may not generalize to a situation in which the mix of covered services is different. Regression coefficients estimated using the set of medical care services as the dependent variable will represent a weighted average response to the price of the component services. The weights are the price shares.<sup>5</sup> If consumers respond differently to price changes in each of the component services, the estimated average price response will generalize only to situations in which the price shares (weights) are identical to those in the experiment. Under current insurance, many services are subject to different cost sharing (own-price) conditions but, in the Health Insurance Experiment, all services are subject to the same cost sharing conditions (for a given family). Price shares may differ in the experimental situation from current insurance, and national health insurance proposals may adjust the relative cost sharing arrangements further. (See Newhouse and Phelps, 1974, for a more complete discussion of the aggregation bias.)

It is desirable for health care measurement to allow separate price elasticity estimates (a) for service subject to different cost sharing conditions under *current* insurance (since NHI may not preserve the existing relative cost sharing characteristics and hence change price shares), (b) for services likely to be subject to different cost sharing conditions in NHI relative to each other, and (c) for services whose price elasticities differ. The policy-relevant disaggregation described earlier is probably sufficient to meet objectives (a) and (b). Since so little is known about the price elasticities of components of the major types of services, the best way of being sure that objective (c) is met is to measure consumption in great detail. Even if the components of the major health services are found not to have very different price elasticities, the generalizations from the experimental data to conditions where the component price shares differ from those observed in the study can be made with much greater confidence.

Policy planners will also need to assess the temporary increases or decreases in demand for specific services under the various NHI proposals.

---

<sup>5</sup>The price share of one service within a set of services is the price of the single service relative to the aggregate price of all services in the set.

Thus the dependent variable measurement should allow separate estimates of those services most likely to be sought under generous NHI coverage or postponed under less generous coverage. These would include elective surgery, well-care examinations, preventive procedures (e.g., fluoride treatments, immunizations), and probably others.

### Scaling

The disaggregated observations must be measured in ratio scale units of quantity in order to be recombined into other analysis units (such as illness episodes, which are discussed below). Two different scaling options can be considered: the simplest is to use the total charge by the provider as an index of the quantity of services dispensed. Dollars are ratio scale units and can be assigned to each of the wide variety of goods and services available in the health care delivery system. The other is to adopt the valuation systems used by the industry to process insurance claims.

Dollar charges are sometimes an inexact proxy for the number of units of health care delivered but provide a common ratio scale metric across a wide variety of services. Newhouse and Phelps (1974a) discuss the problems of using expenditure data, among which is that prices are a function of variables other than intensity units. The other determinants include personal preferences (e.g., for short wait times, geographic proximity, a particular style of care); third party subsidy (e.g., lower or zero prices for care in community, school, or company facilities); short-run supply effects that interact with type of service or physician specialty to distort the relationship between charges and units of service delivered; and pricing structures that provide cross-subsidies among different services (e.g., subsidizing the obstetrics ward from laboratory profits).

The HIS will base its consumption unit scaling both on expenditures and on procedures currently used by the commercial health insurance industry. The procedures involve assigning ratio scale code values<sup>6</sup> to

---

<sup>6</sup>Based, for example, on the average charge for a particular service by all physicians in an area. The use of this kind of scaling removes variance in individual provider charges unrelated to quantities delivered, such as variance due to the amenities, subsidies, etc. discussed above.

each detailed procedure (or item) reported by the health care provider on itemized insurance claims forms. The measurement implications are that each diagnostic and therapeutic procedure, each ancillary service, each prescribed health care item, etc. needs to be made known in order to assign accurate quantity scale values to observed consumption.

#### Aggregation over Time Periods

The concept of demand for services must be related to some time period specification, such as an interval of the insurance accounting period. The measurement task, therefore, requires assigning health care goods and services to a point in the insurance accounting period. Services then may be aggregated over a specified time period to form the dependent variable.

A second time period concept is the episode of illness. For reasons discussed by Keeler et al. (1974) and Newhouse (1974a), analysis of the demand for services per episode of illness will substantially increase the accuracy of the estimated relationship to price.<sup>7</sup> A second measurement task, therefore, is to assign each good and service received to an illness episode. For this purpose, consumption measurement need not be detailed so long as a person does not experience more than one episode at a time. If someone has more than one concurrent episode, the individual diagnostic and therapeutic procedures received must be disaggregated to a fine enough level so that each may be linked to its appropriate episode.

Tentative estimates of the probability of a person experiencing more than one episode of illness at a time can be made using data from various sources. The National Health Survey (Wilder, 1975) estimated that adults<sup>8</sup> experienced between 1.2 and 2.0 acute conditions<sup>9</sup> per person in FY 1972-73. The last time these estimates were published, 50

---

<sup>7</sup> Accuracy gains occur if the coinsurance rate varies over an accounting period and the consumer takes this variation into account in planning his use of health services. He is most likely to plan ahead if he has a diagnosed health problem (episode of illness) requiring future treatments.

<sup>8</sup> Ages 17-64. This HIS does not include persons over 64 years of age in its samples.

<sup>9</sup> Either medically attended or causing activity restriction.



percent of adults had at least one chronic condition. Thus, it seems reasonable to assume that in the HIS adult population, during some time in each accounting year, about half will have at least one chronic and one acute episode simultaneously. From records of a prepaid health maintenance plan in New York, Balamuth et al. (1961) also estimate that about half of their sample have at least one chronic condition (by definition, medically attended) and that 20 percent have at least two possibly chronic conditions (by definition, simultaneous). Thus, for a not insignificant percentage of adults enrolled in the Health Insurance Study, medical care received in any one outpatient visit could be allocated to more than one episode of illness. This linking is possible only if there is a detailed list of services received and information on the relation of each service to each diagnosis.

Tentative plans on obtaining measurements to identify an episode of illness are that an episode is a "primary" diagnosis or set of symptoms (which may become more specific over time) and that related but secondary diagnoses (e.g., "complications") are part of the main illness episode. Variation in the use of reporting terminology among physicians requires expert judgment from medically trained coders to discriminate among "new" primary and secondary diagnoses. Judging by a survey of medical records of primary care physicians in the fee for service system (Tenney et al., 1974), the need for expert coding will not be rare. Detailed knowledge of diagnostic and therapeutic procedures associated with each diagnosis will help in increasing the accuracy of this specialized coding.

#### MEASUREMENT BIAS AND RANDOM ERROR

The economic analysis objectives just discussed determine the scope, detail, and scaling requirements for measurement of health care consumption. The measurement goal is to design a system that meets the economic analysis objectives and minimizes errors, biases, and other effects of measurement on the data--that is, to obtain measurements that minimize the mean squared error of the price-quantity relationships to be estimated. The following discussion focuses on the statistics and parameters to be estimated, the kinds of unwanted measurement effects that

may occur, and the effects that imperfect measurement may have on the analysis.

The effects of imperfect measurement on estimates of the relationship between price and the quantity of service consumed are considered. Two ways of portraying that relationship are the unstandardized regression coefficient in a linear model and the elasticity coefficient obtained from the linear model. Imperfect measurement has slightly different effects on each type of coefficient, and the kinds of measurement imperfections themselves will produce different distortions in each estimate.

The two basic categories of measurement imperfection are bias and random error. When a person reports a fact, such as one visit to a doctor last week, the report may reflect true consumption exactly or it may be either an overstatement or an understatement of truth. Measurement theory defines two kinds of errors that may be contained in the report: Imagine a person who reports the number of visits to a doctor he made last week several times, and further imagine that that person replies independently each time (he does not remember his previous response and nothing else about him changes). The numerical responses given by the person have a frequency distribution, and the deviation of the mean of the response distribution from the true number of visits is the response bias.<sup>10</sup> The deviation of any one observation from the mean of the observations is the random error in that observation (by definition, the expected value of the deviation of any observation from the mean is zero and the variance of the deviations is assumed to be finite). This variance of the deviations of observations from the mean is usually called random response error variance, or simple response variance. The term "random error" is used in this report.

The two basic concepts of measurement imperfection may be represented in model form as follows:

$$M_i = T_i + B_i + e_i , \quad (1)$$

---

<sup>10</sup> Response bias is one component of overall measurement bias. Other sources of systematic deviation of measured values from true values include errors made by interviewers, coders, and computers.

where  $M_i$  = Measured value (report) of individual  $i$ ,

$T_i$  = True value for individual  $i$ ,

$B_i$  = Bias in individual  $i$ 's report,

$e_i$  = Random error in individual  $i$ 's report (with an expected value of 0).

Two kinds of measurement biases need to be distinguished. One is a *response* bias where the information furnished by the respondent is systematically different from the true information. In this case reported (measured) utilization does not equal true utilization as when a respondent systematically fails to report visits to a psychiatrist or faithfully reports clinic visits when the appointments are not kept. A second kind of bias is a Hawthorne or *reactive* effect. It occurs when the consumption observed in the experiment is different from consumption in an otherwise identical, nonexperimental condition because the fact of measurement has systematically influenced actual demand (increasing it or decreasing it).

Both the response bias and the reactive bias can be correlated or uncorrelated with other variables in the analysis. For applied measurement design discussion, the bias parameter in the measurement model (1) will be expanded to include both response and reactive effects and will be partitioned into:

$\delta T_i$ , bias correlated with true utilization where  $T_i$  is true utilization in the absence of measurement,

$TP_i$ , bias correlated with experimental plan, and

$b_i$ , bias uncorrelated with either true score or plan.

The measurement model in (1) becomes:

$$M_i = T_i + \delta T_i + TP_i + b_i + e_i . \quad (2)$$

The quantities  $b_i$  and  $e_i$  are assumed to be uncorrelated with each other, with  $T_i$ , and with  $P_i$ .

### Measurement Effects on Parameter Estimates

Before priorities are set in the measurement design, the effects of the different kinds of measurement imperfections on the parameter estimates are derived. The effects of measurement error and bias can be shown with respect to estimating a price coefficient and a price elasticity. An example of the form of the model of demand to be estimated (from Newhouse, 1974a) might be:

$$Y_i = \alpha + \sum_j \lambda_j Z_{ij} + \beta_y P_{iy} + \beta_x P_{ix} + u_i,$$

where  $Y_i$  represents the true number of units of health care service Y demanded by individual i,

$Z_{ij}$  is a vector of characteristics of individual i such as age and health status,

$P_{iy}$  is the price of service Y facing individual i,

$P_{ix}$  is the price of another service X, and

$u_i$  is the stochastic disturbance term, assumed to be independent

of the  $Z_{ij}$ ,  $P_{iy}$ ,  $P_{ix}$ , and  $b_i$  and  $e_i$  of the measurement model.

The estimated  $\beta_y$  and  $\beta_x$  can then be used to estimate own-price and cross-price elasticities.

To demonstrate the effects of imperfect measurement on estimated relationships, a simplified demand model is examined:<sup>11</sup>

$$Y_i = \alpha + \beta P_i + u_i, \quad (3)$$

where  $P_i$  is the price of service Y for individual i, the subscript y is deleted on the price (P) and its coefficient ( $\beta$ ).

### Measurement Bias Effects on the Price Coefficient

Imperfect measures will bias the estimated price coefficient ( $\beta$ )

<sup>11</sup>The price  $P_i$  is an experimental variable and is orthogonal to all  $Z_{ij}$ . Hence, the conclusions about the effect on the own-price coefficient in this simple model are generalizable to a model in which the vector  $Z_{ij}$  is included as a set of explanatory variables.

in the regression of experimental plan or price (P) on consumption of medical care (Y). Using the measurement model at (2) (and substituting  $Y_i$ , the number of units of health care service, for  $T_i$ , the true value), the demand model (3) can be expressed in observational terms as:

$$M_i = \alpha(1 + \delta) + [\beta(1 + \delta) + T] P_i + u_i(1 + \delta) + b_i + e_i . \quad (4)$$

The observational model (4) may be rewritten as:

$$M_i = \alpha^* + \beta^* P_i + u_i^* . \quad (5)$$

The price coefficient in the observation model  $\beta^*$  is related to the true price coefficient by:

$$\beta^* = (1 + \delta) \beta + T . \quad (6)$$

The price coefficient in the observational model will be equal to the true coefficient,  $\beta$ , if the measurement error is independent of the treatment and true score ( $T = 0$  and  $\delta = 0$ ). Neither random error nor bias that is uncorrelated with true score or plan affects the price coefficient.

#### Effects of Random Measurement Error on the Precision of the Price Coefficient

The variance in true consumption that is not explained by the experimental plan is  $\text{Var}(u)$  from Eq. (3). The variance of  $\beta$  is the  $\text{Var}(u)$  divided by  $\sum (P_i - \bar{P})^2$ . The variance in measured consumption that is not explained by plan is  $\text{Var}(u^*)$ . From (4),

$$\text{Var}(u^*) = (1 + \delta)^2 \text{Var}(u) + \text{Var}(e) + \text{Var}(b) . \quad (7)$$

The variance of  $\beta^*$  is the unexplained variance in (7) divided by  $\sum (P_i - \bar{P})^2$ . The ratio of the variance of  $\beta^*$  to that of  $\beta$ , therefore, is:

$$\frac{\text{Var}(\beta^*)}{\text{Var}(\beta)} = (1 + \delta)^2 + \frac{\text{Var}(e) + \text{Var}(b)}{\text{Var}(u)} .$$

In the usual case where  $\delta$  is positive, random measurement errors increase the variance of the parameter estimate and hence decrease the precision of the estimate.

#### Measurement Bias Effects on the Price Elasticity Estimate

The price regression coefficient can be used to estimate the price elasticity:

$$\eta = \frac{dY}{dP} \cdot \frac{P}{Y} = \beta \frac{P}{Y} . \quad (8)$$

The estimated elasticity based on measured values  $\eta^*$  is<sup>12</sup>

$$\eta^* = \beta^* \frac{P}{M} = \frac{\beta(1 + \delta) P + TP}{(1 + \delta) Y + TP + b} . \quad (9)$$

Errors proportional to true consumption do not bias the elasticity estimate in the absence of other biases. That is, if  $b = 0$  and  $T = 0$ , the elasticity estimate (in (10)) is unbiased even though the errors were systematically related to true demand ( $\delta \neq 0$ ). However, if the measurement biases are related to the experimental plan ( $T \neq 0$ ), the estimated price coefficient and elasticity are not equal to the true population parameters.

If observational biases are independent of both plan and true demand ( $\delta = 0$ ,  $T = 0$ ,  $b \neq 0$ ), the elasticity coefficient is a biased estimate. The price coefficient is not affected by an additive measurement bias; however, the intercept in the linear model will be biased by an amount equal to the average value of  $b_1$ .

#### Measurement Design to Reduce the Bias

The conclusions about bias to be drawn from the above are that:

<sup>12</sup>The random error,  $e$ , has an expected value of zero. Therefore, it does not appear in the formulation.

1. Response and reactive biases that are correlated with experimental treatments (price) will bias the regression and elasticity coefficient estimators of the true relationship between price and consumption in a nonexperimental population.
2. Response and reactive biases correlated with true utilization will bias regression coefficients but, assuming proportional bias, will not affect the magnitude of the elasticity estimate.
3. Uncorrelated response and reactive biases will affect the elasticity and the estimate of the regression intercept but will leave the estimated magnitude of the regression coefficient unaffected.

These derivations suggest that the applied consumption measurement design give highest priority to minimizing measurement biases correlated with experimental treatments and to obtaining reasonably precise estimates of any remaining correlated biases. Resources should be devoted also to minimizing and estimating other biases so the experimental data can be more informative about applied questions depending on accurate estimates of means or the intercept of a regression model.

Measurement biases can be a result of design faults such as excluding one part of utilization (e.g., pre-operative physician visits) from the operational definition of health services and not seeking observations of it. Biases also enter at the data collection stage in the form of systematic reporting distortions (such as forgetting or reporting a single event several times). The HIS measurement design objective is to obtain a record of each use of health services for the data base and to insure that unmeasured consumption is not disproportionately concentrated in one or more of the treatment groups (not correlated with plan). Toward these ends, the scope of health services to be measured is broadly defined (see above discussion). Features of the experimental treatments, however, may introduce correlated measurement biases by providing smaller incentives to report utilizations in some plans relative to others. Preliminary studies investigating the likelihood of utilization reporting biases, therefore, were made.

The major reason that a patient and health service provider will report a health care transaction to the study is that the experimental insurance plan may pay all or part of the charges for that service. The amount of the total bill paid is determined by which experimental financing treatment has been assigned to the patient. Given the treatment, the percentage of charges reimbursed also depends on how much the family has spent on health care to date in the accounting year.<sup>13</sup> One of the unfortunate consequences of creating an experiment that varies the amount of the bill paid by insurance is that, in theory, this also varies the amount of incentive to the patient (and possibly the provider) to furnish utilization information. If reporting compliance is a function of the experimental treatment assigned to the patient, this will create the "correlated" measurement biases discussed earlier. They, in turn, will bias the estimated price and elasticity coefficients.

A primary measurement design question, then, is how large an underreporting bias, if any, is produced by the differential payment schedules resulting from the experimental treatments? If there is a correlated underreporting bias, can a system be devised either to eliminate it or, failing that, to estimate its magnitude? The question is of sufficient importance to warrant a pretest experimental study. The study and results are discussed in Sec. III.

The reduction of random measurement error is probably less important to analysis objectives than reduction of measurement bias. However, it is desirable to do what is possible within cost constraints to increase the precision of the regression coefficient estimates of the various price-quantity relationships by reducing random measurement error.

#### Measurement Precision

As part of the design of the HIS experiment (which also included determinations of sample size, number of experimental sites, and length of the experimental observation periods), the precision of utilization and expenditure measurement was estimated at the then-current state of

---

<sup>13</sup>The latter is not a consideration in the experimental treatment that pays all health care costs beginning with the first dollar.



the art as represented by the Center for Health Administration Studies (CHAS) 1963 survey of health services demand (see Andersen and Anderson, 1967). This single-interview cross-section survey of the U.S. population measured utilization and expenditure by asking household respondents to recall over a one-year period. With these estimates of means and variances, it appears that the Health Insurance Study using moderately sophisticated state-of-the-art analysis models,<sup>14</sup> will be able to obtain reasonably precise estimates of the price parameter both for total health care demand and for gross types of health care services such as physician office visits, prescription drugs, hospitalization, and dental care (Morris, 1973).

The Health Insurance Study should be able to measure health care consumption with somewhat less error than was done in the 1963 cross-section survey and thereby be precise enough to permit analysis of price and quantity demanded at more disaggregated levels of measurement. The proportion of measured variance that is random response variance in the CHAS data is unknown. However, surveys using similar procedures (such as the HIS first year baseline) contain varying amounts of random response error, ranging from very small amounts for hospital visits<sup>15</sup> to 40 percent or more for some of the service-specific, expenditure-per-person estimates.<sup>16</sup>

Two features of the measurement design, incorporated for other (higher priority) reasons, may enable analysts to reduce random measurement error effects at the analysis stage. One is that consumption is measured at very disaggregated levels; reaggregating the component measures to form more broadly defined dependent variables can result in the "derived" measures containing a lower proportion of random measurement error than do the components. This will be true to the extent that the component observations covary positively.<sup>17</sup>

---

<sup>14</sup>See discussion of analysis model options in Newhouse (1974a).

<sup>15</sup>For example, see Feather (1972).

<sup>16</sup>See Marquis, Marquis, and Newhouse (1976).

<sup>17</sup>Assume an index is to be constructed from two components, each measured with error but without bias. Then

The second feature (discussed in detail in Sec. IV) is that there will be more than one measure of the occurrence of each utilization. To the extent that the two measures of each service use can be incorporated into a single analysis strategy and that they covary, estimation precision can be increased. (For a discussion of related estimation theory, see Jöreskog and Goldberger (1975). Marquis, Marquis, and Newhouse (1976) illustrate a simplified version of this strategy in deriving a minimum variance "best estimate" of expenditures for dental care combining two different observations for each case.)

$$M_1 = T_1 + e_1 \quad \text{and} \quad M_2 = T_2 + e_2 .$$

$e_i$  is uncorrelated with  $T_i$  for  $i = 1, 2$ .

$$\begin{aligned} \text{Var}(M_1 + M_2) &= \text{Var } T_1 + \text{Var } e_1 + \text{Var } T_2 + \text{Var } e_2 \\ &\quad + 2 \text{Cov}(T_1, T_2) + 2 \text{Cov}(e_1, e_2) . \end{aligned}$$

The unreliability of each component measure is

$$U(M_i) = \frac{\text{Var } e_i}{\text{Var } T_i + \text{Var } e_i} . \quad (a)$$

Assuming  $\text{Cov}(e_1, e_2) = 0$  the unreliability, denoted by  $U$ , of the combined measure is

$$U(M_1 + M_2) = \frac{\text{Var}(e_1) + \text{Var}(e_2)}{\text{Var}(T_1) + \text{Var}(T_2) + \text{Var}(e_1) + \text{Var}(e_2) + 2 \text{Cov}(T_1, T_2)} .$$

To simplify, ignore the problem of differential weights by assuming that  $\text{Var}(T_1) = \text{Var}(T_2) = \text{Var}(T)$  and that  $\text{Var}(e_1) = \text{Var}(e_2) = \text{Var}(e)$ . Then

$$U(M_1 + M_2) = \frac{2}{2} \frac{\text{Var}(e)}{(\text{Var}(T) + \text{Var}(e) + \text{Cov}(T_1, T_2))} . \quad (b)$$

If the true scores do not covary, the unreliability of the sum (Eq. (b)) is the weighted average of the unreliabilities of the individual components (Eq. (a)). If the component true scores covary positively, then the denominator in (b) is greater than the denominator in (a) by  $\text{Cov}(T_1, T_2)$ . In this case the unreliability of the index score is lower than the average of the individual component unreliabilities and precision is increased. For a discussion of the general principle of increasing reliability by combining imperfect measures of components, consult Cronbach (1951).

#### AN ALTERNATIVE

A frequently raised measurement concern is that the analytical objectives have been allowed to drive utilization measurement well beyond the state of the art. Would it not be more appropriate to trim back some of these objectives and use measurement procedures that involve less risk of undetected biases or large amounts of response error?

At first it might seem that the most relevant information to questions about national health insurance would come from a series of demonstration projects, each simulating principal features of proposed NHI legislation. For this kind of research, the measurement of health care consumption need not be complex. The dependent variable would be the amount of money the demonstration project spent to reimburse people for health care services received and covered by the insurance scheme being tested. Policy planners would then know what each proposed NHI plan would cost taxpayers. The extent that persons used health care not covered by the demonstration plan, used services but did not ask for reimbursement, or used one kind of service in preference for another need not be measured. These kinds of behavior are part of the market mechanisms that would operate the same way under NHI.

If there were only a small number of NHI proposals to be considered and if the increase in demand created by enacting any proposal could be met easily by the existing health care delivery system, the demonstration project approach would be worth considering. However, as mentioned earlier, there are a large number of proposals to be considered, new proposals continue to be advanced, and it is far from certain that the current delivery system could meet changed demand implied by demonstration studies without changing prices, using other means to ration services, or radically reallocating manpower resources and changing other central features of the delivery system. Estimates of demand in these circumstances could be distorted considerably.

These and reasons discussed earlier require estimating demand facing the entire delivery system rather than demand facing the "insurance company." It is necessary to abstract the main NHI issues and investigate them using designs derived from economic theory. It is also

necessary to pay attention to potentially unreported utilization so demand can be estimated directly. A "next best" measurement strategy involves identifying the major measurement-based threats to the validity of price quantity parameter estimates. It requires designing procedures to minimize the threats, monitor them, and estimate any remaining biases. It involves a willingness to proceed sequentially with measurement design to achieve the required degrees of estimation accuracy and precision over the course of the study.

#### UNRESOLVED QUESTIONS ABOUT MEASUREMENT BIAS

The measurement uncertainty judged most important to resolve early is whether there are utilization measurement biases, and if so, are they correlated with the experimental treatments? If they exist to an important degree, can they be removed by incorporating additional features into the measurement design? This is the subject of the next section, in which some pilot research and its results are described. The results suggested both that measurement biases could occur and that correlated biases were present (either as underreporting in the plan with small incentives to report utilization or as reactive effects in the high incentive treatment). The pretest results suggested which of several tested design alternatives would be most effective in reducing the biases. The measurement design for the first year of the experiment, which includes reporting incentives, supplemental measurement and questionnaire design, is discussed in Sec. IV. The final section includes some preliminary results of the first year's experience with measuring utilization, a discussion of some changes made as a result of the evaluation, and the basic features of some additional measurement experiments to be conducted to reduce remaining uncertainties about data quality in the experiment.

### III. PILOT STUDY TO ASSESS NONREPORTING OF UTILIZATION

A pilot experiment was conducted for ten months before the main experiment. It was designed to obtain a preliminary estimate of biases in measuring utilization. Utilization is reported on a document closely resembling a standard health insurance claim form.<sup>1</sup> The pilot research was structured to discover whether all uses of health care were reported on these forms, whether underreporting was correlated with experimental treatment, and which of several supplementary reporting procedures (in addition to MERs) would be most effective in removing whatever measurement biases existed. It was suspected that utilization detected by MERs would contain an underreporting bias correlated with insurance treatment because incentives to report health services vary by insurance treatment.

A small sample of families was selected and assigned to one of two maximally different health insurance treatments. One of two supplementary measurement procedures, in which reporting was independent of incentives to file MERs, was assigned to families in each insurance treatment. Other supplementary reporting procedures were also used. All reporting documents were matched on a service-by-service basis. Estimates of the MER underreporting main effect were made by noting the number of utilizations reported on supplementary documents (and confirmed as true by telephoning providers) but unreported on MERs. Estimates of the measurement biases correlated with insurance plan were obtained by noting the difference between the number of utilizations reported on MERs and the (increased) number reported through the supplementary measurement system in which reporting was not directly contingent on financial incentives to respond.

#### RESEARCH DESIGN

A stratified random sample of 48 families residing in the Dayton, Ohio area were enrolled in the experiment. All remained with the study

---

<sup>1</sup> Called a Medical Expense Report (MER).

during the ten-month pretest phase.<sup>2</sup> The families were actually enrolled for three years.

A family consists of a self-proclaimed head and any or all of the following who reside in the same household: a spouse; persons under 18 years of age related to the head or spouse by blood, marriage, adoption or economic dependency; and adults (18-62) dependent on the head or spouse for more than half of their financial support. Families with annual incomes above \$25,000 are excluded. Persons who are in the military, eligible for military or Veterans Administration medical care, over 17 and full-time students residing at school, in a correctional institution, or receiving inpatient health care at a state-supported institution are excluded.

#### Experimental Insurance Treatments

Thirty of the families (109 people) were enrolled in "generous" insurance plans, which covered a wide variety of health care services (the major exclusions were adult dental care and orthodontia treatments for children). Most of these families received complete reimbursement for all health care expenditures, and a small number were subject to 20 percent coinsurance for inpatient hospital care until the total out-of-pocket family expenditures reached 5 percent of income (or \$1000 if that was lower) during a 12-month accounting year.

The remaining 18 families (52 people) were assigned to a less generous insurance plan called the deductible plan. The same broad range of services were covered; however, families faced a deductible of 15 percent of income (or \$1000 if that was lower) before the experimental insurance paid all subsequent remaining costs of covered services within the one-year accounting period. Families who were potentially less well off with the experimental insurance than with their previously held insurance were paid a cash "participation bonus" in monthly installments to make up the difference. The direct financial incentive to report

---

<sup>2</sup>Two additional families were tentatively enrolled pending their willingness to assign the benefits of their current, employer-furnished health insurance to the study. Both families eventually decided not to enroll in the study. No utilization data were collected from them and they are excluded from the analysis.

utilizations on MERs was highest in the generous treatment and lower in the deductible treatment. For the generous insurance, all charges for covered services listed on MERs were paid by the study. Listed charges for the deductible group were not paid by the study until the total charges in the accounting year exceeded the family's deductible. The families were assigned to insurance treatments such that the groups were comparable across a number of economic, demographic, and health dimensions (measured by personal interview several months before enrollment).

#### Utilization Reporting Procedures

Each patient was asked to have his provider (physician, hospital, pharmacist, etc.) submit a medical expense report (MER)<sup>3</sup> for each use of a health service, regardless of whether the experimental insurance paid for the service. A properly completed MER was the basis for cost reimbursement to the provider or the patient by insurance if applicable. Providers were not paid an additional fee for completing the MER.

Weekly Supplemental Reports. The main pilot study measurement question was whether all uses of health care services are reported on the MERs. Needed was a way of determining the number of services used by each participant that was either completely accurate or contained measurement biases independent of those in the MER system. One candidate was an independent check of provider records. Local area providers (or a sample) would be asked to search their records and report treatments and other services rendered to participants in the pilot study. The provider record check possibility was ruled out because of its complexity and high cost and the unknown effects it might have on provider cooperation in the site during the main experiment period.

The primary alternative adopted was an independent weekly report<sup>4</sup> of outpatient utilization from the families themselves. A short, easily

---

<sup>3</sup>Three different types of MERs were used: Inpatient Hospital, Pharmacy, and Outpatient. These reports are similar in format and content to MERs discussed in the next section.

<sup>4</sup>Available from the Health Insurance Study, The Rand Corporation, 1700 Main St., Santa Monica, Ca., 90406, by requesting the Weekly Information Report.

filled out questionnaire was devised on which a family respondent indicated the family's use of services of "physicians, dentists, psychologists, chiropractors, osteopaths, etc." during the past week (reports of inpatient care and pharmacy purchases were excluded). The blank report was mailed to the household weekly. This served as a stimulus or reminder to fill out the questionnaire. The family was paid \$2.50 for returning the completed form on time. It was felt that families could and would report the bulk of health services received on the questionnaire. Also it was assumed that the weekly mailing and generous payment would be sufficient to maintain their cooperation over the pilot test period. The incentives and difficulty of reporting on this questionnaire were the same across the insurance treatments. Half of the families in each insurance plan treatment were asked to submit the weekly reports.

Monthly Supplementary Reports. Another set of supplementary measurements are the monthly reports.<sup>5</sup> The families not asked to complete weekly reports were sent a monthly list of all outpatient visits detected by MERs and Visit Notices (see below). Families were asked to inspect this list and to write in any additional (undetected) visits made by family members during the one-month reference period.<sup>6</sup>

For measurement evaluation purposes, utilization rates of the groups who made supplementary weekly reports are compared with rates for the groups who made only monthly reports.

The Visit Notice. A more direct check for "missing" MERs was a form called the *Visit Notice*. A person receiving outpatient services was asked to mail in a Visit Notice whenever an outpatient MER was left for the provider to complete and mail to the study. The Visit Notice was attached to the outpatient MER. The patient filled it out, tore it

---

<sup>5</sup> Available by requesting Document HIEI 753 from the Health Insurance Study.

<sup>6</sup> The families reporting utilization at weekly intervals were also sent a list at four-week intervals of outpatient visits they had reported on the weekly questionnaire but for which neither MERs nor Visit Notices had been received. Contrary to the monthly reports, the family was not asked to report additional utilizations on this form. Hence, it is not considered to be a measuring instrument in the following discussion.



off, removed some tape, folded and sealed it, and mailed it (postage paid) to the study. There was no experimental variation in the use of Visit Notices. All families were asked to submit them for each outpatient health care service received (except purchases of pharmacy items and supplies).

#### Summary of Measurement Experiment Design

The basic features of the pilot measurement experiment are shown in Fig. 1. The  $2 \times 2$  factorial design varied the insurance treatment (generous, deductible) and the method of obtaining supplementary reports of outpatient services (weekly, monthly). In addition, all groups were asked to report utilization on MERs and Visit Notices.

Supplementary measurement treatment	Insurance treatment	
	Generous	Deductible
Weekly reports	16	8
Monthly reports	14	10

Fig. 1—Pilot measurement experimental design

(The entries in the cells indicate the number of families assigned to each treatment.)

#### Procedures for Matching Utilization Reports Across Reporting Forms

For inferences to be made about measurement bias, it is necessary to determine on which reporting forms a particular health service use was or was not reported. A set of matching rules and procedures was established for this purpose.

A single outpatient encounter is supposed to be reported on each of three forms: a MER, a Visit Notice, and (as applicable) a weekly or monthly form. "Match rules" were used to judge when the same encounter was reported on the different forms.

The unit of analysis for matching is the outpatient "encounter," consisting of a patient's single visit to a provider or receipt of an

outpatient professional service (e.g., laboratory analysis of blood and urine samples) on a single date. Multiple encounters by one patient to the same provider at the same place of service on a single day are combined into one match analysis unit. A reported encounter is judged to match if three criteria were met.

1. The name of the patient is the same on both forms. (In operational terms, the identification number of the patient must match across the forms. To preserve the privacy of participants, their names are not available to data analysts.)
2. The name of the provider of health services is the same on both forms. (Provider identification numbers are actually used rather than names. Some providers have more than one identification number because they practice in more than one location or for other reasons. A "match" of any of these identification numbers is allowed.)
3. The dates of service or other document dates are within specified time intervals.<sup>7</sup>

#### Analysis Methodology

The analysis is reported in two parts. The first is nonexperimental. An estimate is made of the difference between true outpatient utilization and utilization reported on MERs. It is assumed that a true outpatient encounter occurred if one of three conditions were met: It was reported

- 
- <sup>7</sup> a. MER and Visit Notice: Date of service on the MER is within  $\pm 10$  days of the date of signature on the visit notice.
  - b. MER and weekly report: Date of service on the MER is within the one-week reference period covered by the weekly report or within one week on either side of the weekly reference period.
  - c. MER and monthly report: The MER date of service is within the one-month reference period of the monthly report or within ten days on either side of the reference period month.
  - d. Weekly and Visit Notice: The date of signature on the Visit Notice is within the one-week reference period for the weekly report or within one week on either side of the weekly reference period.
  - e. Monthly and Visit Notice: The date of signature on the Visit Notice is within the monthly report reference month or within ten days on either side of the reference month.

on an MER, it was reported on two supplementary reporting documents, or it was reported on one supplementary document and confirmed by telephoning the provider of the service.

For the second part of the analysis, measurement bias correlated with experimental treatment is inferred by comparing relative effects of the weekly and monthly reporting treatments for each insurance treatment with respect to total MERs submitted and total outpatient utilization reported on all instruments.

The proportions and standard errors mentioned are based on ten months of observations using the person as the unit of analysis. Persons are members of families who may share common utilization patterns. Therefore, the standard errors are corrected for intrafamily correlation.

The participating pilot families were selected from a much larger, geographically clustered sample. The amount of geographic clustering in the pilot sample was felt to be negligible, so the standard errors are not corrected for any sample clustering effects.

Some reported encounters are deleted from the analysis:

1. All encounters taking place outside of the ten-month pilot study period (January through October 1974).
2. Purchase of pharmacy items.
3. Inpatient hospital services and visits.
4. Encounters reported with a missing or invalid person identification number.
5. Free care reported on the monthly form (for which provider identification information was not asked).
6. Encounters involving the same person, same provider, same place and date of service and reported on the same form were combined into a single encounter for analysis.
7. For the second set of analyses, one person with extreme dependent variable (encounter) values is deleted.

## RESULTS

The pilot data suggest two things about the measurement quality of MER information: It is incomplete, and the measurement bias is probably

correlated with the experimental treatment. The results also suggest that the independent weekly reporting procedure may go a long way toward removing the observed correlated measurement biases.

#### Nonresponse to Supplementary Utilization Reports

The cooperation of the pilot families in filling out the weekly and monthly supplementary reports of outpatient visits was good. Weekly response rates averaged 98 percent, monthly response rates averaged 97 percent.<sup>8</sup> There was no followup effort to obtain the missing reports.

#### Sufficiency of the Expense Report Measurement Method

Have participants obtained outpatient health services that are not reported on the Medical Expense Reports (MERs)? The pilot data indicate that the answer is "Yes" and that the number of health services not reported on MERS is not trivial. This conclusion is based both on the empirical results and on several assumptions. One assumption is that at least some of the utilizations reported on the backup forms but not on MERs did take place. To estimate underreporting rates, it is assumed that the likelihood of the supplementary report being valid increases if the report can be confirmed by another source. A final caveat is that the MER underreporting estimates are probably lower-bound estimates to the extent that some utilizations occurred that were not reported on any of the forms.

Table 1 shows on which reporting forms the detected utilizations were obtained for the pilot families. The data are given in terms of "encounters," which are one-time uses of such health services as a visit to a physician, a visit to a clinic, a visit to a dentist, or laboratory services. Excluded are purchases of major health supplies (eyeglasses, hearing aids, etc.); inpatient hospital and physician services; and purchases of medicine, sundries, and other health care supplies. None of the excluded items was queried on the supplementary reporting forms.

---

<sup>8</sup> At the end of the tenth month, all pilot families were converted to another supplementary system of reporting utilizations. Instructions about completing the last monthly report were ambiguous and many were not returned. Therefore, the monthly response rates mentioned above are for the first nine months of the ten-month pilot experiment period. Weekly rates are for the full ten months.

Table 1

DISTRIBUTIONS OF REPORTED OUTPATIENT HEALTH CARE  
ENCOUNTERS BY SOURCE OF REPORT

Source of Report	Frequency	Percent
Medical Expense Report	740	69 <sup>a</sup>
Visit Notice only	185	17
Supplementary report only (weekly or monthly report)	103	10
Visit Notice and weekly or monthly report, not MER	40	4

<sup>a</sup>On MER only 18 percent; on MER and other  
source, 51 percent.

Of the 1068 outpatient encounters detected, 69 percent are reported on the claim forms (the MERs) and the remaining only on the other instruments. Each of the supplementary forms produced a substantial number of unique encounter reports. If these reports are valid, the measurement of outpatient health service use inferred only from claim forms (MERs) is biased.

The validity of the utilizations reported only on supplementary forms is examined next. Some of the reports are found to be valid, so an analysis of the nature of the biases is also made.

Validity of Utilization Not Reported on MERs

Some evidence relevant to the validity of encounters reported only on supplementary forms comes from two small sample telephone verification studies conducted early in the pretest period. Providers' offices were asked if the outpatient service reported on only one supplementary report (a Visit Notice or Weekly Report) took place. These small sample studies indicate that many of the encounters tested could be "verified."

Early in the pilot period (about the fourth month) there were 36 encounters reported only on the weekly forms. The physician's office (dentist's office, clinic) named in each was called and asked about the outpatient encounter reported. The results are shown in Table 2. Eight

Table 2

FREQUENCY DISTRIBUTION OF VERIFICATION RESULTS OF  
ENCOUNTERS REPORTED ONLY ON THE WEEKLY FORM

Verification Result	Frequency
Confirmed, MER located in study files	8
Confirmed, provider did not submit MER <sup>a</sup>	7
Confirmed, patient did not furnish MER	5
Not confirmed	14
Not contacted	<u>2</u>
	36

<sup>a</sup>In three of these seven cases, the provider did not charge separately for the visit in question nor did he indicate that the visit took place on the MER submitted for the entire treatment.

clerical mistakes by the study were discovered (indeed, the MER had been sent in but was not matched correctly to the weekly report). In seven more cases the provider acknowledged the visit but had not submitted the MER. (In three of these cases, there was no additional charge for the visit and, through misunderstanding, the provider had not intended to submit a MER. In another one of these cases the provider had submitted his bill to the patient's previous insurance company.) In five cases the provider said the visit took place but the patient had not furnished an MER to be filled out. In the remaining 14 cases, the provider said the patient had not been seen during the study period. Ignoring the clerical errors and the noncontacted cases, the results suggest that 46 percent  $(7 + 5)/26$  of the encounters detected only on weekly forms are valid and 54 percent are misreported.<sup>9</sup>

Five outpatient visits<sup>10</sup> reported only on visit notices during the first three months had not been confirmed by MERs in the sixth month.

<sup>9</sup>These could be patient overreports, provider underreports, or "true" visits which could not be verified due to misreporting or erroneous recording of details such as patient name, provider name, or date of service.

<sup>10</sup>There were probably quite a few more than five utilizations reported only on Visit Notices during the first three months. However, very "liberal" criteria were used at that time to decide whether an

The study was able to contact both the participant and the provider in four of the five cases with the following results: Two of the visits did not take place, and two did occur but the MER was not filed. In one of the latter cases, the clinic refused to complete the MER. After several attempts by the patient to get the clinic to comply, the patient sent the study a duplicate copy of the bill.

These small scale studies suggest about half of the encounters reported on only one supplementary form did take place. This estimate can be used to make some preliminary inferences about the efficacy of measurement procedures.

It is also assumed that all encounters reported on MERs are valid and that an encounter reported on both the Visit Notice and the weekly or monthly report really did take place.

These assumptions are put together in Table 3; an MER-only reporting system apparently would detect up to 80 percent of outpatient health care encounters. If the supplementary reports remind participants to submit MERs, the 80 percent is overestimated in proportion to the "re-minding" effects.

The validation data are from very small samples, and the conclusions are subject both to sampling error and to the correctness of the inferential assumptions. The conclusions suggest:

1. Supplementary reporting procedures are needed for the main study to reduce underreporting bias in the MER measurements.
2. Contacting providers about encounters reported only on supplementary forms might yield an MER or other confirmation for half of the extra encounters.

---

encounter reported on an MER matched an encounter reported on a Visit Notice. As a result, the false positive match rate was quite high (yielding 3-1/2 times the number of matches of Visit Notice to MER found on reanalysis. In the discussion that follows, it is assumed that the provider confirmation rate of Visit Notice only reports identified using liberal match rules would also be obtained for Visit Notice only reports identified using the stringent match criteria described earlier. This appears to be a conservative assumption and may result in a slight underestimate of the proportion given at the bottom of Table 3.

Table 3

ENCOUNTERS PRESUMED TO BE VALID BY SOURCE OF REPORT

Source of Report	Percent Assumed Valid	Number of Encounters Presumed Valid <sup>a</sup>
MER	100	740
Both Visit Notice and weekly or monthly report	100	40
Only one supplement	50	<u>144</u>
		924
Percent of presumed valid encounters detected only on supplements	20	

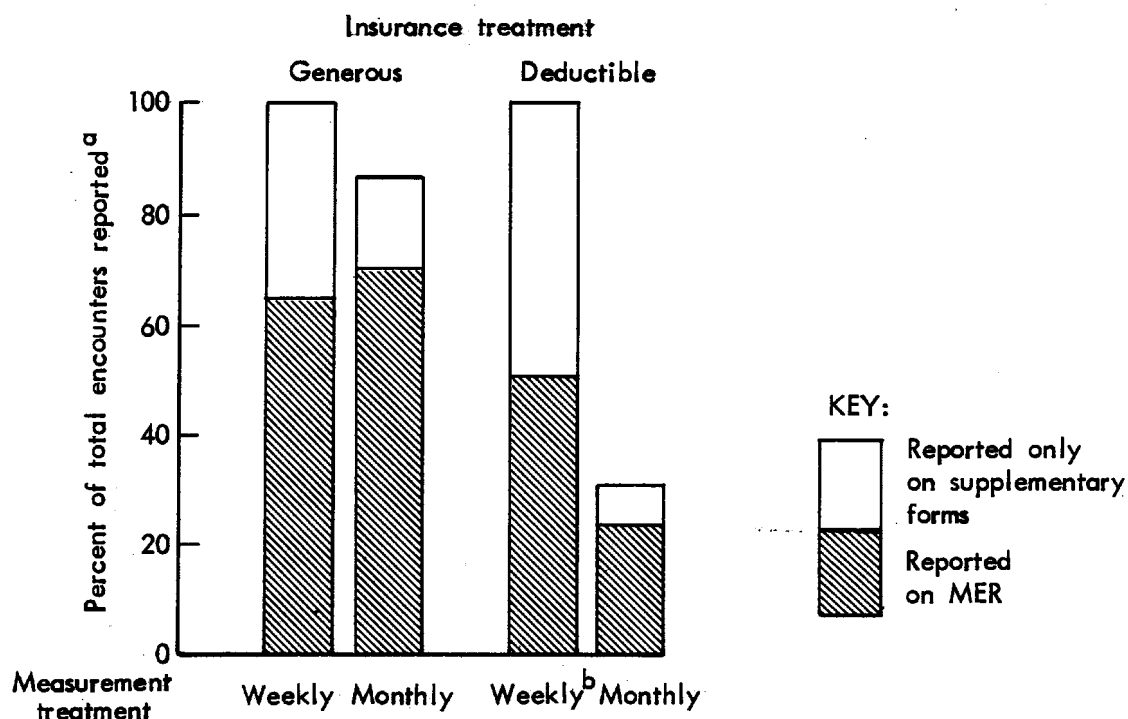
<sup>a</sup> Calculated as percent valid times observed frequency (from Table 1).

It was shown in Sec. II that the effects of measurement bias varied with the form the bias takes. A bias in dependent variable measurements (use of health services) that is correlated with experimental treatments (insurance plans) affects a variety of policy-relevant estimates. The possibility that the measurement biases, discussed above, are correlated with insurance treatment is examined next.

Measurement Bias Correlated with Experimental Treatment

The data collection methods experiment conducted with the pilot families provides some information about the nature and extent of utilization reporting biases correlated with the experimental treatments. The effect of the weekly supplementary measurement in detecting utilizations within each of the insurance treatments is seen in Fig. 2, which shows the percent of encounters detected by MERs and supplementary reporting forms within each of the experimental insurance plan treatments. To prevent premature disclosure of price elasticity results, the raw data have been adjusted: Within each insurance treatment, the total encounters detected for the weekly reporting group are set at 100 percent, and the encounters detected for the monthly reporting group are shown relative to that. Encounters reported on MERs are also shown relative to total encounters reported by people in the weekly treatment.





<sup>a</sup> Total encounters for weekly reporting method for each plan equal 100%.

<sup>b</sup> Excludes one person with a large number of encounters, mostly reported only on MERs.

Fig. 2—Percentage of outpatient health service encounters reported by insurance and measurement treatment

Within each insurance treatment, the percent of encounters reported by the weekly group should equal the percent reported by the monthly group if the weekly reporting treatment has no effect on measurement. Given the balanced sample allocation procedures, there is no a priori reason to suspect that the groups differ in true rates of utilization. The results in Fig. 2 and Table 4 indicate, however, that weekly reporting increases the number of encounters reported on MERs and on all forms within the *deductible* plan; it has no important (or statistically significant) effect on encounters reported within the *generous* plan. That is, the measurement method interacts with plan in producing

Table 4

T-STATISTICS, EFFECTS OF THE MEASUREMENT  
TREATMENTS ON REPORTING

<u>Within Generous Plans</u>	<u>t<sup>a</sup></u>
Number of MERs per person	-0.30
Total encounters per person	0.73
<u>Within Deductible Plan</u>	
Number of MERs per person	1.93 <sup>b</sup>
Total encounters per person	2.70 <sup>b</sup>

<sup>a</sup>The t-statistics are computed using the observed (unadjusted) means and variances as:

$$\frac{(\bar{X} \text{ weekly} - \bar{X} \text{ monthly})}{\sqrt{SE_{\bar{X} \text{ weekly}}^2 + SE_{\bar{X} \text{ monthly}}^2}}$$

The means and variances are estimated from 10 months of data. The person is the unit of analysis. Variances are adjusted for an intrafamily correlation of approximately 0.1.

<sup>b</sup>Degrees of freedom = 49;  $p \leq .05$  (one-tail).

observed utilization rates, and both the price coefficient and elasticity calculated will depend upon the choice of which measurement method is used.

Care should be taken in interpreting these pilot results since there appear to be three possibilities. One interpretation is that the weekly probes are needed to detect utilization in the deductible plan where there is little or no direct financial incentive to report health care encounters on MERs (the families do not get reimbursed for reported health care expenditures below the deductible, and providers are not paid for furnishing the data). According to this interpretation, the weekly probes remind the families who need reminding (because of the nongenerous insurance coverage) to deliver MERs to providers and persuade the providers to furnish the data. The weekly probes also detect a large number of encounters that are not reported on MERs, which implies they aid recall of utilization for which MERs were not submitted.

Another interpretive possibility concerns sampling biases and errors. Despite careful allocation of families to the insurance-reporting method treatment groups in an unbiased (balanced) manner, it may be that the groups do differ on some unmeasured variable that has a large effect on utilization rates.<sup>11</sup> The unobserved variable (sampling bias) hypothesis is not subject to testing except by repeating the study on a new sample.

A third hypothesis is that the measurement methods were unbiased in their detection of utilization but that the weekly reminder probes may have artificially stimulated (true) utilization of health services in the deductible plan. This would be a Hawthorne or "reactive" effect.

The initial hypothesis about a "reactive" measurement effect was that if it were present, it would be the same in both the generous and the deductible plans (e.g., weekly probes remind people in each group to get an annual checkup, which might otherwise be forgotten or postponed) or that stimulation effects would be greater in the generous plan where dollar cost constraints on an increase in the propensity to consume health services are minimal.

The observed effects are contrary to these initial hypotheses (the method effect of weekly probes was observed only in the deductible insurance group). However, the observations are consistent with a variant of the Hawthorne effect hypothesis, which states that generous plan families consume all the health services they possibly can (consuming more would have negative benefits to them) so that an artificial stimulation effect of weekly supplemental reporting can only increase utilization for those who are not close to this maximum consumption threshold.

Since there is no clear way to choose among the three hypotheses, the year-one measurement design incorporated a further test of reporting form methods effects on observed rates of utilization. The experimental contrast during the first year was between a weekly and a biweekly reporting treatment. This affords an opportunity to partly replicate the

---

<sup>11</sup>The plan-by-treatment groups are balanced with respect to measured variables such as age, sex, welfare status, number of family heads, and family income.

small sample experiment on a much larger group. Other record check experiments are planned comparing biweekly supplemental reporting with no supplementary reporting to pin down the nature of the methods effect (e.g., Hypothesis 1, the reporting aid effects; or Hypothesis 3, the Hawthorne effect; or some combination) and to obtain more precise estimates of the form and magnitude of any remaining bias so that appropriate adjustments can be made to the consumption data at the final analysis stage.

#### OTHER INFORMATION FROM THE PILOT STUDIES RELEVANT TO MEASUREMENT DESIGN

The 48-family experiment, described above, provided some information about the completeness of information reported on those Medical Expense Reports that were submitted. The frequency and nature of item nonresponse on submitted MERs is evaluated below.

Halfway through the pilot study, the general nature of the pilot experiment results was already apparent. At that point, work began on designing the supplementary measurement instruments that would be needed for the forthcoming main experiment. (These instruments are described more fully in the next section.) The new instruments were pretested, using a new sample of families for several months before the start of the actual Health Insurance Experiment. Experiences with the pretest form and item nonresponse are described here and the implications for design of the supplementary methods of measuring utilization to the main experiment pointed out.

#### Completeness of MER Data

After the first two months of the pilot experience it was possible to assess how conscientiously physicians and other providers of health care were furnishing data on the MER questionnaire. An MER was considered incomplete if it did not contain sufficient data for the administrative processing of a payment claim or if key research items were missing. Data in Table 5 indicate that, at the end of the first two months, approximately one-third of the MERs received were incomplete. The few reports received from hospitals were all complete, but one in three of those for outpatient services needed clarification of some

Table 5

PERCENT OF MEDICAL EXPENSE REPORTS WITH  
INCOMPLETE DATA BY TYPE OF MER

Type of MER	Number of MERs	Percent Incomplete
Physician/dentist/supplier	73	38
Hospital inpatient	3	0
Pharmacy	<u>23</u>	<u>26</u>
Total	99	34

important item. Among those with incomplete information, an average of 1.8 pieces of important information were missing. The most common missing items on the "Physician-Dentist-Supplier" MER were type of treatment, diagnosis, whether condition was chronic or acute, disaggregation (itemization) of charges, and quantity of prescribed medicine. On the positive side, the basic information about the type of provider, the identity of the patient, the date of service, and the total charges was furnished completely. This information is sufficient for many of the planned analyses. It was apparent, however, that followup would be required to obtain some of the descriptive information needed to estimate demand using the more complex analysis models.

The Pharmacy MERs also contained missing answers. Diagnosis, name of item, strength, form, and quantity of medication were often not provided. Again, such basic information as prescribing physician, date sold, and total charge was usually complete, but some form of recontact with the data source would be needed to obtain the detailed information required by some of the planned analyses.

Item nonresponse was correlated with insurance plan as can be seen in Table 6. Incomplete MER data were more prevalent in MERs for patients in the deductible insurance plan than in the generous plan. To avoid estimation biases that might result from missing data, it will be necessary to edit MERs and followup item nonresponse during the main experiment.

Table 6

PERCENT OF MEDICAL EXPENSE REPORTS WITH  
INCOMPLETE DATA BY TYPE OF INSURANCE

Type of Insurance	Complete	Not Complete	Total Percent
Generous	70	30	100
Deductible	58	42	100

Pretest of a Revised Supplementary Questionnaire  
to Measure Utilization

During the late summer of 1974, 15 families were enrolled in a control group. One purpose was to gain experience with the enrollment, administrative, and measurement problems involved with families who were not offered experimental insurance plans (they remained under their existing health insurance coverage, if any). Families were asked to have their providers fill out and submit MERs for each health service use. Families were not reimbursed for the expenses listed on MERs, nor were providers compensated for furnishing the information.

The measurement goal was to gain experience in the use of the supplementary utilization reporting questionnaire designed to detect utilization that might not be reported on MERs. The questionnaire, a biweekly Health Report, was an independent document on which families reported utilization and other health information<sup>12</sup> every two weeks. It was a compromise between the weekly and monthly systems, using probes to remind respondents of health care received but on a two-week recall basis. Families were paid \$4 every two weeks for filling it out and mailing it in on time. The Health Report is described in more detail in the next section. This brief pretest sought to discover any problems respondents had furnishing information (e.g., item nonresponse) and potential rates of noncooperation (e.g., form nonresponse). Eight weeks of data were evaluated before procedures were specified for the final year-one measurement design. Of the 15 families who accepted

<sup>12</sup>Work loss, other restricted activity, and time spent obtaining health services.

enrollment in the control group, 14 remained with the study for the eight-week pilot test period.

During the eight-week test period, problems with the Health Report became evident: 60 percent of the families (9 of 15) either did not mail in a report at all, did not mail one in on time, or submitted an incomplete report. Four of the 15 families repeated their noncooperation on successive reports. The most common problems appeared to be a lack of understanding of how to answer the questions. Clearly the form needed much more explanation when it was presented at enrollment. Also, some sort of followup and re-explanation procedure was needed to teach families how to respond to the questions when they had information to furnish. Finally, some procedure seemed warranted to encourage families to submit the reports and submit them on time. The two-week recall period may be too long to minimize recall error under normal circumstances, and lengthening that recall period for delayed reports might attenuate even these advantages. These observations led to further specification of the year-one Health Report Edit and Followup systems described in the next section.

#### IMPLICATIONS OF THE PILOT RESULTS

The measurement design implications of the pilot study results are that supplementary measurement efforts are needed to detect uses of health services not reported on MERs and to reduce measurement error correlated with them. However, a backup system will introduce apparent overreports and might stimulate utilization. The most effective backup system for removing underreports appears to be the use of weekly supplementary reports, but cost considerations suggest a biweekly system be tested and used if reporting error is not much worse than under a weekly procedure. If the expanded biweekly report is to be used, respondents may need special training in how to fill it out correctly. To reduce overreports and increase precision, utilizations reported on in the supplementary questionnaire need to be queried further to separate false positive and false negative reports. This may result in contacting patients and providers about as many as 20 to 30 percent of initially detected encounters. Data furnished on MERs that are submitted are often

incomplete and the item nonresponse is also correlated with insurance plan treatment. Health care providers will have to be recontacted to complete the data collection for up to one-third of the outpatient MERs.



#### IV. FIRST YEAR DESIGN FOR MEASURING HEALTH CARE CONSUMPTION

The research objectives, concern about measurement error and bias, pilot study results, and findings from other methodological research were combined with cost considerations to determine the features of the system used to measure utilization in the first year of the HIS experiment.

##### OVERVIEW OF THE MEASUREMENT DESIGN

A schematic representation of the data collection forms, incentives, and procedures used to measure utilization in the first year of the Health Insurance Experiment is given in Fig. 3. Two sets of data collection forms were used: (1) The Medical Expense Reports, containing detailed information about the characteristics of each health service use, and (2) the supplementary reporting forms consisting of Health Reports (on which families made independent reports of utilization every week or two weeks) and the family recordkeeping forms (calendar and visit records) on which the family accumulated information about utilization to be transferred to the Health Report. The recordkeeping forms were not submitted to the study.

The family received \$2 per week or \$4 every two weeks for submitting Health Reports. The charges for health care submitted on MERs were paid by the study subject to the deductible, coinsurance, and coverage provisions of the experimental insurance treatment assigned to the family.

The data forms were edited for completeness and consistency and the respondent (provider or family) asked to furnish any critical missing information. An attempt was also made to obtain missing Health Reports.

Next the utilization information on the Health Report and MER was matched and an attempt made to locate an MER for each utilization mentioned on the Health Report. Families and providers were recontacted in an attempt to obtain an MER for all unmatched utilizations mentioned on the Health Report.

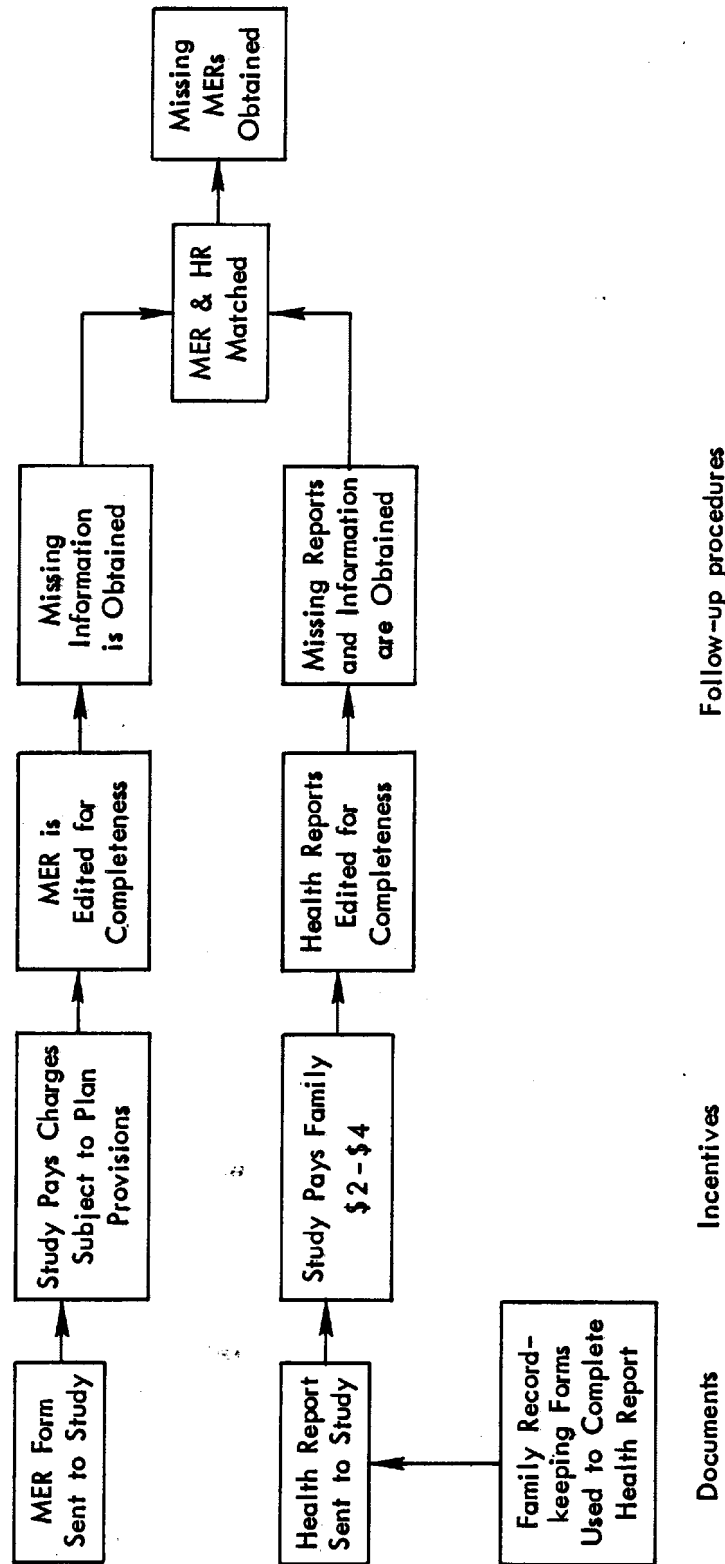


Fig. 3— Documents, incentives, and follow-up procedures used to measure consumption in the first year of the experiment

#### SUPPLEMENTARY REPORTING SYSTEM

Families reported their use of health service every week or every two weeks<sup>1</sup> on the Health Report. The information requested was sufficient to determine whether an MER had been received for the utilization and, if not, to try to obtain the MER from the provider or family.

The objective of the Health Report was to obtain a brief record of each health-related good or service obtained by participating families. The design task was viewed as creating a reporting system in which the respondent understood what was wanted, was able to furnish the information, was willing (motivated) to comply, and could be given immediate feedback about inadequate performance. Record check studies of utilization reporting suggest that response errors and possibly response biases can be fairly large for this kind of reporting task.<sup>2</sup> Previous methodological studies suggest that exercising design options to simplify response requirements, shortening recall periods, using written records, and developing highly structured questions (rather than broad, open-ended queries), detailed definitions, instructions, and examples would reduce both systematic biases and random reporting mistakes.

A supplementary reporting system consisting of a highly structured questionnaire and supplementary recordkeeping forms was developed and used during the first year of the experiment. Families were paid fairly generously for complying with the extra reporting requirements and were contacted by telephone, letter, or in person if they failed to submit the reports or failed to answer the questions properly. The basic features are described more fully below.

#### Health Report<sup>3</sup>

The keystone to the supplementary measurement system for detecting

---

<sup>1</sup>The experimental variation in the length of the supplementary report reference period is discussed at the end of this section.

<sup>2</sup>For example, see Feather (1972); Cannell, Fisher, and Bakker (1965); Cannell and Fowler (1963); Sudman, Wilson, and Ferber (1974); and Parry, Balter, and Cisin (1971).

<sup>3</sup>Available from the Health Insurance Study by requesting documents HIEI 35 (biweekly) and HIEI 51 (weekly). The reports were reprinted

health service consumption was the Health Report. It was a self-administered questionnaire in which families reported all uses of health services by recall and by transcribing from intermediate records. The family was free to select the Health Report respondent. The report was mailed to families periodically as an added stimulus to remember to fill it out and mail it in. Families received a \$2 or \$4 payment for each report submitted on time.<sup>4</sup> The response information was edited closely during the first two months. Families having difficulty responding (or not responding) during this time were contacted and "retrained" in the procedures for supplying requested information. After 3 months, the edit requirements were relaxed somewhat. The Health Report instrument and family recordkeeping forms are described next. Then the related topics of Health Report edit, Health Report followup, and location of missing MERs are discussed.

Health Report Content. The Health Report was an 11-page, self-administered questionnaire filled out by a member of a family participating in the HIS every week or every two weeks (the experiment with reference period lengths is mentioned in more detail below). The questions concerned the various types of health care services a person might have received during the reference period, such as the purchase of prescribed medicines or other items, visits to a doctor's (dentist's) office or clinic, visits to an emergency room or infirmary, medical care received at home from a doctor or nurse, overnight hospital (skilled nursing facility, etc.) stays, and telephone calls for medical advice.<sup>5</sup>

For each utilization, the respondent reported the name of the patient, the name and address of the provider, and the date of service (or

---

toward the end of the first experimental year as document numbers GS&A HR-1 (weekly, experimental), GS&A HR-2 (weekly, control), GS&A HR-3 (biweekly experimental), and GS&A HR-4 (biweekly, control).

<sup>4</sup>In practice, the family also received payment for an occasional late Health Report, but payment was withheld if reports were routinely submitted late. Payment was not withheld if the report was submitted on time but contained item nonresponse.

<sup>5</sup>There were also questions about health-related restricted activity and about delays in getting appointments, travel time, waiting time in the office or clinic, and total treatment time for each outpatient visit. These questions are not discussed here since they are not part of the measurement of the kinds and amounts of health services consumed.

purchase). For MER tracking purposes several other kinds of information were sought: For outpatient utilization, the respondent was asked whether an MER was left with the provider; and for prescribed items, the prescription number and name of the item were requested. For inpatient stays, the dates of admission were obtained as well as the names of all attending physicians. In the fee for service system, physicians usually bill separately (not directly through the hospital) so separate MERs were needed.

A good deal of care was taken in formatting the Health Report so that, wherever possible, the responding task was made as straightforward and comfortable as possible. A primary consideration was to ask specific questions about each category of service rather than for a list of all the goods and services consumed. Specific questions acted as probes to stimulate recall of forgotten utilizations and to provide aids to the family respondent when asking other family members about utilization during the reference period.

A second feature was the use of screening questions at the beginning of each Health Report booklet. A general question was asked for each kind of utilization and activity restriction of interest. If the screening question was answered affirmatively, the respondent was directed to the appropriate section of the booklet in which to report the details of the utilization or activity limitation.

Question construction was guided by the principle of disaggregating the desired information into discrete, small, easily reportable units and asking a question about each unit. From one point of view this increases the response burden because more questions are asked. However, the amount of information the respondent must "process" to furnish each answer is reduced and answers are potentially more accurate and complete.<sup>6</sup> The respondent performed a longer series of simple tasks rather than a few complex response operations.

Further clarification of reporting requirements was attempted by providing additional instructions and examples for each detailed question.

---

<sup>6</sup>For an experimental demonstration of this effect, see Marquis, Marshall, and Oskamp (1972).

Health Report Procedures and Incentives. The Health Report was a retrospective reporting document. The family respondent summarized health care events that took place in the past week or past two weeks. To aid recall and reporting, the family was given intermediate record forms (described below). The family respondent collected the record-keeping forms from family members who used services, and transcribed the information onto the Health Report.

Blank Health Reports were sent by mail to families near the end of each reference period. The arrival of the report was a stimulus or reminder that it should be filled out. Periodic mailings at the end of the reference period are thought to minimize the chance of losing the form.

The HIS did not designate a respondent for the Health Report. It is felt that there is no single best respondent designation rule that would apply equally well to all families. Family units could designate whomever they wished to perform the reporting task, taking into account who among them has good "verbal" skills, adequate knowledge of family health care, adequate time for the task, and the proper degree of conscientiousness. For accounting purposes, the HIS did require the signature of one of the heads of the family to appear on each Health Report.

The family was paid if the report was signed and received by the Health Insurance Study within ten days of the end of the reference period. Families completing weekly reports were paid \$2 per report. Families filling out biweekly reports received \$4 per report. It is estimated that the average report took less than five minutes to complete, so the compensation was generous. Payment was made by check on a monthly schedule.<sup>7</sup> It was contingent on the timely (within ten days) receipt of the questionnaire, not upon the quality of the information furnished. There was one potential disincentive to accurate reporting. As described below, utilizations mentioned on the Health Report but for

---

<sup>7</sup>Families received monthly statements of payments made to them for fulfilling reporting requirements and in connection with other parts of the study (such as what payments were made for health care services covered by experimental insurance).

which no MER was received set off a search for the MER. Part of the search involved contacting the family to ask whether the MER really was given to the provider. If the family admitted it had not delivered the MER, it was asked to do so. No extra compensation for this was given to control families, to families below the deductible (who do not expect expenditures to exceed the deductible during the accounting year), or if the missing MER were for a noncovered service (e.g., orthodontia, or a second pair of eyeglasses this year). The disincentive was even greater if the family had to persuade a reluctant provider to furnish the data or if the provider charged the family a fee for completing the form. Thus, although the intent of the Health Report system was to provide a painless means of reporting all instances of health care consumption (with incentives to comply equal across experimental treatments), potential disincentives were correlated with the experimental treatment.

#### Family Recordkeeping Forms

The pilot research suggested that some uses of health services not entered on MERs could be detected only by the Visit Notice, a half page document attached to the Physician MER that the patient filled out when making an outpatient visit, tore off, and mailed in. On the basis of the utilization measurement methods literature, the apparent beneficial effect of the Visit Notice procedure is due to minimizing recall error in utilization reporting; the patient essentially reports the visit as soon as it occurs. The expensive aspects of the Visit Notice procedure (printed form with pressure sensitive gum, business reply postage) may not contribute significantly to reducing MER underreporting effects. Following this reasoning, the measurement system was redesigned to provide several recording forms that minimized the recall interval between the occurrence of an event and its written report. Instead of participants mailing in each record as the visit was made, they were asked to retain and periodically summarize all of the visits on Health Reports. The intermediate records are described here.

Calendar (Exhibit 1). Each participating family was given an 11" by 17" wall calendar. Each calendar page contained one or two weeks of

# THE HEALTH INSURANCE STUDY

## INSTRUCTIONS:

**SICK DAY:** Who had to cut down on usual daily activities (e.g., job, housework, school) for more than half a day because of health problems? **WRITE IN NAME OF PERSON.**

**MISS WORK:** Who took any time off from a job because of illness, going to the doctor or dentist, or emergency treatment or other health reason? **WRITE IN NAME OF PERSON.**

**PHONE ADVICE: Was a doctor, dentist, or other medical person called for advice?**  
**WRITE NAME OF SICK PERSON AND NAME OF MEDICAL PERSON CALLED**

January 1975											
S	M	T	W	T	F	S					
			1	2	3	4					
5	6	7	8	9	10	11					
12	13	14	15	16	17	18					
19	20	21	22	23	24	25					
26	27	28	29	30	31						

[illegible]




dates (depending on the Health Report reference period length, which is discussed later). There was a large blank space for each day of the one or two week period in which to note doctor appointments or anything else the family wished. Three "probe" questions appeared at the bottom of each day-space. One was a reminder to list telephone calls made to doctors, dentists, or other medical places for advice. The other two probes concerned nonutilization questions on the Health Report. It was suggested that the family hang the calendar in a conspicuous place and enter the requested information as soon as the events occurred.

Physician Visit Record (Exhibit 2). Part 1 of the Physicians, Doctors, Suppliers, and Outpatient MER was an 8-1/2" by 3-1/2" tear-off sheet on which the participant furnished basic identifying information about himself (or minor child), the date, and the name of the doctor (clinic, lab). The participant was instructed to complete the form, tear it off, answer some questions on the back, and retain the small sheet for use in filling out the next Health Report. If the participant complied, the visit eventually got reported to the study along with the correct date of occurrence and the correct provider name.

### Exhibit 2—Physician Visit Record


(FRONT)

6391		<b>FAMILY HEALTH PROTECTION PLAN</b>			
PHYSICIANS, DOCTORS, SUPPLIERS AND OUTPATIENT MEDICAL EXPENSE REPORT					
MAIL TO: FAMILY HEALTH PROTECTION PLAN, P.O. BOX 2393 PRINCETON, NEW JERSEY 08540					
PART I PARTICIPANT TO FILL IN ITEMS 1 THROUGH 13 PLEASE PRINT OR TYPE					
1. Last Name of Patient		First	MI	2. Sex M <input type="checkbox"/> F <input type="checkbox"/>	3. Age 4. Patient's Family No.
5. Patient's Address		City, State, Zip Code			6. Patient's Individual No.
7. What Was The Major Reason or Symptom For This Visit To The Doctor?			8. Was Illness or Injury Employment Related? YES <input type="checkbox"/> NO <input type="checkbox"/>	9. Was Illness or Injury Accident Related? YES <input type="checkbox"/> NO <input type="checkbox"/>	10. Date of Injury or Accident
11. Name of Doctor			12. Has the Patient Ever Visited This Doctor Before? YES <input type="checkbox"/> NO <input type="checkbox"/>		
13. I authorize any holder of medical or other information about the patient to release to the Family Health Protection Plan or its intermediaries any information needed for this or related medical report. I permit a copy of this authorization to be used in place of the original. In conformance with the Family Health Protection Plan contract all medical insurance benefits covering the Patient are hereby assigned to the Family Health Protection Plan.					
SIGN HERE 			Signature of Adult Participant or Guardian of Minor Participant		
			Date Signed		

Dental Visit Record (Exhibit 3). A small sheet similar to the one described above was attached to the Dental Service MER. It was filled out and retained by the participating family and the information eventually transferred to the Health Report.

**Exhibit 3—Dental Visit Record**

(FRONT)

0811		<b>FAMILY HEALTH PROTECTION PLAN</b>			
		<b>DENTAL SERVICE MEDICAL EXPENSE REPORT</b>			
MAIL TO: FAMILY HEALTH PROTECTION PLAN, P.O. BOX 2393 PRINCETON, NEW JERSEY 08540					
<b>PART I PARTICIPANT TO FILL IN ITEMS 1 THROUGH 9 PLEASE PRINT OR TYPE</b>					
1. Last Name of Patient		First	MI	2. Sex M <input type="checkbox"/> F <input type="checkbox"/>	3. Age
4. Patient's Family No.					
5. Patient's Address		City, State, Zip Code			6. Patient's Individual No.
7. What Was the Major Reason or Symptom For This Visit To The Dentist?		8. Name of Dentist			
9. I authorize any holder of medical or other information about the patient to release to the Family Health Protection Plan or its intermediaries any information needed for this or related medical report. I permit a copy of this authorization to be used in place of the original. In conformance with the Family Health Protection Plan contract all medical insurance benefits covering the Patient are hereby assigned to the Family Health Protection Plan.					
SIGN HERE 		Signature of Adult Participant or Guardian of Minor Participant			Date Signed

Pharmacy Purchase Record (Exhibit 4). On the back of the small sheet attached to the Pharmacy or Durable Equipment MER, the patient was asked to record the prescription number assigned to each prescribed item bought, the date of purchase, and the name of the store at which the purchase was made. This information, along with the patient's name, was transferred to the Health Report.

DATA FOLLOWUP

The two sources of utilization data, the MER and the Health Report, went through a checking (editing) process when they were received. If the document did not meet quality specifications, attempts were made to obtain the required information.

Exhibit 4—Pharmacy Purchase Record

Prescription No.	Date of Purchase	Place of Purchase

Health Report Edit

The Health Report received careful checking during the first two months of the experiment in order to identify and correct reporting problems. The edit concentrated on missing answers to questions that should have been answered. When item nonresponse was detected, a telephone call was made to the family, the responding error explained, and instruction given in how to provide the information in the future. Personal visits were made to families with chronic edit problems. Results from the pilot study of the control group (the first time the Health Report was used) suggested that this kind of retraining would be required for over half of the enrolled families during the initial period of the experiment. Actual experience indicated a much lower rate of problems during the main experiment; and by the end of the third month, item nonresponse errors occurred in less than 1 percent of the Health Reports.

Edit requirements were relaxed after the initial training period. The telephone followup method was abandoned in favor of a mailed follow-up procedure in which a copy of the questionnaire page with the response problem was mailed to the family with an explanation of the deficiency. The family was asked to make written corrections and mail the pages back to the HIS. Further followup was handled by telephone.

#### MER Edit

Each of the MERs received by the HIS was also edited for completeness and conformity both to research objectives and to industry standards for insurance claim paying. Special attention was given to the accuracy of patient identifying information, the detailed list of services and charges, diagnostic information, and the "links" between specific services and diagnoses. Problems of incompleteness or lack of clarity were normally called to the attention of the provider in written form with telephone followup as necessary.

Initially, each MER was coded into episodes (see above) and providers contacted for further information when it was not possible to complete the episode coding process satisfactorily. This kind of followup was abandoned when the decision was made to create computer based coding algorithms for episode coding, which would be carried out too far in the future to expect providers to resolve ambiguities accurately on the basis of recall.

The greatest amount of followup was required when outpatient providers submitted a duplicate bill in place of an MER. An additional attempt was then made to obtain research information (e.g., diagnoses and their relationship to goods and services dispensed).

#### Locating Missing MERs

After editing, an attempt was made to match each MER received with an entry on a Health Report. After 12 weeks, efforts were made to obtain MERs for each Health Report entry that did not have a matching MER. Because of the potentially large biases due to a 12-week recall period, no attempt was made to obtain Health Report information (e.g., delay to appointment, travel time) for visits reported only on MERs.

The HR-MER match rules required that the independent reports agree on:

- Name of patient
- Type and name of provider
- Date of service
- Name of item (for medicines and supplies) and, if applicable, prescription number.

Some discretion was exercised by operations personnel in locating missing MERs, but the general steps were as follows:

1. A telephone call<sup>8</sup> was made to the family to ask the whereabouts of the missing MER.
2. If the family said it did not give the MER to the provider, it was asked to do so. If it refused, the family was asked to send a copy of the provider's bill to the HIS. (It went through routine editing, which may have resulted in further followup contacts.)
3. If the MER had been delivered to the provider, the provider was contacted by telephone and asked to submit the document. If the provider was unwilling, a duplicate itemized bill was requested (which, if submitted, was edited and might result in further contact).
4. In the rare cases where neither an MER nor a duplicate bill could be obtained, the family was recontacted by telephone and asked to provide as much of the relevant information about procedures, diagnoses, and costs as could be remembered.

The MER location activities were stopped approximately six months after the reported date of service (or purchase). At that time whatever information was available about the utilization was entered into the data base. In some cases this included information that a Health Report item could not be matched to an MER. As a result, researchers have the option of treating this information as a "false positive" entry for purposes of data analysis.

#### THE MEDICAL EXPENSE REPORT

The principal set of data collection instruments for measuring uses of and expenditures for health care services were the Medical Expense Reports (MERs). Four different medical expense reports were

---

<sup>8</sup> Mailed forms were substituted for initial telephone contacts part way through the first year. The net effect of the substitution was to increase the clarity of communication and reduce the contact costs.

used during the first year: inpatient hospital and extended care services, physician or supplier services, dental services, and pharmacy or durable equipment expense reports. There were two versions of each type of report, one for participants in the experimental health insurance groups and one for control group people who continued to hold their previous health insurance policies.

#### General Format and Content

The individual forms were one or two page questionnaires to be filled out by the provider of health services. They corresponded as far as possible to standard claim forms in the insurance industry. The question items were a combination of forced choice and short answer questions to obtain information for episode coding, demand analysis, tracking of related utilization information, and claims payment administration. The forms were printed on chemically treated paper so that copies were created automatically.<sup>9</sup> The control group forms were similar to the experiment forms in content. They did not contain an item for the provider to indicate whether he wished to be paid directly by the HIS for covered services.

#### Participant Incentives to Furnish MER Data

The participating family was responsible for taking or mailing the MER to the provider (hospital, doctor, pharmacy, etc.), entering the identifying information on the top part of the form, and asking the provider to complete the remaining items. In return for completed MERs, the experimental insurance plan paid for the listed goods and services for eligible participants, subject to deductible, coinsurance, and coverage provisions. Families in the control group (who were not eligible for experimental insurance) received no direct benefit. Families with

---

<sup>9</sup> A second copy is provided in case the provider wishes to give it to the patient. The desirability of this type of communication of diagnostic and therapeutic information to the patient was left to the discretion of the provider. The availability of this kind of medical detail could affect the patient's planning for future medical expenses and could have an undetected effect on consumption.

deductible plans did not receive direct benefits until total family health care expenditures (for covered services) exceeded the deductible threshold for that accounting year. The latter group used the MERs to document current expenses that applied toward the deductibles. They were informed, at enrollment, that the study administrator reserved the right not to recognize an MER submitted more than 30 days after the goods or services were received.<sup>10</sup> The control group family met the requirement to submit MERs as part of their general obligations to the study for which other benefits (payments for periodic interviews and questionnaires) were received. Families on other plans received all or part reimbursement (50 percent or 75 percent) of charges listed on MERs for covered services. Those on nonzero coinsurance plans received total reimbursement after their out-of-pocket expenses exceeded a predetermined maximum dollar expenditure during the accounting year.

#### Provider Incentives to Furnish MER Data

The provider filled out the MER so that he could obtain payment for his services. However, reimbursement to the provider could come directly from the patient if the provider refused to accept the patient's assignment of insurance benefits over to him<sup>11</sup> or if some or all of the costs of the services rendered were not reimbursible by the HIS. Providers were not compensated further for filling out MERs and any charges for filing forms were disallowed.<sup>12</sup>

This means that there was really no provider financial incentive to complete forms for families below the deductible threshold or for control families. In the latter group, the MER sometimes represented

---

<sup>10</sup>In practice, this provision was not enforced. It will be necessary to enforce it at the end of the experiment.

<sup>11</sup>Physicians and pharmacies did not accept assignment of benefits for about two-thirds of the claims during the first two months of the pilot research. This rate stabilized at around 50 percent of outpatient claims by the end of the first year of the main experiment in Dayton.

<sup>12</sup>During the first year, some physicians were charging control group families up to \$5 extra for completing an MER. Some of the families paid the fee in order to remain eligible for other benefits of participation in the study and as their contribution as good citizens to an important policy research effort.

a second set of paperwork for the provider since he also had to complete the required claim forms for the control family's private insurance company. Special data collection provisions were available if the health care provider was unwilling to fill out an MER.

#### ALTERNATIVES TO MEDICAL EXPENSE REPORTS

Both the claim-paying and research components of the HIS accepted a copy of an itemized bill for health care. However, the bill usually does not contain diagnosis, episode coding, or MER tracking information and results in missing data in the dependent variables correlated with predictor variables (price parameters).<sup>13</sup> If neither the MER nor an itemized bill could be obtained, the patient (or parent proxy) was interviewed by telephone and asked to recall the details (procedures, diagnoses, dates, costs) of the utilization.

#### Inpatient Hospital and Extended Care Medical Expense Report (Exhibits 5a and b)<sup>14</sup>

This form obtained information to estimate admission, length of stay, and services demand equations. It included admission and discharge dates, number of days in each of eight types of accommodations, and a detailed listing of each hospital service received (laboratory, radiology, pharmacy or supplies, operating room, delivery room, private duty nurse, ambulance, etc.). Detailed codes for laboratory and radiology goods and services were provided.

For episode coding, the date of admission and diagnosis information was obtained. The patient was asked to supply his own diagnosis label and the date symptoms first appeared. Because of the potential complexity of the task and the possibility of substantial amounts of error, the hospital was not asked to match service line items to appropriate illness episode diagnoses.

For MER tracking, the form obtained names of the admitting and attending physicians (who generally submit separate expense reports for

---

<sup>13</sup> This creates heteroscedastic error in the dependent variable measurements. The effect on parameter estimates is not yet known.

<sup>14</sup> Available from the Health Insurance Study by requesting Documents HIEI 752 5M Rev. 11/73 (experimentals) and HIEI 752 7/74 (controls).



# Exhibit 5a—Inpatient Hospital and Extended Care Medical Expense Report

FAMILY HEALTH PROTECTION PLAN									
INPATIENT HOSPITAL AND EXTENDED CARE MEDICAL EXPENSE REPORT									
MAIL TO: FAMILY HEALTH PROTECTION PLAN, P.O. BOX 2393 PRINCETON, NEW JERSEY 08540									
PART 1 - PARTICIPANT TO FILL IN ITEMS 1 THROUGH 11									
Copy From Identification Card	1. Last Name of Patient		First	MI	2. Sex	3 Age	4. Patient's Family No.		
	5. Patient's Address		City, State, Zip Code		M <input type="checkbox"/> F <input type="checkbox"/>	6. Patient's Individual No.			
7. Briefly describe the illness or injury				8. Date symptoms first occurred		9. Was illness or injury accident related?		10. Was your illness or injury connected with your employment?	
						Yes <input type="checkbox"/> No <input type="checkbox"/>		Yes <input type="checkbox"/> No <input type="checkbox"/>	
11. I authorize any holder of medical or other information about the patient to release the Family Health Protection Plan or its intermediaries any information needed for this or related medical report. I permit a copy of this authorization to be used in the event of a lawsuit, and request payment of medical benefits to the party who accepts assignment before or the planholder, if not assigned. In conformance with the Family Health Protection Plan, the payment of medical benefits to the party who accepts assignment before or the planholder, if not assigned, in conformance with the Family Health Protection Plan, the payment of medical benefits to the party who accepts assignment before or the planholder, if not assigned, in conformance with the Family Health Protection Plan.									
SIGN HERE		Signature of Adult Participant or Guardian of Minor Participant						Date Signed	
PART 2 - HOSPITAL OR EXTENDED CARE FACILITY MUST COMPLETE ITEMS 12 THROUGH 33									
12. Provider Name and Address			13. Employer Number		17. Admitting Physician		19. Primary (Discharge) Diagnosis		
			14. Medical Record No.		18. Attending Physician(s) (if different from above)		ICDA Code		
			15. Date of This Admission				20. Secondary Diagnosis		
			16. Date of Discharge				ICDA Code		
							21. Other Diagnosis Contributing To Stay		
							ICDA Code		
PART 3 - ATTACH ITEMIZED BILL AND COMPLETE ITEMS 31 - 32							22. Other Diagnosis Not Contributing To Stay		
31. ACCOMMODATION      DAYS      RATE      TOTAL CHARGE							ICDA Code		
A 1-Bed							23. Period Covered by Billing		
B 2-3-4 Bed							From (Mo/Day/Yr)      Thru (Mo/Day/Yr)		
C 5 or More Beds							24. Billing Date (Mo/Day/Yr)		
D Intensive Care									
E Coronary Care									
F Nursery									
G Per Diem Charge									
H Per Diem Ancillary Charge									
32. DATE      SERVICE      RATE							25. If Physician does not bill separately, please indicate surgical or medical procedures		
A							CPT Code      Mo      Day      Yr		
B									
C									
D									
E									
F									
G									
H									
I									
J									
K									
L									
M									
N									
O									
P									
Q PHARMACY: Specify-See Last Page and Attach Bill							26. Patient Status		
R RADIOLOGY: Specify-See Last Page and Attach Bill							Discharged      Deceased      Still Patient		
S LABORATORY: Specify-See Last Page and Attach Bill							<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>		
T TOTALS							27. If Patient Discharged, Indicate Destination		
U LESS PREPAYMENT							<input type="checkbox"/> Home <input type="checkbox"/> Child Care Inst.		
V NET TOTALS							<input type="checkbox"/> Home Health Plan <input type="checkbox"/> Intermediate Care Facility		
							<input type="checkbox"/> Hospital <input type="checkbox"/> Skilled Nursing Facility		
							<input type="checkbox"/> Facility <input type="checkbox"/> Other		
							28. Discharged Destination Provider:		
							29. Assignment of Patient's Payment		
							<input type="checkbox"/> Accept Assignment		
							<input type="checkbox"/> Do Not Accept Assignment		
							30. It is hereby certified that the services and/or supplies listed above have been provided on the date(s) shown.		
							Signature of Provider's Representative		
							Date Signed		

Form HIEI - 752 5M Rev. 11/73

**LABORATORY, RADIOLOGY, PHARMACY SERVICES**

[illegible]

CODE	NAME	CODE	NAME	CODE	NAME	CODE	NAME
1	HCT	11	Bili; SGPT; ALK Phos.	21	Pap Smear	31	Plane Films
2	HGB	12	SGOT; LDH	22	Biopsy		Abdomen or Pelvis
3	U/A	13	Cholesterol/Triglyceride	23	Other Pathology	32	IVP/VCU
4	Clotting Studies	14	Other Chemistry	24	EKG	33	Isotope Scan
5	CBC With Diff.	15	Urine Culture	25	EEG	34	Other Diagnostic
6	Other HEME	16	Throat Culture	26	EMG		Radiology
7	Electrolyte	17	Other Bact.	27	Other Electrical	35	Any Therapeutic
8	BUN/Creat	18	STS	28	Other Laboratory		Radiology
9	Ca/Po4	19	Sickle Prep.	29	Chest Films		
10	T4/PBI	20	Other Serology	30	Skull Films		

## PHARMACY SERVICES

their services in connection with illness episodes treated in the hospital) and the name, address, and type of health care facility, if any, to which the patient had been discharged.

Physicians, Doctors, Suppliers, and Outpatient Medical Expense Report (Exhibit 6)<sup>15</sup>

This form was designed to obtain details of services rendered by physicians and other licensed practitioners on either an inpatient or an outpatient basis. On the basis of pilot study experience, dental services were excluded from this form and a separate dental report (described below) was created. Three types of information, in addition to disaggregated charges, were sought for the analysis of quantity demanded: (a) consultation (or visit) services including a special code to reflect the "intensity" of these services on an ordinal scale,<sup>16</sup> (b) medical and surgical procedures, and (c) specific lab tests, drugs injected, and supplies furnished. Although the quality of care analysis would benefit greatly from knowing the results of lab tests and other diagnostic and therapeutic procedures, it was judged infeasible to request this kind of information on the expense form.

The Outpatient MER was formatted to classify services into illness episodes. Providers were asked to list all diagnoses and problems treated or monitored during the visit, to indicate the date the symptoms first occurred, the first consultation date for the problem, and the treatment history for the problem.<sup>17</sup> The diagnosis or problem information was entered on separate lines, each with a letter code designation (A, B, C, D). The provider was asked to link each consultation, service, and supplies item to one or more of the diagnoses by recording the letter code next to each listed service item. The provider was also asked to

---

<sup>15</sup> Available from the Health Insurance Study by requesting documents HIEI 751 7/74 (experimentals) and HIEI 761 7/74 (controls).

<sup>16</sup> Minimal service, brief examination, limited examination, intermediate history and physical examination, extended examination, comprehensive history and physical examination.

<sup>17</sup> The following nominal scale treatment history codes were made available to ensure comprehensive and standardized responses across providers: initial acute, initial chronic, repeat acute, chronic routine, chronic flareup, well care.


# Exhibit 6—Physicians, Doctors, Suppliers and Outpatient Medical Expense Report

6391



## **FAMILY HEALTH PROTECTION PLAN** PHYSICIANS, DOCTORS, SUPPLIERS AND OUTPATIENT MEDICAL EXPENSE REPORT MAIL TO: FAMILY HEALTH PROTECTION PLAN, P.O. BOX 2393 PRINCETON, NEW JERSEY 08540

PART 1

PART 1 PARTICIPANT TO FILL IN ITEMS 1 THROUGH 13 PLEASE PRINT OR TYPE			
1. Last Name of Patient	First	MI	2. Sex M <input type="checkbox"/> F <input type="checkbox"/>
3. Age		4. Patient's Family No.	
5. Patient's Address		City, State, Zip Code	
6. Patient's Individual No.			
7. What Was The Major Reason or Symptom For This Visit To The Doctor?		8. Was Illness or Injury Employment Related? YES <input type="checkbox"/> NO <input type="checkbox"/>	9. Was Illness or Injury Accident Related? YES <input type="checkbox"/> NO <input type="checkbox"/>
10. Date of Injury or Accident			
11. Name of Doctor		12. Has the Patient Ever Visited This Doctor Before? YES <input type="checkbox"/> NO <input type="checkbox"/>	
13. I authorize any holder of medical or other information about the patient to release to the Family Health Protection Plan or its intermediaries any information needed for this or related medical report. I permit a copy of this authorization to be used in place of the original. In conformance with the Family Health Protection Plan contract all medical insurance benefits covering the Patient are hereby assigned to the Family Health Protection Plan.			

SIGN HERE 	Signature of Adult Participant or Guardian of Minor Participant	Date Signed
---	---	-------------

PART 2

PART 2 DOCTOR OR SUPPLIER TO FILL IN ITEMS 14 THROUGH 27 PLEASE PRINT OR TYPE							
14. Name of Referring Doctor. If None, Write None				15. Name of Providers to Whom You Referred Patient for Consultation, Lab. Tests, or Other Services. If None, Write None.			
16. Describe the Primary Problem or Diagnosis Which Brought the Patient To Your Office and Any Co-existent Problem for Which You Supplied Treatment. Please List Primary Problem or Diagnosis on Line A.				17. Date Symptoms First Occurred, For Pregnancy Estimate Delivery	18. Date You Were First Consulted for This Problem	19. Treatment History for This Problem * (1)	
A.						1 ( ) 2 ( ) 3 ( ) 4 ( ) 5 ( ) 6 ( )	
B.						1 ( ) 2 ( ) 3 ( ) 4 ( ) 5 ( ) 6 ( )	
C.						1 ( ) 2 ( ) 3 ( ) 4 ( ) 5 ( ) 6 ( )	
D.						1 ( ) 2 ( ) 3 ( ) 4 ( ) 5 ( ) 6 ( )	
20. A. Date Of Service	B. Place of Service (2)	C. Describe Each Medical or Surgical Procedure and Other Service or Supplies Furnished For Each Date, Including Specific Lab Tests and the Specific Name of Any Drug Injected. If Office Visit, Write Visit and Indicate Type According to Codes J, K, Below	D. Reuse Treatment to Problem by Ref to 15 A, B, C or D above	E. Charge	21. Were Any Drugs Prescribed? A. If Yes, Specify Prescribed Drug (s)		B. Refer to Problem by Referring to 15 A, B, C or D above
1					YES <input type="checkbox"/> NO <input type="checkbox"/>		
2							
3							
4							
5							
6							
22. Name and Address of Doctor or Supplier			23. Social Security or Employee Number	24. TOTAL CHARGE			
			Telephone Number	25. LESS PREPAYMENT			
				26. NET TOTAL			
27. I hereby certify that the services and/or supplies listed above have been provided on the date(s) shown.							Date Signed
PROVIDER'S SIGNATURE 							
28. I hereby authorize payment directly to the above-named doctor to the benefits otherwise payable to me, but not to exceed the charges shown. I understand that I am financially responsible for any charges not covered by the Family Health Protection Plan							Date Signed
PARTICIPANT'S SIGNATURE 							
<p>1. Treatment History: 1 = Initial Acute; 2 = Initial Chronic; 3 = Repeat Acute; 4 = Chronic Routine; 5 = Chronic Flareup; 6 = Well Care</p> <p>2. Place of Service Codes: O = Doctor's Office; IL = Independent Laboratory; H = Patient's Home; IH = Inpatient Hospital; NH = Nursing Home or SNF; OH = Outpatient Hospital; CL = Clinic; OL = Other Location</p> <p>3. 1 = Minimal Service; 2 = Brief Examination; 3 = Limited Examination; 4 = Intermediate History and Physical Examination; 5 = Extended Examination; 6 = Comprehensive History and Physical Examination</p>							

list the names of prescribed medications and to link these to diagnoses by the letter code system.

Location of other expense reports related to health episodes listed on the Physician MER was facilitated by answers to questions about name of referring doctor, names of providers to whom patient was referred (for consultation, examination laboratory tests, and other services), and the name of prescribed medication.

Dental Service Medical Expense Report (Exhibit 7)<sup>18</sup>

A detailed listing of dental procedures and costs provided the data for estimating demand. There is specific policy interest in the effects of periodic examination and cleaning on later need for dental services. To minimize measurement error at this level of disaggregation, the dental form listed these two items and requested providers to furnish information about each separately. Quality of care analysis requirements led to identifying the specific teeth involved in each service (a standard dental chart with codes for each tooth is provided on the form) and an additional question was included about the type of fluoride used in treatments.

Episode coding information was obtained with a question about primary symptom or diagnosis (not linked to specific services). Neither treatment history nor onset of symptoms was asked. Episode MER tracking information was confined to an item about name of prescribed drug.<sup>19</sup> No information about referring or referred dentists was obtained since referral is infrequent.

Pharmacy or Durable Equipment Medical Expense Report (Exhibit 8)<sup>20</sup>

The Pharmacy or Durable Equipment MER included a list of each

---

<sup>18</sup> Available from the Health Insurance Study by requesting Documents HIEI 759 7/74 (experimentals) and HIEI 764 7/74 (controls).

<sup>19</sup> In the second year of the experiment, dental insurance coverage was expanded. At that time, providers were asked to submit a treatment plan for multi-visit services. The treatment plan was used to classify visits and services into treatment episodes for analysis.

<sup>20</sup> Available from the Health Insurance Study by requesting Documents HIEI 753 Rev. 7/74 (experimentals) and HIEI 763 7/74 (controls).

# FAMILY HEALTH PROTECTION PLAN

## DENTAL SERVICE MEDICAL EXPENSE REPORT

MAIL TO: FAMILY HEALTH PROTECTION PLAN, P.O. BOX 2393 PRINCETON, NEW JERSEY 08540

FORM HIEI #759 7/74

# **Exhibit 8—Pharmacy or Durable Equipment Medical Expense Report**

4642

## **FAMILY HEALTH PROTECTION PLAN PHARMACY OR DURABLE EQUIPMENT MEDICAL EXPENSE REPORT**

MAIL TO: FAMILY HEALTH PROTECTION PLAN, P.O. BOX 2393 PRINCETON, NEW JERSEY 08540

PART 1 — PARTICIPANT TO FILL IN ITEMS 1 THROUGH 7 PLEASE PRINT OR TYPE							
1. Last Name of Patient		First		M.I.	2. Sex M <input type="checkbox"/> F <input type="checkbox"/>	3. Age	4. Patient's Family No.
5. Patient's Address		City, State, Zip Code				6. Patient's Individual No.	
7. I authorize any holder of medical or other information about the patient to release to the Family Health Protection Plan or its intermediaries any information needed for this or related medical report. I permit a copy of this authorization to be used in place of the original. In conformance with the Family Health Protection Plan contract all medical insurance benefits covering the Patient are hereby assigned to the Family Health Protection Plan.							
SIGN HERE  Signature of Adult Participant or Guardian of Minor Participant						Date Signed	
8. PART 2—DRUGGIST TO COMPLETE PART 2-FILL IN ITEMS 8 THROUGH 15. PLEASE PRINT OR TYPE							
Date Item Dispensed	Prescription No.	Refill No. _____ of _____	Prescribing Physician		Prescription Phoned Yes <input type="checkbox"/> No <input type="checkbox"/>	Manufacturer	
1 Date of Prescription	Physician's Dosage Instructions	Brand Or <input type="checkbox"/> Generic <input type="checkbox"/>	Name of Drug Dispensed (See Item 9 below)		Dosage Form	Strength	Metric Quantity Charge
Date Item Dispensed	Prescription No.	Refill No. _____ of _____	Prescribing Physician		Prescription Phoned Yes <input type="checkbox"/> No <input type="checkbox"/>	Manufacturer	
2 Date of Prescription	Physician's Dosage Instructions	Brand Or <input type="checkbox"/> Generic <input type="checkbox"/>	Name of Drug Dispensed (See Item 9 below)		Dosage Form	Strength	Metric Quantity Charge
Date Item Dispensed	Prescription No.	Refill No. _____ of _____	Prescribing Physician		Prescription Phoned Yes <input type="checkbox"/> No <input type="checkbox"/>	Manufacturer	
3 Date of Prescription	Physician's Dosage Instructions	Brand Or <input type="checkbox"/> Generic <input type="checkbox"/>	Name of Drug Dispensed (See Item 9 below)		Dosage Form	Strength	Metric Quantity Charge
Date Item Dispensed	Prescription No.	Refill No. _____ of _____	Prescribing Physician		Prescription Phoned Yes <input type="checkbox"/> No <input type="checkbox"/>	Manufacturer	
4 Date of Prescription	Physician's Dosage Instructions	Brand Or <input type="checkbox"/> Generic <input type="checkbox"/>	Name of Drug Dispensed (See Item 9 below)		Dosage Form	Strength	Metric Quantity Charge
9. For Compounded Prescriptions, List the Ingredients and Total Cost of the Item						10. TOTAL CHARGE	
						11. LESS PREPAYMENT	
						12. NET TOTAL	
13. Name and Address of Provider						14. Employer I.D. Number	
						Telephone No.	
15. I hereby certify that the services and/or supplies listed above have been provided on the date(s) shown.  PROVIDER'S SIGNATURE						Date Signed:	
16. I hereby authorize payment directly to the above-named Druggist of the benefits otherwise payable to me, but not to exceed the charges shown. I understand that I am financially responsible for any charges not covered by the Family Health Protection Plan.  PARTICIPANT'S SIGNATURE						Date Signed:	

prescribed item sold and its cost for the demand analysis. To estimate that part of medication cost not related to physical quantity units (see discussion of amenities, Sec. II) information was also obtained on whether the good was generic or a brand item and the form of dosage. Information about physician dosage instructions, name of medication or list of compounded ingredients, strength, and metric quantity was obtained for quality of care assessment.

The main task for episode analysis was to link each pharmacy item to a particular problem or diagnosis on a physician MER. Therefore, the item name, date the prescription was filled out, and the name of the prescribing physician was included. The form also asked whether the prescription was phoned to the pharmacist. If yes, the episode coders were altered to the possibility that a linkable Physician MER might not exist. The date on which the item was purchased and the refill history were included for the expenditure targeting part of the episode analysis.

#### MEASURES OF UTILIZATION FOR WHICH MERs ARE NOT REQUIRED

Participating families were not required to submit MERs to document medical consultation by telephone or purchases of nonprescription medicines and supplies (e.g., aspirin or bandages).<sup>21</sup> Telephone consultations, excluding calls only to make an appointment, were reported on the Health Report. Nonprescribed items (except as noted) were not measured during year one by any of the data collection procedures. They were added to the Health Report in year two.

#### HEALTH REPORT REFERENCE PERIOD EXPERIMENT

One question left unresolved by the pilot experiment is the length of the reporting period to be used with the Health Report. To be maximally effective, the Health Report should detect as many "true" uses of health care services as actually take place. The length of the time period for which utilization is reported (called the reference period) should be designed to maximize the probability that every true utilization is reported.

---

<sup>21</sup>If prior authorization were obtained from the HIS, nonprescribed items judged medically necessary could be covered by the experimental insurance. MERs were required for these purchases.



The pilot research results suggest that the reference period should not be as long as a month. The monthly supplementary reporting treatment obtained many fewer reports of utilization than the weekly treatment for the deductible insurance plan. The measurement methodology literature contains a good deal of evidence that reporting error increases with elapsed time between the occurrence of an event and when it is reported. Long reference periods increase the probability that the elapsed times are long. Cannell and Fowler (1963) report that about 15 percent of known visits occurring within a week of the interview are not reported, and 30 percent occurring between one and two weeks before the interview were not mentioned. Taken at face value these studies strongly suggest that a one-week reference period, or even a shorter one, is needed to eliminate underreporting bias in questionnaire reports of health care consumption. Two-week or one-month reference periods permit levels of underreporting that are clearly unacceptable for measurement purposes.

There are several reasons not to accept these research results at face value. For example, the monthly measurement treatment in the pilot study included none of the record and reporting aids discussed above (in connection with the design of the Health Report). Had the monthly form included these assists, the difference between it and the weekly questionnaire might have been much smaller. The Cannell and Fowler research results might not apply to the HIS because they deal with reporting in a one-time, personal interview. The Health Insurance Study involves continuous contact with respondents, and reporting is done on a self-administered questionnaire. Research by Sudman, Wilson, and Ferber (1974), however, suggests self-administration and panel participation may not alter the underreporting tendency. Their panel study estimate of underreporting of health care consumption using a diary (19 percent) is not much different from the overall underreporting estimate of Cannell and Fowler (23 percent).

The most important reason for questioning the applicability of the Cannell and Fowler results to the HIS is that they used a one-directional, retrospective record check design to infer response bias in respondent data. Marquis, Marquis, and Newhouse (1976) have shown that this kind

of design can cause random response and matching errors to appear as an underreporting bias. Several other record check studies reporting physician and dentist visits containing minimal design biases<sup>22</sup> suggest that survey data (or the record data or both) contain mostly random errors rather than a net underreporting bias. It is quite possible that what Cannell and Fowler interpreted as increased underreporting bias for two-week recall versus one-week recall was really an increase in random reporting error.

This could mean one of at least two things, if true:

1. Respondents individually report the correct number of utilizations but, as the reference period becomes longer, they are more likely to misreport details of utilization (such as the doctor's name, the patient's name, or when the visit took place). Getting these details confused when reporting them can account for the high rates of over- and underreporting obtained by the full-design record check studies (but the absence of a net reporting bias) because these details are used to match survey and record information about utilization. Incorrect details result in mismatches even if the basic fact of utilization frequency is reported correctly. If this is the case, the HIS might use short reference periods merely to increase the probability of obtaining the correct details of each utilization. Correct details make it easier for the study to know if an MER has been received for a utilization on the Health Report and facilitate obtaining an MER (for the correct utilization) if one has not been submitted. However, using longer reference periods may only mean that extra communication with the respondent about the details of a utilization is needed before reporting discrepancies can be resolved.

2. Respondents may misreport the true number of utilizations more often when longer reference periods are used. The error is random across respondents because for every overreported visit by one respondent, another respondent underreports a visit. This pattern of errors

---

<sup>22</sup>Cartwright (1963), Sudman, Wilson, and Ferber (1974), Marquis, Marquis, and Newhouse (1976), and the HIP clinic visit component of Lowenstein's research (1969). The full design study reported by Feather (1972) suggests there may be a net overreporting bias.

has more serious implications for HIS reference period design. If some utilizations really are not reported at all on Health Reports, the efficacy of the questionnaire in identifying all utilizations for MER tracking purposes is diminished. In this case it is much more important to use short reference periods for the supplementary questionnaires.

Unfortunately, the published studies do not provide much insight into the type of random response error to be expected, and it is always possible that the introduction of calendars and visits or purchase records will eliminate recall error, mooted the decision about Health Report reference period. A "safe" design choice would be to use short reference periods. However, the costs of weekly questionnaires are substantially above the costs of biweekly questionnaires (adding as much as \$1-2 million to the measurement costs over the duration of the experiment).

As a result, the first year supplementary measurement design includes an experiment using one-week and two-week Health Report reference periods. Half of the enrolled families are allocated to the weekly treatment and half to the biweekly period. Effects of the different reference periods on the amount of utilization detected within each insurance treatment will be examined. If effects are found, further experiments are planned to isolate their nature (e.g., short reference periods aid recall, frequent supplementary reporting stimulates demand for health services, or both). A final step will be to estimate the magnitude of any measurement effects and use these estimates to adjust the empirically derived demand parameter estimates.

## V. PRELIMINARY EVALUATIONS

During the first year of the HIS experiment, a broad scope of kinds of utilization was measured at a fairly fine level of disaggregation by two measurement instruments: the Medical Expense Report, which obtained technical utilization data from medical care providers, and the Health Report (plus intermediate records), which provided an independent stream of data from families about use of gross categories of health care services. The two streams of data were used to detect as much health care consumption as possible and to obtain disaggregated information about procedures and expenditures for each utilization.

From administrative reports and other sources, two tentative conclusions have emerged from the first year measurement of utilization.

1. The control group utilizations during the experiment were severely underreported. Using baseline data covering the year before enrollment, the control group appeared to use health services at approximately the national per person rates published by the National Center for Health Statistics from their Health Interview Survey. During the first year of the experiment, the control group utilization rates ascertained from Health Reports dropped by about one-third. Utilization rates calculated from MER-only data were 50 percent or less of the expected rates.

These are very preliminary analyses of the aggregated operations monitoring data. The conclusions may change considerably upon analyses of the individual and family data in the data base. At present, however, a researcher considering the collection of health services utilization and expenditure data from a cross-section population sample using panel techniques would be advised to look into procedures used by the Current Medicare Survey of the Social Security Administration or the Medical Expenditure Study of the National Center for Health Services Research and the National Center for Health Statistics.

2. A back-up system for detecting missing MERs appears needed but there is doubt about which kind of system is best. In year one of the HIS, analysis of operations monitoring data suggests that weekly probes

obtain slightly more reports of utilization on Health Reports than the biweekly probes. The effect is confined to the less generous insurance plans. The weekly probes detect more utilization unreported on MERs than the biweekly measurement treatment. They may also be (or only be) stimulating utilization artificially.

Again, when it is possible to analyze the full range of data, including MERs, these conclusions may change considerably. In addition, the Health Insurance Study has made plans to test both the stimulation effects of probes and the remaining extent of utilization underreporting on health reports using records of utilization in a prepaid group practice plan and records of providers in other sites. An outline of these designs is reported in Newhouse et al. (1976). An additional feature, investigating the effectiveness of the procedures that attempt to obtain MERs for all utilization mentioned in the Health Report, has also been incorporated recently.

The potential problem with the year one measurement system is that measurement biases may be present that are correlated with the price treatments: Patients in the coinsurance plans, deductible plans, and control groups (and their providers) have less incentive to report utilizations than other participants because reporting on an MER involves duplicate billing and need not lead to any expenditure reimbursement by the HIS. The problem is potentially greater in the control group because participants must sometimes ask the doctor, pharmacist, dentist, etc. to fill out two detailed claims forms for a single visit (one for the study and one for the family's health insurance company).<sup>1</sup>

The Health Report system may be one possible way to compensate for measurement biases in the MER-only utilization detection system. It provides a simple way to report all uses of services for participants in all plans. However, its effectiveness may diminish as families realize that the study requires an MER for each utilization reported on the HR. Families may learn not to report utilizations on the Health Report if they do not want to take the trouble of initiating an MER.

---

<sup>1</sup>The utilization measurement problems in the control group contributed to the eventual decision to drop this feature of the experimental design.

Such underreporting has no direct consequences to control group families since they do not forgo any payments when utilization is underreported. Families in the deductible plans lose by underreporting only if their true out-of-pocket health expenditures exceed the deductible limit. Pilot results and preliminary results from the first year weekly or biweekly methods experiment suggest that frequent Health Report probes may overcome at least some of the correlated measurement biases. However, the Health Report system does involve the risk of both artificially stimulating utilization and measurement biases, so the net result could be unsatisfactory in terms of generalizing findings to a national health insurance context. Alternative procedures to remove bias include:

- Paying providers a fee for filling out MERs. Preliminary estimates indicate that this would add over \$1 million to the lifetime costs of the research and not necessarily remove all of the measurement biases. Respondents may still neglect to deliver MERs to providers, clinics delivering free care may find the incentive insufficient, and the fee may not be large enough to persuade some providers to cooperate.
- Reimbursing patients for MERs. It is possible to pay families some flat fee or proportion of each MER-reported expense to gain their cooperation in delivering MERs and persuading providers to complete them. The danger in this approach is that the payment is contingent on the use of services and, indeed, alters the out-of-pocket price of each service to the family. Out-of-pocket price is the experimental variable and such a procedure risks confounding the experimental demand observations.<sup>2</sup>
- An independent record search. In theory it is possible to search all records of doctors, dentists, pharmacies, etc. in

---

<sup>2</sup> A variant of this reimbursement procedure was instituted in year two for the deductible insurance treatments. The HIS pays 5 percent of the bill for covered services before the deductible is met. It is felt that the effects of the extra 5 percent reimbursement can be modeled satisfactorily.

a limited geographical area for details of utilization for families enrolled in the study. Since costs of a census of all records would be prohibitive, a sampling procedure would be used. The measurement literature suggests that provider records are likely to contain error although possibly not bias. The sampling and measurement errors can be expected to combine and result in somewhat imprecise measures of the utilization dependent variables and of the consequent demand estimates. There are ways of increasing the efficiency of record check designs to improve precision at the expense of an unmeasurable (but possibly small) increase in bias. Nevertheless, precision of data derived from records probably would not reach that obtained from existing MERs, which are often based on fairly immediate recall by providers, and the sampling errors would remain.<sup>3</sup>

- Rely solely on Health Report data for analysis, using MERs only for administration of the insurance reimbursement obligations of the study. This option can always be exercised but at a major sacrifice in the ability to generate useful knowledge about consumption of health care services.

At present, cost-effective measurement procedures of known high reliability and validity do not exist for the measurement of health care consumption, so the HIS must devote extra resources both to its measurement and to evaluation of its measurement. New and different procedures are being tried and evaluated in other research efforts also. In five or ten years we can expect that this collection of knowledge and experience will reduce many of the measurement design uncertainties that contemporary researchers face. For the present the Health Insurance Study plans to continue to evaluate its measurement for error and bias using the kinds of procedures mentioned above. At the end of the data collection period we expect to have sufficiently precise estimates

---

<sup>3</sup>The HIS is considering the possibility of conducting record check studies in its smaller sites, where a census of certain kinds of providers may be possible.

of the measurement effects (if any) to adjust the price parameter estimates for them. Precision problems created by the need to adjust parameter estimates for measurement bias may be lessened by using the MER and Health Report streams of utilization data with the combinatorial principles of reaggregation mentioned in Sec. II.



REFERENCES

- Andersen, Ronald and Odin Anderson, *A Decade of Health Services*, University of Chicago Press, 1967.
- Andersen, Ronald et al., *Total Survey Error: Bias and Random Error in Health Survey Estimates*, Final Report submitted to the National Health Center for Health Services Research under contract PHS HRA 106-74-24, 1977.
- Balamuth, Eve, Sam Shapiro, and Paul M. Densen, "Health Interview Responses Compared with Medical Records," *Vital and Health Statistics*, June 1961, Series D-No. 5, U.S. Public Health Service Publication No. 584-D5.
- Brehm, J. W. and A. R. Cohen, *Explorations in Cognitive Dissonance*, Wiley, New York, 1962.
- Cannell, Charles F., Gordon Fisher, and Thomas Bakker, "Reporting of Hospitalization in the Health Interview Survey," *Vital and Health Statistics*, July 1965, Series 2, No. 6.
- Cannell, Charles F. and Floyd J. Fowler, "A Study of the Reporting of Visits to Doctors in the National Health Survey," Institute for Social Research, University of Michigan, October 1963.
- Cartwright, Ann, "Memory Errors in a Morbidity Survey," *Milbank Memorial Fund Quarterly*, Vol. 41, 1963, pp. 5-24.
- Clasquin, Lorraine A., *Mental Health, Dental Services, and Other Coverage in the Health Insurance Study*, The Rand Corporation, R-1216-OEO, November 1973.
- Cronbach, Lee J., "Coefficient Alpha and the Internal Structure of Tests," *Psychometrika*, Vol. 16, 1951, pp. 297-334.
- Deutscher, Irwin, "Words and Deeds: Social Science and Social Policy," *Social Problems*, Vol. 13, 1966, pp. 235-254.
- , *What We Say/What We Do: Sentiments and Acts*, Scott, Foresman and Co., Glenview, Illinois, 1973.
- Feather, Joan, *A Response/Record Discrepancy Study*, University of Saskatchewan, Saskatoon, November 1972.
- Festinger, Leon, "Behavioral Support for Opinion Change," *Public Opinion Quarterly*, Vol. 28, 1964, pp. 404-417.
- , *A Theory of Cognitive Dissonance*, Row, Peterson, New York, 1957.

- Fishbein, Martin, "Attitude and the Prediction of Behavior," in M. Fishbein (ed.), *Readings in Attitude Theory and Measurement*, Wiley, New York, 1967.
- Jöreskog, Karl G. and Arthur S. Goldberger, "Estimation of a Model with Multiple Indicators and Multiple Causes of a Single Latent Variable," *Journal of the American Statistical Association*, Vol. 70, 1975, pp. 631-639.
- Keeler, Emmett B., Joseph P. Newhouse, and Charles E. Phelps, *Deductibles and the Demand for Medical Services: The Theory of the Consumer Facing a Variable Price Schedule Under Uncertainty*, The Rand Corporation, R-1514-OEO/NC, December 1974.
- LaPiere, Richard T., "Attitudes vs. Actions," *Social Forces*, Vol. 13, 1934, pp. 230-237.
- McNeil, John C., *Census Bureau Programs to Measure Consumer Purchase Expectations: 1959-1973*, Technical Paper 36, Bureau of the Census, Washington, D.C., December 1974.
- Marquis, Kent H., James Marshall, and Stuart Oskamp, "Testimony Validity as a Function of Question Form, Atmosphere and Item Difficulty," *Journal of Applied Social Psychology*, Vol. 2, 1972, pp. 167-186.
- Marquis, Kent H., Susan Marquis, and Joseph P. Newhouse, "The Measurement of Expenditures for Outpatient Physician and Dental Services: Methodological Findings from the Health Insurance Study," *Medical Care*, Vol. 14, 1976, pp. 913-931.
- Morris, Carl, "Initial Precision of Estimates for the Health Insurance Study," unpublished Rand research, 1973.
- Newhouse, Joseph P., *The Health Insurance Study--A Summary*, The Rand Corporation, R-961-1-OEO, March 1974a.
- , "A Design for a Health Insurance Experiment," *Inquiry*, Vol. 11, 1974, pp. 5-27.
- Newhouse, Joseph P. and Charles E. Phelps, *Price and Income Elasticities for Medical Care Services*, The Rand Corporation, R-1197-NC-OEO, June 1974.
- , *On Having Your Cake and Eating It Too: Econometric Problems in Estimating the Demand for Health Services*, The Rand Corporation, R-1149-NC, April 1974.
- Newhouse, Joseph P. et al., "Measurement Issues in the Second Generation of Social Experiments: The Health Insurance Study," *Proceedings*, Social Statistics Section, American Statistical Association, 1976.
- Orr, Larry L., "Experimentation and Health Financing Policy," *Inquiry*, Vol. 11, 1974, pp. 28-39.

- Parry, H. J., M. B. Balter, and I. H. Cisin, "Primary Levels of Under-reporting Psychotropic Drug Use," *Public Opinion Quarterly*, Vol. 23, 1971, pp. 582-592.
- Sudman, Seymour, Wallace Wilson, and Robert Ferber, The Cost-Effectiveness of Using the Diary as an Instrument for Collecting Health Data in Household Surveys," Survey Research Laboratory, University of Illinois, October 1974.
- Tenney, James B., Kerr L. White, and John W. Williamson, "National Ambulatory Medical Care Survey: Background and Methodology," *Vital and Health Statistics*, Series 2, No. 61, DHEW Publication No. (HRA) 74-1335, April 1974.
- Webb, Eugene et al., *Unobtrusive Measures: Non-Reactive Research in the Social Sciences*, Rand-McNally, Chicago, 1966.
- Wicker, Allan W., "Attitudes vs. Actions: The Relationship of Verbal and Overt Behavioral Responses to Attitude Objects," *Journal of Social Issues*, Vol. 25, 1969, pp. 41-78.
- , "An Examination of the 'Other Variables' Explanation of Attitude--Behavior Inconsistency," *Journal of Personality and Social Psychology*, Vol. 19, 1971, pp. 18-30.
- Wilder, Charles S., "Acute Conditions: Incidence and Associated Disability: United States, July 1972-June 1973," *Vital and Health Statistics, Data from the National Health Survey*, Series 10, No. 98, DHEW Publication No. (HRA) 75-1525, January 1975.