# THE DESIGN AND METHODOLOGY FOR THE NATIONAL MEDICAL CARE EXPENDITURE SURVEY

Julia D. Oliver,[1] National Center for Health Statistics

## 1. INTRODUCTION

### 1.1 Background

The rapid escalation of medical care costs and their concomitant economic and social implications has received more national attention in recent years than any other health care issue. Since World War II the cost of medical care has risen dramatically: total personal health care expenditures have increased more than sevenfold; third-party payments have increased fiftenfold,[2] while the cost of living has more than doubled.[3] In just the four-year period of 1965-1969, the daily cost of a hospital room increased by 69.7 percent.[4] The effect of these expenditures on the health service delivery system and on the 75 million households has far reaching implications for our economy and all our citizenry.

At the present time, speculation about enactment of national health insurance legislation and its social and economic consequences is a major factor in generating the increased attention to the need for reliable information and research on the use of health care services, the expenditures for these, and the payment sources employed. Furthermore, if national health insurance legislation is passed, relevant and valid data are needed to evaluate the effect of the legislation.

The Health Interview Survey (HIS), conducted by the National Center for Health Statistics (NCHS) since 1957, has been a major vehicle to obtain operational, policy related, and evaluative data about the health system and its substantive aspects. However, even before the HIS, considerable interest had been directed to the problem of measuring and assessing medical and health expenditures. Studies were instituted as early as 1933 by I. S. Falk.[5] In 1953 the National Opinion Research Center (NORC) in collaboration with the Health Information Foundation (HIF) began a series of studies that were continued in 1958, 1963, and 1970. Subsequent studies such as the Federal Employees Health Benefits Program Utilization Study in 1972-1973 and the continuing studies by the HIS in 1971, 1975, and 1976 have monitored and documented individual and family health services utilization and expenditures.

However, previous surveys on health and medical expenditures have had trouble achieving desired levels of completeness and accuracy of reporting. Problems with recall periods and underreporting or incorrect reporting from the households have been major obstacles to obtaining valid health cost data in household interview surveys. The household interview process, by itself, has been thought to be an inadequate mechanism to obtain accurate data on the total costs of medical care. Collecting supplementary data from medical care providers, third-party payers, and employers can ameliorate the problems resulting from underreporting and incorrect reporting. Even where households report information on medically attended illness, however, most households cannot report diagnosis accurately.

In 1975, in the interest of improving medical expense data, NCHS conducted a pilot study to determine the most feasible survey design for a national study of medical care utilization, expenditures, and health insurance coverage.[6] During the same period, research on medical care utilization and expenditures was being conducted by the National Center for Health Services Research (NCHSR). In its program of intramural research, special emphasis was being given to the impact of possible changes in Federal income tax treatment of medical care expenses; the costs and benefits of various national health insurance proposals; the impact of Medicare and Medicaid programs on various subgroups of the population such as the poor and rural citizens; the impact of existing Federal health programs on access to care; the costs of illness for American families by diagnostic category and the impact of such costs on these families.

Given the shared interest in health expenditure data, NCHSR and NCHS decided to jointly sponsor the National Medical Care Expenditure Survey (NMCES). The principal funds for this research are provided through the Division of Intramural Research, NCHSR.

In June of 1976 a contract was awarded to the Research Triangle Institute (RTI), Research Triangle Park, North Carolina.[7] Dr. Daniel G. Horvitz, Vice President of RTI's Statistical Sciences Group, is project director. RTI has awarded subcontracts for components of the data collection to Abt Associates, Inc. (AAI), Cambridge, Massachusetts, and to NORC of the University of Chicago. RTI as the prime contractor has overall responsibility for the management of the contract. In conducting the survey, RTI and NORC shared responsibility for collecting the data. Each organization used their own probability samples and supervised their own field staff.

## 2. OVERALL DESIGN OF THE NMCES

NMCES is, in reality, three surveys: a survey of households, a survey of those medical providers and hospitals who provided services to the sampled population, and a survey of health insurers/ employers who provided health insurance for them. Data on conditions, utilization and expenditures are being collected from about 11,500 households, oversampled for the uninsured population, during calendar year 1977.

The surveys of the medical providers and health insurers will begin in 1978. Questionnaires will be mailed to physicians, hospitals, and insurance companies or employers for each NMCES respondent who has signed a permission form authorizing the release of data.

## 2.1 Overall Design of the Household Survey

The original study design called for a national probability sample of the civilian noninstitutionalized population in the fifty States and the District of Columbia residing in 10,000 households. These households were to be interviewed 8 times--once every 9 weeks over a period of 16 months about medical care received during the 1977 calendar year. The first two interviews are to be in person, the next four by telephone, the seventh in person, and a final contact by telephone in April 1978. A diary was to be left with respondents and a computer generated summary was to be mailed to the households between interviewing rounds.

### 2.1.1 Sampling Protocol

Similar area probability designs were used by both RTI and NORC for the national survey. Both utilize a stratified multi-stage area probability sample. Subsets or half-samples of the full complement of Primary Sampling Units (PSU's) were selected from both organizations' national general purpose samples. The RTI and NORC half-samples contain 59 and 76 PSU's respectively and each represents the United States. Together the two independent samples overalp so that 106 separate locations are sampled. The second stage sampling units are large clusters of households (60 or more for RTI and 100 or more for NORC) from which 12 households were subsampled.

Given the analytical goals of the study, it was decided that the families and individuals identified for the sample in the initial round would be followed throughout the year even if they moved, unless they died, entered the armed forces, entered an institution, or left the country. Any new individuals who join the initial set of families or who join with a sampled individual to form a new family will be included for purposes of analyzing families but will not be included in any analysis of individuals.[8]

The sample was originally designed to provide eight rounds of data on 10,000 households. At the request of HEW's Assistant Secretary for Planning and Evaluation, this sample was supplemented with an oversample of 1,500 households which contained one or more persons without any health insurance. This was accomplished by selecting more households in the segments with a higher expected proportion of uninsured households. In order to produce a final sample of 11,500, an initial sample of approximately 16,000 households was selected to allow for an overall attrition of about 25 percent due to (1) unusable selected sampling units (e.g., vacant or demolished housing units), (2) unavailable respondents 19 years of age or older, (3) not-at-homes, (4) refusals and interview breakoffs, and (5) dropouts before end of all eight interviews.

### 2.1.2 Questionnaire Design

The questionnaire includes both items that are collected on special supplements and a repetitive core of questions asked on each round and on special one-time supplements. The core questions include the number of bed days, restricted activity days, hospital admissions, medical and dental out-patient visits, other medical care encounters, prescribed medicines, and coverage by private or public health insurance plans such as Medicaid. For each contact with the medical care system, data are obtained on health conditions, the characteristics of the provider, the services provided, the charges and methods of payment. Questions on the special one-time supplements include detailed data on health insurance coverage, access to medical care, limitation of activities, employment, income history, and socio-demographic characteristics.

### 2.1.3 Memory Aids

One of the ways NMCES addresses the problem of response error in the household survey is with the use of memory aids. A calendar/diary and a device called the summary have been developed for use in the household survey.

The calendar/diary is designed as a memory aid as opposed to a data collection instrument. It was developed and successfully used in the pilot study. It is left at the household at the end of the first personal interview. The field interviewer instructs the respondent in the use of the calender/diary and tells her that she will be asked to refer to it during the next interview. Instructions and examples are printed on the diary itself.

The diary consists of a monthly calendar with numbered date squares. Directly underneath the calendar is a ledger with boxes provided for each type of utilization reported: prescribed medicines, doctors, dentists, and other medical persons, hospitals, care in a nursing home or other similar place, and other medical items. It further identifies the reported utilization by household member, date obtained, name of provider or description of illness or injury and the cost. Beneath the ledger is a pocket with the printed instruction "Keep your medical bills in here."

The summary is a computer generated document that presents in a standard format the cumulative utilization and expenditures reported in each household for each round.

Data on utilization and expenditures collected during the household interview is keyed, processed through the computer and mailed to the household before the next interview.

The summary has separate computer printed pages for each household member. Each page contains a section for each type of utilization data collected: medical provider, hospital, prescription medicine, and other medical expenses. For each event of utilization the date obtained, the provider's name or source of care, the costs and source of payment are presented.

During the next interview, anywhere from 9 to 15 weeks later, the field interviewer systematically reviews and verifies the information with the

household. The interviewer enters any corrections, such as changes in dollar amounts or unreported visits and returns the corrected summary to RTI for rekeying and reprocessing to update the summary prior to the next interview.

Although, the primary purpose of the summary is to improve household reporting by collecting information that was forgotten or not available at the time of the survey, it has also served to legitimize the survey to respondents and to correct interviewer, respondent and processing errors.

The summary has been designed as an adjunct to the main questionnaire. Both flat fee and health insurance data from the summary are used in followup interviews if the reporting unit has reported either flat fees or health insurance plans/programs in previous interviews.

2.1.4 Achieving Desired Response Levels

In order to achieve the goal of a 95 percent response rate for the household survey and an 85 percent overall completion rate, it was decided to pay participating households a $20 incentive fee in addition to agressively pursuing traditional conversion procedures. Respondents are paid $5 after completing each of the first two rounds and $10 after the seventh. The conversion procedures include making calls at different times of the day for not-at-homes, making appointments at unusual hours for unavailables, and the use of supervisors as telephone converters for refusals.

2.2 Design of the Medical Provider Survey

The Medical Provider Survey is designed to obtain information on diagnosis, utilization, third-party payment, and total charges, and verification of payment from (1) hospitals and other institutional medical/health care facilities reported as utilized by sample individuals, and (2) physicians and doctors of osteopathy. It is necessary to ask providers of care for this information because in most instances respondents simply cannot provide charge and diagnostic information at the level of detail and accuracy necessary for our analysis plans. Data collection from providers of medical care such as dentists, nurses, chiropractors, etc., is not planned. Health insurance rarely directly covers the services of such providers and, hence, the costs of such care are usually paid directly and, therefore, known by patients. The data collection from doctors and hospitals will serve not only as a means for obtaining the required data but will also permit an assessment of the accuracy of household respondent reports on the utilization of medical care. Both household respondent over- and underreporting will be assessed using the data.

2.2.1 Physician Data Collection

The data element definitions and format for the physician data collection were developed in light of the pilot study and extensive consultation with the American Medical Association and knowledgeable individuals. In order to be responsive to the complexities of the medical environment, it was necessary to develop four separate data collection forms: the Medical Provider Questionnaire--the basic data collection instrument--and the Pregnancy Related Visit Form, the Inpatient Related Care Form, and the Repetitive Visit Form. These forms are designed to verify dates of visits reported by households and to gather detailed diganostic care information and the costs associated with that care. The Pregnancy Related Form, the Inpatient Related Care Form, and the Repetitive Visit Form are all designed to deal with the particular data collection complexities in each particular situation.

2.2.2 Hospital Data Collection

The design of these forms was influenced primarily by the experience of the pilot study and by the advice of outside consultants. Extensive consultation with the technical and research staff of the American Hospital Association and knowledgeable individuals in the hospital industry was undertakn on the format and data element definition.

Since in most hospitals there is not one central source for obtaining both cost and diagnostic data, the forms have been designed so that the cost data can be collected from a hospital accounting or billing department--the Financial Information Questionnaire--and the diagnostic data can be collected from the hospital medical records department--the Medical Information Questionnaire.

The Medical Information Questionnaire obtains the following basic information: (1) verification of the reported admission and discharge dates, (2) the respondent's chief complaint at the time of admission, (3) discharge diagnostic data including the H-ICDA or ICDA code, (4) operational and diagnostic procedural data, and (5) the names of physicians who provided care.

The Financial Information Questionnaire obtains information on the hospital charges by type for the admission, the amount of the bill already collected and the source of payment, the expected amounts of remaining payments and the sources of such payment. The questionnaire also includes a question that addresses the discounting of payments to hospitals.

2.3 The Health Insurer/Employer Survey

The insurance survey is designed to collect additional information on health insurance coverage, benefits, and premiums from two possible sources: (1) the employer, union, or group carrying the insurance; or (2) from the insurance company providing the coverage. Given the cost and complexity of collecting claims data in the 1970 Andersen survey and pilot study, it was decided not to collect claims data.

Detailed insurance coverage information is a crucial component of this study for several reasons. First, the analytical plans to simulate national health insurance proposals cannot be carried out without details of coverage, limits,

etc., which the household respondent cannot normally provide. Second, the information obtained from third-party payers will permit a comparison of how people perceive their coverage vs. their actual coverage and an assessment of the correlation of these measures with utilization. Third, premiums are a major component of medical costs each year. Both the pilot study and the pretest indicated that a majority of household respondents cannot provide accurately the amount of the premium paid by employers or other sources. The employer or insurance company is also the best source of information about premiums.

## 3. CURRENT STATUS OF THE SURVEY

### 3.1 Household

A scaled down version of the national study was pretested in Charlotte, North Carolina and Dayton, Ohio in the fall of 1976. Round 1 of the national household study began on January 17, 1977 and ended on April 1, 1977. The response rate for Round 1 was about 93 percent.

Round 2 of the national study was completed in June 1977. Preliminary results indicate that 96 percent of the households completing Round 1 completed Round 2.

During Round 1 it became obvious that processing and turning around the data for over 14,000 families within 9 weeks was next to impossible. This problem forced the delay of the start of Round 2 for three weeks and would consequently delay the start of every succeeding round, as well as start of the medical provider and health insurance surveys.

Given the necessity of keeping the survey on schedule and of collecting data for calendar year 1977, two telephone interviews were eliminated. These telephone interviews consisted only of the main questionnaire and contained no special supplements. This effectively extends the recall period between interviews, but the use of the summary and the use of the calendar/diary is expected to offset some of the memory problems.

#### 3.1.1 Uninsured Sample

Of the 1,500 additional households added to the sample to increase the uninsured segment of the sample approximately 800 households turned out to actually have one or more household members who were uninsured. The 700 insured households were dropped from the survey after the first interview to reduce costs since they could not be used to offset losses due to attrition.

### 3.2 Current-Status--Medical Provider and Hospital Survey

The physician and hospital pretest began on September 17, 1977. Althogether 215 unique physicians and 5 hospitals were identified in the Charlotte pretest.

The physicians and hospitals that are being contacted have been identified from signed permission forms obtained during the household pretest interview carried out in Charlotte during the fall of 1976. Each adult and minor age 14 and over was asked to sign a permission form authorizing the NMCES contractor to obtain additional data from physicians and hospitals on medical care and cost. For minors under age 14, the parent or guardian was asked to sign.

An attempt was made to obtain signed permission forms for every physician who provided medical care during the pretest. In addition, signed forms were also sought for those physicians who were identified as a regular source of care in the "Access to Medical Care Supplement," even if no visits to the regular provider were mentioned during the pretest. This procedure was followed so that some measure of the extent of unreported visits could be obtained. Also, all respondents who reported a hospital stay of at least one night were asked to sign permission forms. Five of the 169 Charlotte household pretest families refused to sign permission forms: these five families reported eight doctor visits and no hospital admissions.

During the medical provider pretest, the forms are mailed to providers. Three to four weeks after mailing the forms, field interviewers are to telephone the medical providers. Depending on the situation, the field interviewer will then (1) arrange to pick up the forms, (2) remind the medical providers to pick up the forms and then arrange pickup, (3) offer to assist the provider in filling out the forms, or (4) in limited cases, where there is a small case load, offer to conduct the interview over the telephone.

The procedures for the Hospital Survey are identical to those developed for the medical provider surveys.

The medical provider/hospital surveys will end on October 21, 1977.

### 3.3 Health Insurer/Employer Survey

During the Charlotte household pretest, 41 unique insurers/employers were identified. For each health insurance plan reported in the pretest, the policyholder was asked, in Round 2, to sign a form authorizing the NMCES data collection contractor to obtain additional information from the employer/third-party payer. For all policies carried through employer, unions, or some other groups for which authorization was obtained, the employer/group was mailed a copy of the combined authorization form and questionnaire. If the employer or other groups was unable or unwilling to complete the questionnaire, the insurance company was contacted.

For individual policies for which signed authorization has been obtained, the insurance companies are being mailed a copy of the

combined authorization form and questionnaire directly. All providers of the insurance coverage are being asked to provide a copy of the policyholder's (or respondent's) insurance policy. These policies are being coded at AAI to test the codebook procedures. For contacts, whether with employers or other groups or with insurance companies, who did not respond to the first mailing, the study design calls for a follow-up effort consisting of two mailings, two telephone calls, and finally a personal visit from a field representative to conduct a face-to-face interview. However, preliminary results indicate that a high level of response was obtained without the use of the full regimen of follow-up procedures.

## 4. CONCLUDING REMARKS

Preliminary data from the NMCES will be published as soon as they become available in the form of joint publications from the National Center for Health Statistics and the National Center for Health Services Research. It is anticipated that some preliminary data will be available in 1978.

## FOOTNOTES

1. The bulk of this report is drawn from internal working documents from the National Medical Care Expenditure Survey prepared by the staff of the National Center for Health Statistics and the National Center for Health Services Research, RTI, NORC, and Abt Associates, Inc. Dr. Daniel C. Walden, Mr. Robert A. Wright and Dr. Gordon S. Bonham reviewed the document and made many helpful comments.


2. Worthington, N. L. "National Health Expenditures, 1929-1974." Social Security Bulletin, 38, No. 2 (1975), p. 3.

3. U.S. Department of Commerce, Bureau of Economic Analysis. Survey of Current Business, May 1975, and U.S. Department of Commerce, Bureau of Economic Analysis. Business Statistics, 1975.

4. Jones, Sidney L. "Measuring National Wealth and Well Being." Included in Technical Paper No. 37 of the Bureau of the Census, June 1975.

5. Falk, I. S., et al. "The Incidence of Illness and the Receipt of Medical Care among Representative Families." Committee on the Cost of Medical Care, Publication No. 26 (Chicago: University of Chicago Press, 1933).


6. For a detailed description of the interviewing results of the pilot study, see Shapiro, S., Yaffe, R., Fuchsberg, R., and Corpeno, H. "Medical Economics Survey--Methods Study," a paper presented at the American Public Health Association, Chicago, Illinois, November 1975. This research was supported through Contract No. HRA 106-74-150 from the National Center for Health Statistics, Health Resources Administration, DHEW.

7. Contract No. HRA 230-76-0268 from the National Center for Health Services Research, Health Resources Administration, DHEW.

8. Additional information on the sampling design is available on request.