SAMPLE DESIGN OF THE NATIONAL MEDICAL EXPENDITURE SURVEY-HOUSEHOLD COMPONENT

Steven B. Cohen, National Center for Health Services Research and Health Care Technology Assessment, Rockville, Maryland 20857; Ralph DiGaetano and Joseph Waksberg, Westat, Incorporated

Introduction

The Household Component of the National Medical Expenditure Survey (NMES) was established to provide an assessment of the health care utilization, costs, sources of payment and health insurance coverage of the U.S. civilian noninstitutional population. The period of assessment covers calendar year 1987, during which data are collected from a national probability sample of dwelling units. The survey was designed to provide data for a major research effort in the Division of Intramural Research of the National Center for Health Services Research and the Health Care Technology Assessment (NCHSR), and was cosponsored with the Health Care Finance Administration (HCFA). The data will meet the needs of government agencies, legislative bodies, and health professionals for more comprehensive national data required for the analysis and formulation of national health policies. NMES succeeds a series of national medical expenditure surveys, most notably the 1980 National Medical Care Utilization and Expenditure Survey (NMCUES) and the 1977 National Medical Care Expenditure Survey (NMCES). Unlike the previous efforts, the NMES household sample of approximately 16,000 dwelling units includes the oversampling of groups of particular policy interest: blacks, Hispanics, the poor and near poor, the elderly and persons with functional limitations.

The NMES household survey is a year long panel, collecting measures of health status, use of health care services, expenditures and sources of payment, insurance coverage, employment, income and assets, as well as demographic information. An initial screening interview was conducted in the fall of 1986 for a sample of approximately 35,000 addresses, to obtain information required for the oversampling of specific policy relevant population subgroups. Field operations for the NMES household component consist of four core interviews conducted with the selected households at three to four month intervals over a fifteen month period. The first two interviews will be conducted in person, the third by telephone (if a telephone is available and the procedure is acceptable to the household), and the final core interview will again be in person.

The sample design can be characterized as a stratified multi-stage area probability design. The survey was conducted from two independently drawn national samples selected from two distinct survey organizations, Westat, Inc., and the National Opinion Research Center (NORC). This paper provides a detailed description of the NMES household survey design (HHS), sample yields from the screener interview and a discussion of survey design complexities which require special consideration for analysis.

The NMES Sample Design

An examination of the analytical goals of the study and budget constraints indicated that the sample design for the NMES household survey should satisfy the following requirements:

1. The full series of interviews should be completed in approximately 14,000 households.
2. The sample should be spread over at least 100 separate areas to represent the civilian noninstitutionalized population of the 50 states and the District of Columbia.
3. The sample shall be designed to produce unbiased national estimates and unbiased estimates with adequate precision for the four census regions.
4. The sample shall be designed to meet predesignated precision specifications for the following population subgroups of analytical interest: blacks, Hispanics, the poor and near poor, the elderly (65+), and the functionally impaired. It was recognized that in order to achieve the NMES precision requirements for the overall population and specified domains of interest, an initial sample selection substantially larger than 14,000 households with subsequent screening would be required.

The numerical constraint on a minimum number of geographical localities was specified to ensure sufficient geographic dispersion of the sample and allow for separate regional estimates. The precision specifications were provided to insure that the design would meet analytical objectives and to facilitate stage-specific sample size determination.

The adopted NMES household survey sample design is a stratified area probability design with four stages of sample selection: (1) selection of primary sampling units (PSU's), which are counties, parts of counties or groups of contiguous counties; (2) selection of segments within PSU's; (3) selection and screening of dwelling units within segments; and (4) selection of dwelling units based on demographic characteristics (both household and individual level) from the set of screened dwelling units.

The sample of PSU's represents a union of the national sample frames of Westat, Inc., and the National Opinion Research Center (NORC). The national general purpose area samples from Westat and NORC that comprise the NMES household sample are similar in structure, thus simplifying the development of sampling, listing and interview procedures and permitting a single management control system. Since the sampled PSU's, segments, and dwelling units selected for screening were representative of two independently drawn national samples, except for difficulties associated with survey nonresponse and other nonsampling errors, statistically unbiased national and domain specific estimates can be produced from each sample or from the two samples combined. The combined sample includes 165 PSU's located in 127 distinct sites. The number of separate primary areas is less than the total primary sampling units in the two national samples because some areas are in both samples.

More specifically, the Westat first stage sample was stratified by social, economic and demographic characteristics which included re-

gion, SMSA status, percent of population employed, percent white, and percent over age 65. In all, 81 PSU's were selected for NMES. Similarly, the NORC first stage sample included the following stratification measures: region, SMSA status, and population size. The NMES sample consisted of 84 NORC PSU's. Within PSU's, a two or three stage sample design was used to select dwelling units for the screening sample. The first stage consisted of 1980 Census Enumeration Districts (ED's) or individual block or block combinations. The second stage was only used when ED's or block groups were exceptionally large in area or number of households, respectively. Then "chunking" or the partitioning of the ED or block group was employed, dividing the selected area into several smaller segments of approximately equal size in terms of households, one of which was randomly selected. The third stage consists of the selection of all or a systematically selected subsample of households within each area segment. The ED's or blocks were selected with probability proportionate to size, with a systematic procedure allowing for implicit geographic stratification. The sampling and subsampling rates were specified so that all dwelling units in the U.S. had an equal probability of selection.

Within the sampled PSU's, 2,317 segments were selected (1,150 for Westat, 1,167 for NORC). The segment sampling process resulted in a set of maps showing the boundaries of the sampled segments and their associated probabilities of selection. The addresses within the boundaries of sample segments were then listed by trained interviewers, and served as the sampling frame from which the address sample for the NMES screener interview was selected. Approximately 35,000 addresses were selected for screening, within the sampled segments. Following the screening interview a subsample of dwelling units were selected for the full panel household survey from those screened, according to person and household level demographic characteristics. Subsampling rates were specified to obtain the required sample size to satisfy NMES precision specifications for person level estimates.

Screener Sample

As indicated, the 1987 NMES household survey was designed to provide an assessment of the annual health care utilization, expenditures, sources of payment and health insurance coverage in 1987, for the civilian non-institutionalized U.S. population. In addition, the NMES specifically targeted population subgroups of special policy interest for oversampling. More specifically, the NMES design required selective oversampling of blacks, Hispanics, the poor and near poor, those 65 years of age or older, and the functionally impaired.

The functionally impaired are defined as persons needing or receiving personal assistance or the help of special equipment for a period of three months or more to perform one or more specifically defined activities of daily living (eating, transferring from bed or chair, dressing, bathing,) because of a health or physical problem.

A set of strict precision requirements were specified to allow for detailed analyses of these policy relevant population subgroups. These precision specifications are listed in Table 1, for a 20 percent person level estimate characterizing the respective analytical domains.

The 1977 NMCES consisted of 1,290 segments with an average segment size of 12.9 per segment. Design effects generally ranged from 2 to 5 for survey estimates (Cohen, 1983). The NMES requirement of an average design effect of 1.7 for survey estimates was specified to insure an efficient survey design. To satisfy this requirement, an average segment size of 6 households per segment was planned for NMES. Given the overall sample size requirement of 14,000 household interviews, 2,317 segments were selected, nearly twice the number considered in NMCES.

Due to the complex sampling requirements in the NMES, a separate screening interview was used to facilitate sample identification of population subgroups targeted for sampling at differential selection rates. The separate screening operation allowed for analyses of the results of the screener interview prior to the selection of the NMES Round 1 household sample. More specifically, it allowed for analysis of three features of the sample before the design was finalized: (1) a determination of whether the number of households screened was sufficient to provide the required sample sizes for each analytical domain; (2) whether the person based domain requirements were consistent with a sample size specification of 14,000 households completing the full series of interviews; and (3) development of the subsampling rates needed to attain the desired sample size in each domain. An advanced screening operation would make available precise counts of the outcome of the NMES subsampling process, prior to fielding the Round 1 household sample.

A self-weighting sample design was developed for the NMES screener interview to insure an efficient sample. The number of screening interviews required to meet targeted precision specifications was determined by identification of the demographic category which required the highest sampling rate for inclusion in the NMES household survey. The estimated sample size requirement for the NMES household screening interview was driven by the precision requirement for the black, 65 +, poor population subgroup. The households that completed the screening interview also served as a base from which individuals associated with the remaining demographic categories could be identified for inclusion in the NMES.

Since the NMES precision specifications are based on individuals completing the full series of NMES household interviews (4 rounds of data collection), additional adjustments which control for survey nonresponse and vacancy rates had to be incorporated in the screener sample size specifications. Using the experience of Current Population Survey, the estimated occupancy rate in the U.S. was 92.5 percent. Further, the expected response rate for the screener interview was 95 percent, and the expected overall response rate for Rounds 1 through 4 of the household survey was 85 percent. Consequently, the number of addresses to be selected for the NMES screener sample was specified as 36,150, to obtain 210 respondents in the black, 65 + poor subgroup in

round 4 (Table 2).

The targeted screener sample of 36,150 addresses was equally divided across data collection organizations requiring a selection of 18,075 sample addresses from the 1,150 sample segments that characterized the Westat national area probability sample, and a selection of 18,075 sample addresses from the 1,167 sample segments that characterized the NORC national area probability sample. For both organizations, the probability of selection is constant for all addresses sampled for the screener interview, regardless of the PSU or segment in which an address was located.

The screener sample consisted of dwelling units, although the basic analysis units are persons. The sample dwelling units (DU's) include housing units, group quarters, and other non-institutional (non-group) living quarters. All civilians who considered the selected DU as their usual place of residence were included in the interview.

Analysis is planned at both the individual and what is referred to as the reporting unit level. A reporting unit consists of one person living alone or unrelated to others in the same dwelling unit, and two or more persons related to each other by blood, marriage, adoption, or foster care whose usual residence is the assigned DU. More than one reporting unit may be living in the selected DU. Reporting units are comparable to census classifications of unrelated individuals plus families.

NORC selected its sample of 18,075 addresses after all the listing information was established on a computer data base. Due to time constraints, Westat selected its sample of addresses from listed segments on a flow basis as they became available. An overall sampling rate was chosen based on the estimated number of dwelling units for 1986. Each segment was then sampled individually at a rate which provided a constant overall sampling fraction. Since the total number of listed units was not known precisely in advance, the exact overall sample size could not be controlled.

Westat's initial sample selection consisted of 17,016 addresses. An additional 264 cases were added during Westat's quality control operations. The 4.5 percent short fall in sample size was primarily due to an over-estimate for 1986 of the number of dwelling units and a missed rate of 2.5 percent in the listing operation.

NMES Household Screening Interview - Field Results

The NMES screener interview was conducted during the fall of 1986 (October to December) at dwelling units selected from the respective Westat/NORC national area samples. An advanced letter describing the purpose of the NMES was mailed to each selected address. Interviewers visited each selected address to determine whether it was occupied as a primary residence, and attempted to conduct the interview with a household member. The household screener enumerated persons living in the selected dwelling unit, identified reporting units within the dwelling unit, and obtained demographic and other information to permit the subsampling of dwelling units to achieve the desired sample sizes for various domains of interest. The person identified as the reference person by the household member, and all other persons in the household related to the reference person formed a primary reporting unit (RU). Persons unrelated to the reference person also living in the sampled dwelling unit were grouped into other reporting units, with all related persons in the same reporting units. A separate screening interview was to be completed for each reporting unit, identified in a dwelling unit.

The final NMES household screener sample consisted of 35,634 addresses, of which 3,091 were identified as vacant and another 1,085 identified as not a dwelling unit. Of the 31,458 dwelling units eligible for the NMES, 28,708 responded to the screener interview. A dwelling unit was classified as responding if any of its component reporting units completed the screener interview. Consequently, the overall NMES household screener response rate was 91.3 percent. A more detailed breakdown of the final status classifications for the NMES screener sample is provided in Table 3.

NMES Round 1 Household Sample

The sample design had several features not usually required in area samples. First, initial precision requirements for the household survey were stated in terms of national estimates at the person level (Table 1). To meet or exceed these requirements, the survey must include at least the precalculated number of persons in each subdomain of interest. However, the unit of interviewing was the reporting unit and subsampling was done at the dwelling unit level. Thus, screened households were selected for the full panel HHS on the basis of the characteristics of the persons they included. Secondly, for both cost-efficiency, and to maximize the response rate, Round 1 was characterized by a "housing unit" design. That is, while households were selected based on the characteristics of their members, the Round 1 sample unit was the address of the household. The Round 1 interview was thus conducted with persons residing at the sample address at the time of the interview whether or not they lived there at the time of screening. These persons are then included throughout the full panel HHS, so long as they are in the civilian non-institutionalized population. Finally, a further complication was that subdomains are not mutually exclusive -- for example, a person can be both elderly and Hispanic. Consequently, in determining required sample sizes it was necessary to group the United States population into mutually exclusive cells by cross-classifying the variables that define the subdomains of interest. This cross-classification is presented in Table 4, in addition to the desired yields of sample respondents completing the full 4 rounds of NMES data collection, and their targeted subsampling rates.

The expected number of sample respondents at round 4 were derived by applying the approximate subsampling rates to a completed screener sample of 31800 households under the following assumptions: (1) 2.7 persons per household ,(2) an 85 percent overall Rounds 1-4 response rate, and (3) the March 1985 CPS estimated population distribution for the specified subdomains also character-

ized the 1987 distribution (Table 4).

To achieve a reduction in the overall costs of the NMES household survey, a modest reduction in the sample was implemented prior to the selection of the Round 1 sample. Generally, a six percent reduction in the Round 1 sample from initial specifications was implemented across each of the mutually exclusive population subdomains with one exception. An initial reduction of 16.67 percent in the Hispanic sample was specified, prior to the additional reduction of 6 percent. Consequently, the NMES Hispanic sample was reduced 21.67 percent from original specifications. In addition, the NMES sample of functionally impaired individuals 65 years of age or over was increased slightly over original specifications. Table 4 also presents the expected number of sample respondents at round 4, after implementing the sample reduction.

In the actual sample draw, households were selected sequentially. Households containing individuals targeted for selection at the highest rates were selected at the initial stages, one demographic category at a time. During the initial stages of sample selection, other members of sampled households who do not possess the demographic characteristic of the household classification were tallied in accordance with the remaining demographic categories, and the counts were subtracted from the sample quotas for the other domains. Thus, sampling rates for each category will change after each cycle of selection, from initial specifications based on the cumulative samples obtained for each category.

This sequential selection procedure permits control of sample sizes obtained for all demographic categories of interest so as to meet the specified precision levels. Because such a sequential procedure results in variation in sampling rates within demographic categories, there will be a slight increase in the variance of sample estimates attributable to the differential rates. However, for most groups this increase is quite small and has been accounted for in determining the sample sizes needed to meet the specified precision levels.

The NMES Round 1 household sample was selected after the data collection phase for the NMES screening interview was completed and all relevant demographic data necessary for the sample selection of responding dwelling units (race/ethnicity, age, functional impairment level, and income) was entered into the screener data base. To allow for the monitoring of sample yields from the NMES screener interview on an ongoing basis during the data collection phase of the screening operation, and to facilitate the availability of a complete screener data base in early January, 1987, the screener data was entered into a preliminary data base by a computer assisted data entry system upon receipt.

All eligible individuals within an eligible DU were assigned to one of the 23 mutually exclusive categories specified in Table 4 (and also to a functionally impaired category when appropriate). All individuals in the same reporting unit were given the same income classification. For the purposes of sample selection, an imputation strategy was considered for missing values at the person level for the following key sampling variables: age, functional impairment, poverty status, and income. The modal value for sample respondents was used as the imputed value for the NORC sample. The same strategy was considered for the Westat sample for age and functional impairment. Westat data at the segment level on race/ethnicity and income, used for sample selection, was also used to impute a value of race/ethnicity or poverty status for individuals with missing data who resided in the selected segment.

Results of the NMES monitoring reports indicated a substantial departure from expectation in the poverty status distribution. In particular, sample yields for the near poor classes were generally larger than expected, and the corresponding yields for the other income classes were slightly smaller, while the "at or below poverty" classes showed an erratic pattern when compared to CPS estimates. The discrepancies in the income distribution appeared to reflect reporting errors rather than sampling problems. Since there were only small differences in the required sampling rates for many of the income classes within demographic groups, it was decided, for sampling purposes, to eliminate some of the domains. Income groups were therefore mostly collapsed within the following demographic groups: blacks, Hispanics, white 65-79, white 80+, and white less than 65 (poor/near poor). In addition, to minimize the effect of differential sampling rates on the precision of survey estimates, the "other races, near poor/other income" was combined with the "white, less than 65, other income" class. Prior to sample selection, dwelling units in the collapsed domains were sorted by income and race/ethnicity. Further, the separate domain for the functionally impaired, under 65 category was eliminated as a distinct sampling domain. As a result of collapsing the initial set of sampling domains, ten final sampling classes were specified.

The final ten collapsed sampling domains are presented in Table 5. With only one exception, dwelling units within each of these sampling classes were hierarchically sorted on the dimensions of race, poverty status, PSU, segment and household size. For the functionally impaired domain, dwelling units were only sorted by PSU, segment, and household size. A systematic sample of dwelling units was then selected within each of these classes. Since it seemed likely that some of the white persons under 65 who reported themselves as near poor actually had higher incomes, the targeted sample size for this domain was increased by approximately 300, to make sure the desired sample size was attained. The targeted sample for white, under 65, other income was reduced by the same margin.

The Round 1 household sample of dwelling units was then selected by the following sampling strategy (DiGaetano,1987):

1. Each screened DU was classified into one of the ten sampling categories based on the demographic characteristics of its "highest priority" individual. This was the household member requiring the highest sampling rate to meet sample size targets. The sampling rate was defined as the ratio of the required number of individuals in a sampling class to the available number of screener respondents. The

sampling classes presented in Table 5 are arranged in order of highest priority.
2. A systematic sample of DU's was then selected from the highest priority class using the person level sampling rate.
3. The number of individuals in all sampled DU's that possessed the demographic characteristics that defined the remaining sampling classes were counted. The sample size requirements for the remaining sampling classes were then reduced by these enumerated totals.
4. DU's associated with the demographic category that was sampled were then removed from the sampling process.
5. DU's associated with the next highest priority sampling class were then selected at a sampling rate designed to meet sample size requirements, repeating the process outlined above.
6. The sampling process is repeated until all ten sampling categories have been sampled.

Sample of Addresses Vacant During the Screener Field Period

It was expected that some of the addresses contacted which were vacant in the screener field period would become occupied at the time of the Round One interview. Excluding them from the sample would understate the number of recent moves in the sample. Consequently, a sample of vacant addresses was selected to supplement the occupied addresses sampled from all screened households. Since eight percent of household addresses are expected to be vacant at any time, a sample of 1464 vacant addresses identified during the screening field period were selected, representing approximately eight percent of the total addresses to be visited in Round 1. It was assumed the number of vacant addresses becoming occupied would compensate for the number of occupied dwelling units becoming vacant in Round 1.

The set of 3,091 addresses identified as vacant during the screener field period were classified into one of four categories, based on the demographic characteristics of the sample segments from which the vacant addresses were drawn. The four classes were defined as black, Hispanic, white poor, and white other, reflecting the dominant characteristic of the individuals in the segments in which the vacant addresses were located. The sampling rates that were applied for the sample of vacant addresses were the minimum sampling rate associated with the stratum from the main sample draw. For example, if the category "Hispanic, under 65" had a lower sampling rate than "Hispanics, 65+", then the rate for 'Hispanics, under 65" was used to sample vacant addresses in segments classified in the Hispanic stratum.

Sample of Screener Refusals and Other Nonresponding Dwelling Units

A supplemental sample of refusals to the screener and other nonresponding dwelling units was also included in the sample, to improve the overall NMES response rate which reflects four rounds of data collection. In addition to 1600 dwelling units classified as screener interview refusals, another 916 dwelling units were categorized as "other nonresponse". The other nonresponse classification included dwelling units where no one was home after 4 calls, where the potential household respondents were unavailable during the screener field period, and where respondents were considered too ill to complete the interview.

Refusals and other nonresponding dwelling units were classified into the same four strata used for vacant addresses. A sample of 645 DU's categorized as refusals to the screener were then selected from the 1600 eligible DU's and a sample of 376 DU's categorized as other nonresponding DU's were selected from 916 eligible DU's with this classification. Consequently, the only set of screener nonrespondents that were not considered for selection in the first round of the NMES household survey consisted of DU's with a language problem, DU's where the interviewer was unable to enter the structure, and the remaining nonresponse classifications. These 234 DU's were not targeted for selection in Round 1 because of the low probability of converting them to participate in the four rounds of the household survey. Theoretically, the potential response rate for the Round 1 household interview is 99.26 percent, if all screener refusals and other nonresponding DU's participated in this interview. The final Round 1 sample is summarized in Table 6.

Summary

The complex survey design of the household component of the National Medical Expenditure Survey has been described in detail. Particular attention has been given to the disproportionate sampling strategy used to oversample blacks, Hispanics, the functionally impaired, the poor and near poor, and the elderly. Field results from the screener interview are also presented in addition to addresses vacant during the screener field period, of screener nonrespondents, and of screener refusals.

References

Cohen, S.B., and W.D. Kalsbeek (1981). National Medical Expenditure Survey: Estimation and Sampling Variances in the Household Survey. National Center for Health Services Research, Instruments and Procedures Series, No. 2. Department of Health and Human Services Publication No. (PHS) 81-3281. Washington, D.C.: Government Printing Office.

Cohen, S.B. (1983). Design Effect Variation in the National Medical Care Expenditure Survey. Proceedings of the Survey Research Section, 748-756.

Cox, B.G., and S.B. Cohen (1985). Methodological Issues for Health Care Surveys. New York, Basel: Marcel Dekker, Inc.

DiGaetano, R.G. (1987). Sampling Report for the Selection of Dwelling Units for the Household Survey of the NMES. NMES Report No. 1.00. Westat, Inc., Rockville, Maryland.

DiGaetano, R.G., and J. Waksberg (1986). Sample Design of the Household Survey (HHS) and the Survey of American Indians and Alaska Natives (SAIAN). NMES Report No. 1.46. Westat, Inc., Rockville, Maryland.

Kish, L. (1965). Survey Sampling. New York: John Wiley and Sons.

The views expressed in this paper are those of the author and no official endorsement by the National Center for Health Services Research and Health Care Technology Assessment is intended or should be inferred. The authors wish to thank Ms. Evelyn Manstof, Mr. Bill Ashby,and Ms. Mary Gore for their skillfull typing of the manuscript. The referenced Tables are not presented in this paper due to space limitations. They may be obtained from the author by writing to: Dr. Steven B. Cohen, NCHSR, 5600 Fishers Lane, Room18-A55, Rockville, MD 20857.