SAMPLE DESIGN OF THE NMES SURVEY OF AMERICAN INDIANS AND ALASKA NATIVES

Steven B. Cohen, NCHSR, Ralph DiGaetano, and Joseph Waksberg, WESTAT
Steven B. Cohen, 5600 Fishers Lane Room 18 A55, Rockville, MD 20857

Introduction

The Survey of American Indians and Alaska Natives (SAIAN) component of the National Medical Expenditure Survey was designed to provide an assessment of the health care utilization, costs, sources of payment and health insurance coverage for American Indians and Alaska Natives living on or near reservations and eligible for Indian Health Service (IHS) care. The period of assessment covers calendar year 1987, during which data were collected from a national probability sample of dwelling units with at least one member of the civilian, non-institutionalized American Indian and Alaska Native population who was eligible for IHS services. As a household survey, the SAIAN complements the Household Component of the National Medical Expenditure Survey (NMES), which was established to provide an assessment of the health care experience of the U.S. civilian noninstitutionalized population.

The survey was designed to provide data for a major research effort in the Division of Research of the National Center for Health Services Research and Health Care Technology Assessment (NCHSR), and was cosponsored with the Indian Health Service (IHS). The data will meet the need of government agencies, legislative bodies, and health professionals for more comprehensive national data required for the formulation of policies relating to the Indian Health Service. The 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). The SAIAN component of the NMES, however, reflects a new dimension to the previous efforts, and is the first national survey of its type to be conducted.

The SAIAN 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. In addition to obtaining data comparable to the NMES household survey, the SAIAN also collected information on the use of IHS facilities by this population relative to other health care providers, and obtained associated charges and source of payment data. An initial screening interview was conducted in the winter and spring of 1987, to identify households with at least one member who was an American Indian or Alaska Native and also eligible to receive Indian Health Service care.

Field operations for the SAIAN component consisted of three core interviews conducted with selected households at five to six month periods over a sixteen month interval. Westat, Inc. was the prime data collection organization responsible for fielding the survey. The sample design can be described as a stratified multi-stage area probability design. Due to cost constraints a disproportionate sampling strategy was adopted whereby households in high Indian density areas were sampled at higher rates than Indian households in low Indian density areas (DiGaetano, 1987; DiGaetano and Waksberg, 1986).

Since the survey is the first national health care survey of this scope for the American Indian and Alaska Native population, its design should be of particular interest to survey planners that will conduct future surveys for comparable population subgroups. This paper provides a detailed description of the SAIAN survey design, the sample yields from a screener interview that determined survey eligibility, and a discussion of survey design complexities that require special consideration for analysis. In addition, the field results that characterize the respective rounds of data collection for the core SAIAN interviews are also presented.

The SAIAN Sample Design

An examination of the analytical goals of the study and budget constraints indicated that the sample design of the NMES Survey of American Indians and Alaska Natives should satisfy the following objectives:
(a) The sample should be designed to produce national estimates, and to the extent possible, separate estimates for the following seven groups of IHS Areas/Programs: (1) Navajo, (2) Alaska, (3) California, (4) Bemidji, Nashville, and Portland, (5) Oklahoma, (6) Aberdeen and Billings, and (7) Albuquerque, Phoenix, and Tuscon.
(b) The precision specification for the national estimates was that the relative standard error for a 20 percent person level estimate would not exceed 6 percent, for the overall SAIAN population.
(c) A cost efficient sample design was to be used.

These objectives were met by a sample design with the following features:
1. The full series of interviews should be completed in approximately 2,000 households.
2. The sample should be spread over 20 separate areas, selected from a list of eligible counties identified by the Indian Health Service, to represent the civilian noninstitutionalized population of American Indians and Alaska Natives living on or near federal reservations and eligible for IHS care. Two of the separate areas were to be certainty selections (Apache County, Arizona, and McKinley County, New Mexico), by virtue of their large Indian populations and their high density rate of Indian population relative to their respective total county population.
3. The sample of addresses selected for screening, in order to identify the households eligible for the SAIAN, was to be constrained at 17,500 addresses.

The adopted SAIAN sample design is a stratified area probability design with three stages of sample selection: (1) selection of primary

sampling units (PSU's), which are counties, or groups of contiguous counties (in Alaska, the county equivalents developed by the State of Alaska and the Census Bureau for statistical purposes were employed) ; (2) selection of segments within PSU's; (3) selection and screening of dwelling units within segments. Based on the results of the screening interview, all dwelling units that include at least one American Indian or Alaska Native eligible for IHS care were to be selected in the SAIAN.

The SAIAN sample frame initially consisted of 482 counties in the United States served by the Indian Health Service, with an estimated American Indian and Alaska Native population of 1,013,000 projected for 1987. For cost efficiency, the frame was truncated to exclude counties with fewer than 400 American Indians or Alaska Natives. This strategy removed 184 counties from consideration for the SAIAN sample, with an estimated target population of 28,000. Consequently, the truncated frame included 97.2 percent of the population of interest.

Primary sampling units for the SAIAN were then formed from the remaining 298 counties. Twenty four counties with American Indian and Alaska Native populations of between 401-500 were paired with larger neighboring counties to form 274 PSU's available for sample selection. As noted, two PSU's (Apache County, Arizona, and McKinley County, New Mexico) were selected with certainty. The remaining PSU's were stratified and an additional 18 PSU's selected in the SAIAN.

Selection of Non-Certainty PSU's

The initial stratification measure employed for selecting the non-certainty PSU's was the location of the IHS Area/Program. Seven groupings were defined for the SAIAN: (1) Navajo, (2) Alaska, (3) California, (4) Bemidji, Nashville, and Portland, (5) Oklahoma, (6) Aberdeen and Billings, and (7) Albuquerque, Phoenix, and Tucson. Except for the IHS areas of California and Alaska, the PSU's were further explicitly stratified by the degree of health service usage for the geographic area (high usage, other usage), as assigned by the Indian Health Service. Within these ten strata, the PSU's were sorted by dichotomous variables (high, low) representing the percent of the population that was urban and the percent that was American Indian or Alaska Native (the PSU's for each IHS area were classified as above or below the median for the respective measures within the area). A final sort variable was the median household income for the PSU.

Since no PSU was considered a "high" health service usage PSU in Alaska, a stratification variable was formed reflecting level of urbanization. More specifically, one contained highly urban PSU's (Juneau, Anchorage, and Fairbanks) and the other included all remaining SAIAN eligible PSU's. PSU's were then further sorted by high/low classification of PSU's in terms of percent of American Indian and Alaska Native population and median income.

The number of non-certainty PSU's targeted for sample selection in the explicit strata closely followed the PSU allocation to be expected under proportionate allocation, based on the SAIAN

eligible population. Overall, one sample PSU was selected to represent the California IHS Area, and two PSU's were selected to represent the Alaska IHS service area. For the remaining IHS service areas, the 15 non-certainty PSU's were to be selected under a proportional allocation scheme based on population distribution of American Indians and Alaska Natives.

The actual number of PSU's selected for each stratum using a modification of Bryant's allocation procedure (Bryant et al., 1960) are presented in Table 1. Using this procedure, the expected number of PSU's assigned to a stratum is proportional to the stratum population, but the marginal allocation could be modified.

After the explicit and implicit stratification of PSU's was complete, a systematic random sample of PSU's was selected with probability proportionate to the total number of American Indians and Alaska Natives. The measure of size for each PSU was standardized so that the sum of the measures of size within a particular explicit stratum equaled the number of PSU's to be sampled for the stratum. Since the sample allocation considered a probablistic method, the initial SAIAN PSU weights (defined as the reciprocal of their selection probability) were adjusted to reflect the sample distribution across explicit strata expected under a proportional allocation scheme. Table 2 identifies the PSUs selected for the SAIAN, in addition to the projected 1987 American Indian and Alaska Native population within the IHS Program Areas prior to frame truncation.

Selection of the Sample of Segments

Within each sample PSU, a one or two step process was used to create the second stage units, or SAIAN area segments. Segments were generally defined as 1980 Census Enumeration Districts (ED's) or individual blocks or block combinations. A two step process was considered when an ED or block grouping was exceptionally large in area or in number of households. The initially identified second stage units were then partitioned, whereby the ED or block group was divided into several smaller segments of approximately equal size in terms of households, one of which was to be randomly selected.

Initially, a self-weighting design was proposed for the SAIAN, with an average segment size of 6 households per segment was planned. Given the overall sample size requirement of 2,000 completed household interviews, 333 segments were required to meet the average segment size specification. As design work progressed in the SAIAN, it became evident that if proportional allocation were employed to represent the SAIAN eligible population in the 20 sample PSU's, the costs associated with screening interviews would be prohibitive. The level of screening required for a self-weighting design would seriously reduce the number of core SAIAN interviews that would be permitted under a fixed budget.

To address this budgetary restriction, the sample of addresses to be screened was constrained at 17,500, while the requirement of obtaining the full series of interviews in 2,000 households was maintained. The most efficient modification under these constraints was to select segments at disproportionate rates, with

741

lower density segments sampled at rates lower than would be expected under proportionate allocation. However, the number of segments to be selected was maintained at 333.

To further reduce expected costs associated with inefficient screening activities, the sample frame for SAIAN was further restricted by excluding segments with very low relative concentrations of American Indians and Alaska Natives from the sampling frame. More specifically, segments with less than 0.5 percent population representation of American Indians and Alaskan Natives were excluded from consideration for SAIAN segment selection. By restricting the exclusion of segments to the very low density ones, only a small proportion of the eligible population is missed, and the potential for biased estimates is minimal. Approximately 95 percent of the American Indian and Alaska Native population remained eligible for sample selection, when the effects of excluding counties with small Indian populations and segments with low concentrations of Indians are jointly considered.

As noted, the disproportionate sampling strategy adopted for the selection of SAIAN sample segments was driven by a sample size constraint in the number of screening interviews. The differential rates of selection were constrained, however, to control for the increase in variability in SAIAN population estimates relative to what would be expected under a proportionate allocation scheme. Here, the design effect due to the disproportionate sampling strategy was constrained at 1.75. Furthermore, the SAIAN sample design was constrained by an overall precision specification, whereby the relative standard error for a 20 percent population estimate would not exceed 6 percent.

Prior to selecting the SAIAN sample segments, six strata were defined to facilitate the disproportionate sampling strategy. The low density stratum included segments with American Indian population concentrations of between 0.5 to less than 5 percent. The medium density stratum included segments with American Indian population concentrations of between 5 to less than 10 percent. A 10 percent or greater concentration was required for the high density stratum. In addition, a separate stratum was specified for Alaska, and two separate strata were created for Delaware, Oklahoma (medium density, high density). Within strata, the segments were implicitly sorted on the following additional stratifiers: on or off reservations, percentage of American Indians or Alaska Natives in the ED or block group, and rent or value of home. A systematic sample of segments within explicit strata was then selected, with an overall target of 333 segments. Here, the measure of size that was considered within each stratum was proportional to the number of Indians within the segment, further adjusted by the PSU weight.

The large probabilities of selection in high density Indian areas, combined with the large average size of ED's, resulted in a substantial representation of the Indian population living in certainty ED's and block groups. Many of these certainty segments were selected multiple times, yielding 251 separate locations. Under this sampling scheme, the proportion of sampled households selected from the certainty segments will correspond to the proportion of the population represented by these segments. The final allocation scheme for segment selection across the six explicit sampling strata, in addition to relative sampling rates and expected households completing all data collection rounds of the SAIAN, is presented in Table 3.

For the Delaware, Oklahoma PSU, the expected number of sample segments under the disproportionate allocation scheme was 51, which was inordinately high for a single PSU allocation. Prior to segment selection, the number was reduced to 34. In a similar manner, the expected number of sample segment under the disproportionate sampling strategy for Alaska was 36, which exceeded the number of segments initially budgeted for this relatively remote location by 44 percent. Consequently, the number of segments specified for selection in Alaska was reduced to 25. To preserve the overall number of sample segments, there was a redistribution of sample segments from Delaware, Oklahoma and Alaska, to the remaining PSU's.

To further reduce costs associated with screening in the selected segments, the average number households per segment was varied across strata. As can be noted in Table 3, the average number of households targeted per low density segment was reduced from 6.0 to 5.2, which translates to a 13.33 percent reduction in the overall number of completed interviews targeted for this stratum. To preserve the overall number of completed SAIAN interviews at 2,000, there was a redistribution of sample households to the segments which characterized the high density stratum.

SAIAN Screener Sample

To facilitate the selection of the SAIAN screener sample, the addresses within the boundaries of the sampled segments were listed by trained interviewers during the summer and fall of 1986. These addresses then served as the sampling frame from which the address sample for the SAIAN screener interview was selected. Since the SAIAN precision specifications are based on individuals completing the full series of SAIAN interviews, additional adjustments which control for screening rates and survey nonresponse, had to be incorporated in the screener sample size specifications.

During the listing and screening effort, there were thirteen sampled segments (11 unique segments) associated with the White Mountain Apache tribe that refused to participate in the survey. These segments were all located in either Navajo county or Gila county, Arizona. The effect of this exclusion of segments associated with the White Mountain Apache tribe was to further restrict the SAIAN target population. Consequently, these segments were removed from the frame for SAIAN sample household selection. Furthermore, no reduction in the overall SAIAN household sample size specification was considered.

The expected response rate for the screener interview was 95 percent, and the expected overall response rate for all rounds of SAIAN data collection was 85 percent. Consequently, to

obtain approximately 2000 households completing the full series of interviews (the exact specification was 1,984 households), it was necessary to target interviews for 2,457 SAIAN eligible households (Table 4). It was further determined that 17,416 addresses would need to be designated for screening, to obtain 2,457 dwelling units eligible for the SAIAN survey.

As noted, disproportionate sampling rates were employed in the SAIAN in order to achieve the dual targets of (1) 15,000 completed screening interviews and (2) 2,334 completed screening interviews for SAIAN eligible dwelling units, while keeping the increase in variance due to the departure from a proportionate allocation scheme to a reasonable level. Table 5 presents the sampling rates employed for the sample selection of addresses, the actual number of addresses selected and the estimated yield of SAIAN eligible dwelling units, in addition to information used in establishing the sampling rates. By specifying sampling rates that were function of occupied American Indian and Alaska Native dwelling units, the sampling rates specified for the selection of addresses implicitly adjusted for expected vacancies and non-dwelling units.

(1) Ratio of high density sampling rate to stratum sampling rate.
The 1987 estimate of the number of American Indian and Alaska Native dwelling units was driven by assumptions related to (1) the number of persons per occupied dwelling unit and to (2) the growth rate in the number of dwelling units from 1980. Using 1980 Census data for the sampled segments, an unweighted estimate of the mean number of persons per occupied dwelling unit was obtained by dividing the number of people in the segments by the corresponding number of occupied dwelling units. This was done separately for each of the six explicit sampling stata. The estimate derived for the general population was then assumed to also apply to the SAIAN eligible population. Similarly, the growth rate was estimated by the ratio of addresses listed in the sampled segment in the latter period of 1986, to the number of occupied and vacant dwelling units found in the 1980 Census. This was also done separately for each of the sampling strata, and the general population ratio was assumed to hold for the SAIAN eligible population.

In SAIAN, the probability of selection is constant for all sample dwelling units within each of the six sampling strata, regardless of the PSU or segment in which a household is located. Once the sample segments were listed, the segment weights were applied to the number of addresses listed in the respective segments, to obtain an estimate of the total population of dwelling units eligible for the screening interview within each of the sampling strata. The sampling interval for each stratum was then specified as the reciprocal of the stratum specific selection rate, and after the selection of a random start, the sample of addresses was selected in the following manner. Within each explicit sampling stratum, the listed addresses were sorted by PSU and segment. The segment weights were then applied to the addresses listed within the sampled segments, and the weighted sum

cumulated across the ordered addresses. The address whose cumulated weighted sum included the value indicated by the cumulated skip interval was then selected for the SAIAN sample.

SAIAN Field Procedures and Definitions
The screening sample for the SAIAN consisted of dwelling units, although the basic analysis units for the survey are individuals. 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.

For the SAIAN, a housing unit is defined as a house, apartment, group of rooms or a single room which is occupied as separate living quarters or vacant but intended for occupancy as a separate living quarters. The people who live there must eat and live separately from everyone else in the building, and the room must have either a separate entrance directly from the outside of the building or through a common hall, or kitchen facilities for the use of the household only.

A group quarter is defined as a single civilian structure in which five or more unrelated persons reside and where inhabitants are not considered a part of any other housing unit. Examples are boarding and rooming houses, hospital staff quarters, monasteries and convents, communes, bunk houses, flop houses and missions.

Based on the results of the SAIAN screener interview, a dwelling unit was considered eligible for the SAIAN if it contained at least one American Indian or Alaska Native who was eligible to receive health services provided by or paid for by the Indian Health Service. Included in a sample household within the dwelling units are persons who are considered to be a part of the household, but who are temporarily residing elsewhere. Unmarried college students under age 22 were selected at their parents' sample addresses rather than at colleges in order to allow for the derivation of health care utilization and expenditure estimates at the family level. College students living at college are interviewed at their college address for the core SAIAN interviews.

Analysis is planned at both the individual and what is referred to as the reporting unit (RU) level. A reporting unit consists of a 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.

Definitions of respondents considered to be "key" or "non-key" were established for the SAIAN, with the understanding that data would be collected for households with key individuals for all of 1987. For the purposes of SAIAN, an individual was considered to be key if the following conditions held:
(1) the individual was an American Indian or Alaska Native who is eligible for IHS services, and
(2) the individual is either present at the first round of data collection or moved into the

reporting unit after the first round but did not live in an IHS-eligible county between January 1, 1987 and the time he or she moved into the reporting unit. New babies and persons not in the civilian, noninstitutionalized population on January 1, 1987, who join sample households after Round 1 and also meet criterion (1) will be considered as key individuals for the purposes of SAIAN data collection.

SAIAN Screening Interview - Field Results

The SAIAN screener interview was conducted during the winter and spring of 1987 at dwelling units identified from the selected address sample. Due to the large number of screening interviews that had to be conducted over a limited time period in the low density segments and some of the medium density segments, a separate screening interview was carried out in dwelling units in these areas. A concurrent screening interview was conducted at the dwelling units in the remaining segments, whereby interviewers conducted the screening interview and immediately proceeded to ask for participation of eligible households in the core SAIAN Round 1 interview.

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 SAIAN household screening interview enumerated persons living in the selected dwelling unit, identified reporting units in the selected dwelling unit, and obtained critical information on SAIAN eligibility. The person identified as the reference person by the household member, and all other persons in the the household related to the reference person formed a primary reporting unit (RU). Persons unrelated to the reference person also living in the sample dwelling unit were grouped into other reporting units, with all related persons in the same reporting units. A separate screening interview was attempted for each reporting unit identified in the sample dwelling units. The following PSU's contained segments that were targeted for the separate screening operation: (1) Clark, NV, (2) Bernalillo, NM, (3) San Bernardino, CA, (4) Washington, RI, (5) St. Louis, MN, (6) Itasca, MN, (7) Wagoner, OK, (8) Oklahoma City, OK, and (9) Tulsa, OK. Of the 17,416 addresses selected for the SAIAN screener sample, 12,446 were targeted for the independent screening interview. While in the field, the interviewers implemented a procedure that checked the accuracy of the listing operations. The procedure also insured that missed dwelling units discovered in the field were selected in the sample at the same rate as the listed dwelling units. On occasion, entire segments were required to be relisted. Consequently, the actual number of dwelling units that comprised the independent screening interview was somewhat greater than the initial target.

The final independent SAIAN screener sample consisted of 12,856 addresses, of which 1,273 were identified as vacant and another 393 identified as not a dwelling unit. Of the 11,190 dwelling units eligible the the interview, 10,217 responded to the interview. A dwelling unit was classified as responding if any of its component reporting units completed the screener interview. Consequently, the independent screener response rate was 91.3 percent. However, it was determined that 554 of the nonresponding dwelling units were ineligible for the SAIAN, from information provided to interviewers by neighbors. From that perspective, information was available to determine eligibility for the SAIAN for 96.3 percent of the eligible dwelling units. In all, 274 of the responding dwelling units were determined to be eligible for the SAIAN. A more detailed breakdown of the final status classifications for the independent SAIAN screener interview is provided in Table 6.

For the remaining dwelling units, a concurrent screening interview was conducted, to be followed with the Round One SAIAN household interview. Of the 17,416 addresses selected for the SAIAN screener sample, 4,970 were targeted for the concurrent screening interview. As in the independent screening interview, application of the missed dwelling unit procedure and occasional relisting of segments resulted in a modification from the initial target. The final concurrent SAIAN screener sample consisted of 5,025 addresses, of which 767 were identified as vacant and another 388 identified as not a dwelling unit. Of the 3,870 dwelling units eligible for the interview, 3,522 responded to the interview. Consequently, the concurrent screener response rate was 91 percent. However, it was determined that another 52 dwelling units were ineligible for the SAIAN, from information provided to neighbors on nonresponding households. This strategy yielded information to determine SAIAN eligibility for 92.4 percent of the eligible dwelling units. In all, 1,756 of the responding dwelling units were determined to be eligible for the SAIAN. A more detailed breakdown of the final status classifications for the concurrent screener interview is provided in Table 7. When these results are combined with the independent screener, the overall SAIAN screener response rate was 91.2 percent (13,739 reponding DUs out of 15,060) and information was available for 95.3 percent of the dwelling units (14,345) to determine eligibility for the SAIAN.

SAIAN Round Specific Field Results

Based on the overall results of the SAIAN screening interview, there were 2,030 dwelling units eligible for the SAIAN core Round One interview. These 2,030 dwelling units translated into 2101 reporting units for which SAIAN Round One interviews were attempted. The Round One field period spanned March to June, 1987. According to field reports, 2005 reporting units completed the interview, out of 2080 reporting units determined to be eligible for the Round One interview, yielding a 96.4 percent round specific response rate (Table 8). When the results of the screener interview are incorporated into the determination of the overall response rate, the joint screener-round one response rate was 87.9 percent (.9123 x .9639).

The resultant number of households completing the Round One interview (2,005) was approximately 10 percent lower then expected from original projections (2,217). This differential was most likely a function of departures from the

assumptions used to estimate the growth rate in the number of Indian dwelling units from the time of the 1980 Census, and the assumptions used to estimate the number of persons per occupied dwelling unit. Since the primary level of analysis for SAIAN is at the person level, and preliminary estimates indicated a larger number of persons per household than expected, the effect of the shortfall in households on the precision of survey estimates was minimal.

The set of 2,005 reporting units responding to the Round One interview constituted the initial Round Two sample. The Round Two field period ran from July, 1987 through February, 1988. In all, 2,212 reporting units were identified. This departure from the initial target was a consequence of the creation of new reporting units in the field. A new reporting unit was created when a key member of an original reporting unit moved away or "split" from the household. The movers were followed to obtain complete information on their 1987 medical experience. Westat field reports indicated there were 2,116 reporting units that completed the interview out of 2,192 determined to be eligible, yielding a 96.5 percent round specific response rate (Table 9). This translated to an overall response rate of 84.9 percent (.9123 x .9639 x .9653).

The Round Three field period ran from February through June, 1988, with a primary objective to obtain data on medical care utilization and expenditures for sampled individuals for the period between the Round Two interview and December 31, 1987. According to field reports, the final Round Three sample consisted of 2,201 reporting units. Field results indicated that 2,194 reporting units were determined to be eligible for the SAIAN, of which 2,152 completed the interview, yielding a round specific response rate of 98.1 percent (Table 10). When the results of all rounds of interviewing are considered, the overall response rate for the SAIAN at the reporting unit level was 83.3 percent (.9123 x .9639 x .9653 x .9809). It should be noted, however, that this response rate does not reflect the exclusion of 11 unique segments associated with the White Mountain Apache tribe that refused to participate in the survey.

Estimation and Analysis

The SAIAN was designed to yield national estimates of demographic and health care parameters for the 1987 civilian noninstitutionalized population of American Indians and Alaska Natives living on or near federal reservations and eligible for IHS care. Analyses are planned at the person level, at the family level, and at the health care event level (e.g. physician visits, hospital stays, use of prescribed medicines). The estimation strategy for the SAIAN will include adjustments for all levels of nonresponse. Nonresponse adjustments will be made at the housing unit level, the reporting unit level, and the person level for complete and partial nonresponse for a person's entire period of eligibility in 1987. To further improve the precision of survey estimates that characterize the SAIAN sample, a poststratification adjustment at the person level will be considered, based on IHS estimates of the SAIAN eligible population.

The SAIAN is characterized by a complex design that includes stratification, clustering, and disproportionate sampling. These departures from simple random sampling assumptions requires special consideration with respect to variance estimation and analysis. Several methods for approximating sampling variances, which incorporate the components of a complex survey design, have been developed and will be considered for application to NMES data (Cohen, Burt and Jones, 1986). The three most generally accepted and frequently used techniques are the method of balanced repeated replication (BRR), the "jackknife" method, and the Taylor series linearization method (Cox and Cohen, 1985). These variance estimation strategies have been incorporated as procedures in several of the widely used statistical package programs.

Summary

The complex survey design of the household component of the Survey of American Indians and Alaska Natives has been described in detail. The report includes a discussion of the departures from the original SAIAN design that were driven by cost constraints. Particular attention has been given to the disproportionate sampling strategy used to oversample the SAIAN eligible population in high density areas, in order to reduce survey costs due to screening.

Field results from the screener interview are also presented, in addition to the results for the core SAIAN household interviews, which spanned three rounds of data collection. As indicated, the SAIAN achieved an 83.3 percent overall response rate. The report includes a comparison of targeted sample size goals with those that were achieved for the survey. To conclude, appropriate estimation techniques and methods of variance estimation which consider SAIAN survey design complexities, have been suggested for use.

References

Bryant, E.C., H.O. Hartley, and R.J. Jessen (1960). "Design and Estimation in Two-Way Stratification". Journal of the American Statistical Association, 55, 105-124.

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

DiGaetano, R.G., and J. Waksberg (1986). Sample Design of the NMES Household Survey (HHS) and the Survey of American Indians and Alaska Natives (SAIAN). NMES Report No. 1.100. Westat, Inc., Rockville, Maryland.

DiGaetano, R.G. (1987). Sampling Report for the Selection of Dwelling Units for SAIAN. NMES Report No. 1.101. Westat, Inc., Rockville, Maryland.

To obtain the tables that are referenced in this paper, contact Dr. Steven B. Cohen, Senior Research Manager, NCHSR, Room 18A55, 5600 Fishers Lane, Rockville, MD 20857