THE DESIGN OF THE NATIONAL IMMUNIZATION SURVEY

Trena M. Ezzati-Rice, Centers for Disease Control and Prevention (CDC); Elizabeth R. Zell, CDC; Michael P. Battaglia, Abt Associates, Inc; Pamela L.Y.H. Ching, CDC; Robert A. Wright, CDC
Trena M. Ezzati-Rice, National Center for Health Statistics, CDC, 6525 Belcrest Road, Hyattsville, MD 20782

1. Introduction

This paper presents a comprehensive overview of the design features of the National Immunization Survey (NIS), formerly known as the State and Local Area Immunization Coverage and Health Survey. The National Immunization Program and the National Center for Health Statistics, Centers for Disease Control and Prevention (CDC), and its contractor, Abt Associates, Inc., implemented the NIS as one element of President Clinton's Childhood Immunization Initiative (CII). Details of the CII are given in a previous report by CDC (1994). After outlining the objectives of the CII, we discuss the sample design of the NIS; data collection procedures, including methods for maximizing response rates; procedures for assessing data quality; and key methodological issues.

One of the Healthy People 2000 objectives is to have at least 90% of 2-year-old children fully vaccinated with the recommended schedule of vaccines. The CII has been set up to 1) improve the delivery of vaccines to children; 2) reduce the cost of vaccines for parents; 3) enhance awareness, partnerships, and community participation; 4) monitor vaccination coverage and occurrences of disease; and 5) improve vaccinations and their use. Specific objectives of the CII include a 1996 goal of 90% coverage for four of five critical vaccines — diphtheria and tetanus toxoids and pertussis (DTP), poliovirus (polio), measles-mumps-rubella (MMR), and *Haemophilus Influenzae* type B (Hib), and 70% coverage for hepatitis B (Hep B).

In order to measure progress toward achieving these goals and to assist the CDC in directing resources, the NIS was initiated in April 1994 to monitor vaccination coverage levels on an ongoing basis. The target population is the noninstitutionalized population 19 to 35 months of age. The NIS covers 78 Immunization Action Plan (IAP) areas which include the 50 states, the District of Columbia, and 27 metropolitan areas. The NIS has been designed to provide estimates of vaccination coverage levels for the 78 individual IAP areas through the use of a random-digit-dialing (RDD) telephone survey methodology. Although approximately 5% of households in the United States do not have a telephone, for households with 2-year-old children the nontelephone rate is 12% ranging from 2% to 25% among the various IAP areas.

The NIS includes independent quarterly surveys in each of the 78 IAP areas. With this design, any four consecutive quarters of data can be combined to produce annual estimates of vaccination coverage. The NIS data collection methodology and sample design use a list-assisted RDD technique to provide the sample of telephone numbers in each IAP area. The screening for eligible households and the immunization interview are conducted by computer-assisted telephone interviewing (CATI) The NIS seeks to attain a coefficient of variation no larger than 5% for the annual vaccination coverage estimates in each IAP area This precision requires a simple random sample of 400 (100 per quarter) age-eligible children per IAP area. However, to compensate for factors such as households with multiple telephone numbers, households with more than one age-eligible child, household telephone numbers assigned to an IAP area for sampling purposes but actually located in an adjacent IAP area (crossover interviews), and differential weights from nonresponse adjustments, the sample size target for the NIS is set at 440 per IAP area for each four-quarter period or 110 per quarter.

2. Sample Design

The NIS uses a two-phase sample design. For the first-phase a quarterly random sample of telephone numbers for each IAP area is drawn, and a screening questionnairre is administered to locate households with one or more children 19 to 35 months of age. Vaccination information is then collected for all age-eligible children. In order to efficiently screen the large number of households required to obtain a sample of age-eligible children, the first-phase sample design uses list-assisted RDD sampling methodology (Lepkowski, 1988). This RDD technique uses information on the distribution of residential directory-listed telephone numbers across banks of 100 consecutive telephone numbers. Banks with no residential directory-listed telephone numbers are removed from the sampling frame. A simple random sample of ten-digit telephone numbers is then drawn from the banks containing one or more residential

directory-listed numbers. The resulting sample includes households with listed and unlisted telephone numbers and has a high percentage of telephone numbers that are residential, as opposed to business or nonworking.

The selected sample is matched against a commercial data file of residential directory-listed telephone numbers to identify those sample numbers that will be mailed an advance letter explaining the survey. The sample numbers that are not residential directory-listed numbers are matched against a commercial data file of business telephone numbers in an effort to remove a portion of the business numbers before the sample is dialed by interviewers. The sample numbers that are neither residential directory-listed numbers or business numbers are sent to an auto-dialer that is designed to detect signals from nonworking-number recordings, modems, and facsimile machines. this step in the sample design removes a portion of the nonworking numbers from the sample. This pre-screening of the sample removes 34.1% of the nonworking and business numbers from 'the initial sample, thereby raising the percentage of working residential numbers from an average of 52% to 61%. As reported by Battaglia *et al.,* only about 1.4% of the residential numbers in the sample were misidentified as nonworking or business numbers.

Previous studies by Goldstein *et al.* (1993) and Valadez and Weld (1992) indicate that parents tend to underestimate the number of doses received for multiple-dose vaccines and to overestimate coverage for single-dose vaccines. In addition, results from the National Immunization Provider Record Check Study (NIPRCS), conducted by CDC in conjunction with the 1994 National Health Interview Survey (NHIS), indicate that household reports of vaccination often contain errors and that a combiion of parental reports and information from health care providers' records gives a more accurate estimate of vaccination coverage levels than either source alone (CDC, unpublished data, 1995). Examples of reporting error include memory recall errors and the use of a shot card that fails to show all the vaccinations that the child received.

Therefore, in the 1994 NIS, a second-phase sample of health care providers of vaccinations of children in surveyed households was included to supply the data needed to improve the accuracy of the vaccination coverage estimates. Because the provider component was incorporated into the NIS late in the fourth quarter of 1994, respondents who had already completed an RDD immunization interview were recontacted to obtain information on the name and address of the health care providers of their children's vaccinations and verbal consent to contact the named providers. After implementation of the provider component into the NIS, provider names, addresses, and consent were obtained concurrently with the CATI immunization interview. Because of time and resource constraints, children with household reported vaccinations from a shot card which documented receipt of 4 DTP, 3 polio, 1 MMR, and 3 Hib were excluded from recontact Results from the NIPRCS show such shot card reports to be highly accurate (CDC, unpublished data, 1995). A questionnaire mailed to providers collects vaccination information for the identified children. When compared with the household telephone reports of vaccinations, the provider data form the basis for a set of ratio adjustment factors that are applied to the vaccination reports from the first-phase sample. Zell *et al.* (1995) give further details of the provider component of the NIS. Standard two-phase sample estimation procedures are used to form the estimates of vaccination coverage for each IAP area (Cochran, 1977). Procedures to adjust for potential coverage bias resulting from noncoverage of nontelephone households in the NIS are discussed by Battaglia et al. (1995).

The NIS target sample size is 8,580 completed interviews per quarter for all 78 IAPs (110 per IAP area); therefore, the target for the three quarters of data collection in 1994 was 25,740 completed interviews. For 1994, an average of 512,800 sample telephone numbers were drawn per quarter for a total of 1,538,400 sample telephone numbers for the three quarters. On average, 420,500 sample telephone numbers were actually dialed per quarter by interviewers after the prescreening for business and non-working numbers, thus removing 18% of the initial sample. About 40% were residential directory-listed telephone numbers, and hence, eligible to receive an advance letter explaining the purpose of the NIS and encouraging the cooperation of the household. To improve the accuracy of the estimates for 1994, the second-phase sample of providers had. a target sample size of 50 children with complete provider data per IAP area (excluding children with a household report from a shot card which documented receipt of 4 DTP, 3 polio, 1 MMR, and 3 Hib because of time and resource constraints and the retrospective nature of data collection for the provider component). For future quarters of data collection the target sample size will be about 80 children per IAP area with provider vaccination reports.

3. Data Collection Procedures

Given the size of the NIS, the data collection process is automated as much as possible. The interviewing via CATI is conducted from a single telephone center. Because the sample telephone numbers cover six time zones, the telephone center is open from early morning to past midnight seven days per week During a quarter, simultaneous interviewing takes place in the 78 IAP areas, The sample for each IAP area is divided into replicates (subsamples) in order to control the amount of sample released throughout the quarter toward the goal of obtaining approximately 110 interviews from each IAP area.

The CATI interview has five main components. First, the screening questions determine whether the telephone number reached is a residence and whether the household has any children 12 months to 3 years. The birth date and gender of children in that age range are obtained, and the CATI system then determines whether any of the children are 19 to 35 months of age. The final section of the screening instrument identifies the respondent most knowledgeable about the child's vaccination status. If that person is not available, a time to call back is arranged.

Second, in the immunization interview, the immunization questions have been adapted from the in-person NHIS Immunization Supplement. The immunization interview prompts the respondent to locate the child's shot card If a shot card is available, the interviewer asks the respondent to report the number and dates on which the child received vaccinations for DTP, polio, MMR, Hib, and Hep B. The respondent is also asked to report any additional shots received by the child but not recorded on the shot card. If a shot card is not available, the interviewer asks the respondent to recall from memory the number of shots, but not the dates, for each of the vaccinations mentioned previously. The immunization instrument also allows for an unsolicited response of "all" as the number of shots received for a particular vaccine.

Upon completion of the immunization questions, the interviewer then proceeds to the third component, a section that collects demographics, current residence and residence at birth, and information on the number of phone lines in the household. Fourth, in the provider information and consent section, the interviewer asks the respondent to identity the health care providers who have provided vaccinations for the sample child(ren) and to give their addresses and telephone numbers. Verbal consent to contact the named provider(s) is also requested from the respondent. A packet of materials is mailed to the named providers, requesting the vaccination history of the child(ren)

Fifth, the nonresponse interview, an abbreviated version of the screening questionnaire, is used when the respondent refuses the screening or immunization interview, when the most knowledgeable person is not available, and when a call-back must be arranged because the person needs to locate the child's shot card. The questions include whether any children 18 months to 3 years of age live in the household, the number of doses of DTP received by the oldest child in that age range, and the race/ethnicity and date of birth of that child This information is used in the estimation process to make adjustments for unit nonresponse.

4. Response Rate Goals and Methods to Maximize Survey Participation

The NIS has several goals for components of the overall household telephone response rate. First, the percentage of sample numbers with an unkown residential status should be less than 2% at the end of a quarter. Second, for each IAP area, the screening interview completion rate for households contacted should be 90% and the immunization interview completion rate among eligible households should be 85%. The overall target response rate is 77%. In addition, the NIS has an expected residential working number rate of 70% and an expected household eligibility rate of 5%. With data collection in 78 separate IAP areas, it is important that a comprehensive monitoring system be developed and used. An automated monitoring process for the NIS determines eligibility rates, working number rates, telephone number resolution rates, and response rates by IAP area on an ongoing basis throughout the quarter.

The response rate goals for the NIS require special procedures to enhance participation in the survey. First, as noted earlier, the sample telephone numbers are matched against a commercial data file of residential directory-listed numbers, and an advance letter is mailed to about 40% of the telephone numbers called by inerviewers. The advance letter explains the purpose of the study, identifies the federal agency and organization conducting the survey, encourages households to participate in the survey, and requests respondents to locate the vaccination records of any young children in the household. Mailings are coordinated so that the advance letter arrives shortly before the first call attempt is made to the sample telephone number. In an experiment embedded in the NIS sample, Camburn et al. report that the advance letter increased the overall survey response rate by about 7 percentage points.

Second, in the NIS, for 20% of the sample telephone numbers, at least one attempt reaches an

answering machine.   After reaching an answering machine for the third time, a brief message is left describing the NIS, and a toll-free 800 number is provided for the person to call with questions about the survey and/or to schedule an interview. The message left by the interviewer also asks that, if someone is present in the household, they pick up the telephone. The message has been customized to deliver the same information provided in the advance letter. The primary objective is to differentiate the NIS from telemarketing calls.

Thiid, the NIS uses several approaches to accommodate the nearly 3 million linguistically isolated households in the U.S. The questionnaire has been translated into Spanish, and Spanish speaking interviewers and supervisors are available during all work shifts. The NIS also uses real-time language translation services. The AT&T Language Line provides an ability to conduct the interview in more than 140 languages, 24 hours per day.

Fourth, the NIS identifies sample numbers that have an unresolved residential status after a preset number of attempts.   These numbers. are matched against an AT&T data file to identity the local telephone company; Specially trained interviewers make calls to local telephone companies in an effort to determine whether the number is residential, business, or not assigned. Typically, 70% of the numbers that result in a "ring - no answer" to all attempts are found to be a business or not assigned, and these do not receive further dialing attempts.

5. Assessment of Data Quality

Assessing the quality of the data from the NIS is a critical and ongoing aspect of the survey. Because of the high level of screening required, monitoring of the household eligibility and response rates is essential for maintaining quality and for minimiig costs. The response rates used for monitoring the NIS are described by Massey (1995).   To assess potential nonresponse bias resulting from nonparticipation of health care providers, characteristics of children with both household and provider data are compared with those with only household data. To assess the validity of estimates from the NIS, findings are compared with vaccination coverage data from the NHIS, a national household survey of the U.S. civilian, noninstitutionalized population, supplemented with provider information in the same manner as in the NIS survey. In addition, comparisons are made with estimates   of vaccination   coverage derived independently by selected states.   Finally, the

vaccination coverage levels at the national and state level are monitored on a quarterly basis.

6. Conclusions

The NIS is one of the largest rnadom-digit-dialing surveys conducted in the U.S. The provider component will also collect tens of thousands of vaccination reports from providers. As a result, the NIS faces several methodological issues that relate to both its large size and sampling and data collection procedures. Key issues include assessing and adjusting for potential nontelephone coverage bias and unit nonresponse bias, as well as for response error in household vaccination reports, and developing procedures to produce accurate coverage estimates for the nation and for each state and metropolitan area. In addition, maximizing the use of a shot card during the interview is important because information from a shot card tends to be more accurate than reporting from memory. Further, obtaining verbal consent to contact providers presents several logistical and legal issues.

In summary, for the first time, a standard methodology is available for producing current, population-based state and national rates of vaccination coverage for all children 19 to 35 months to monitor the progress toward national vaccination goals. The NIS provides the most reliable estimates of vaccination coverage through the use of household vaccination information and data from the records of health care providers. However, as with any sample survey, potential limitations exist. For the NIS, these include possible biases associated with the exclusion of households without telephones, household and provider nonresponse, and inaccurate reporting by household respondents. An adjustment for the exclusion of households without telephones applies findings in the 1992 and 1993 NHIS that-showed coverage levels for 4 DTP, 3 polio, and 1 MMR of about 18 percentage points lower among children in households without telephones than among children in telephone households (CDC, unpublished data, 1995). An adjustment for household nonresponse is also made. To improve the accuracy of the estimates and reduce potential response errors in the reports of vaccinations by household respondents, theNIS includes a second-phase sample of providers.   Although provider information was not available for all children in the 1994 NIS, the demographic characteristics and immunization status of those children whose providers were not included in the survey were similar to children whose providers were included, suggesting that the use of the provider data in producing coverage estimates will not introduce bias into the estimates. Finally,

because the 1994 NIS collected only three quarters of data, small sample sizes for individual states result in a relatively large variance for the vaccination coverage estimates. However, future analyses will include data for four quarters, thereby reducing the size of the sampling error.

References

1. CDC. (1994). "Reported vaccine-preventable diseases-United States, 1993, and the Childhood Immunization Initiative." *Morbidity and Mortality Weekly Report 43:57-60.*
2. Lepkowski, J.M. (1988). "Telephone sampling methods in the United States." In: *Telephone Survey Methodology (eds. R Groves et al.), 73-98. New York:* John Wiley & Sons, Inc.
3. Battaglia, UP., Starer, A., Oberkofler, J., Zell, E. " Pre-identification of nonworking and business telephone numbers in list-assisted random digit dialing samples." Presentation at the 1995 Annual Meeting of the American Association for Public Opinion Research, Tampa, FL.
4. Goldstein, K.P., Kviz, F.J., Daum, RS. (1993). " Accuracy of immunization histories provided by adults accompanying preschool children to a pediatric emergency department."*Journal of the American Medical Association* 270:2190-2194
5. Valadez, J.J., Weld, L.H. (1992). "Maternal recall error of child vaccination status in a developing nation." *American Journal of Public Health 82:120-122.*
6. Zell, E.R, Ezzati-Rice, T.M., Hoaglin, D.C., Massey, J.T. (1995). " Adjusting for respondent bias on vaccination status in a telephone survey." 1995 *Proceedings of the Section on Survey Research Methods* Alexandria, VA.: American Statistical Association.
7. Cochran W.G. (1977). *Sampling Techniques.* 3rd ed., New York: John Wiley & Sons, Inc.
8. Battaglia, M.P., Malec, D., Spencer, B., Hoaglin, D.C., Sedransk, J. (1995). "Adjusting for noncoverage of nontelephone households in the National Immunization Survey." 1995 *Proceedings of the Section on Survey Research Methods* Alexandria, VA.: American Statistical Association.
9. Camburn, D.P., Lavarkas, P.J., Battaglia, M.P., Massey, J.T., Wright, RA. "Using advance respondent letters in random-digit-dialing telephone surveys." Presentation at the 1995 Annual Meeting of the American Association for Public Opinion Research, Tampa, FL.
10. Massey, J.T. " Estimating the response rate in a two-stage telephone survey." (1995). 1995 *Proceedings of the Section on Survey Research methods.* Alexandria, VA.: American Statistical Association.