

SPLIT SAMPLING DESIGN FOR TOPICAL MODULES IN THE NATIONAL IMMUNIZATION SURVEY

Philip J. Smith, CDC/NIP; Diane Simpson, CDC/NIP; Michael P. Battaglia, Abt Associates Inc.;
Vicki Huggins, Abt Associates Inc.; Lance Rodewald, CDC/NIP; Monina Klevens, CDC/NIV/AIDS;
Marilyn Wilkinson, Abt Associates Inc.; Ann-Sofi Rodén, Abt Associates Inc.; Meena Khare, CDC/NCHS;
Robert A. Wright, CDC/NCHS

Philip J. Smith, CDC/NIP, Corporate Square Blvd., Bldg. 12, Room 2109, Atlanta, GA 30333

1. Introduction

The National Immunization Survey (NIS) was initiated in April 1994 to monitor vaccination coverage rates on an ongoing basis. The NIS covers 78 Immunization Action Plan (IAP) areas that include the 50 states and 28 metropolitan areas, including the District of Columbia. Each IAP area represents a stratum of the sampling design within which the NIS samples independently. Within each IAP area, the design of the NIS includes 2 phases of sampling: a random digit dialing (RDD) phase of households with children between the ages of 19 and 35 months of age, followed by a mail survey of vaccination providers of these "age-eligible" children in sampled households. Smith et al.^{1,2} and Zell et al.³ give a more detailed description of the design of the NIS. One of the primary functions of the NIS has been to provide surveillance to monitor progress toward achieving vaccination goals defined by the Childhood Immunization Initiative⁴ and to assist the CDC in directing resources to states for the purposes of improving vaccination coverage rates in IAP areas.

In addition to providing estimated coverage rates for IAP areas, there has been recognition that additional information is required to assist the National Immunization Program at the Centers for Disease Control and Prevention in improving vaccination coverage rates. Specific topics that have been identified where additional information is required include: (i) parents' ability to pay for vaccinations; (ii) parental knowledge and experiences about immunization; and (iii) day care arrangements/breastfeeding practices/WIC participation. All of these are important topics in which data will be collected and analyzed to improve our understanding of vaccination in the U.S. This information could lead to new or improved programs that are designed to increase vaccination rates further.

However, the need to collect information on these topics should be carefully balanced with the burden that respondents bear in participating in the NIS RDD interview. Approximately 85% of all NIS respondents in the RDD phase of the survey are known to be mothers with at least 1 child between 19 and 35 months of age. Also, the current average interview length for the NIS household interview conducted over the

telephone is approximately 24 minutes for households with an age eligible child. While the need for information on additional topics is great, if all respondents were asked additional questions, the length of the NIS interview would become considerably longer and would place even greater burden upon respondents.

To control respondent burden, a split sampling design will be used. In this design, each respondent with an age eligible child will be randomly assigned to receive questions pertaining to only one of the three topics. We refer to the questions that pertain to a single topic as a "topical module." The overall goal of the split sampling design is to control interview time with respondents, yet to enhance the NIS interview to collect additional analytically important information.

To make room for the new topical module questions and ensure that the overall household burden does not increase unduly in the NIS, several questions administered to all RDD sampling phase respondents will be dropped. Specifically, detailed questions on participation in the Women, Infants and Children Program (WIC) and age of child during WIC participation will be no longer asked of all respondents. Also, several NIS questions that ask about the parent/guardian's perception of whether their child is up-to-date on recommended vaccinations and who took the child for most of his/her vaccinations will be dropped.

The three proposed modules are scheduled to be introduced into the NIS in 2001 and continue to be asked for 4 consecutive quarters. These modules may be repeated or may be replaced with other modules as new topics are proposed for study. The rapid availability of quarterly data on a national level is a feature that makes the NIS unique among national health surveys. Adding topical modules into the NIS will make information on current events rapidly available and increase the pool of timely data from which knowledge may be obtained about behavioral, social, demographic and economic correlates of vaccination practices. Topical modules are used routinely in the National Health Interview Survey (NHIS). This paper describes:

- the primary analytic purpose of each topical module;
- a discussion of methodological issues for each module; and
- the statistical design that will be used for each module to insure that the statistics obtained from each module will be suitably precise for each module's primary analytic purpose.

2. Analytic Objectives of the Topical Modules

Ability to Pay for Vaccinations. Costs to parents and providers are a known barrier to vaccination. Specifically, administration of vaccines by private providers depends on insurance reimbursement rates, the availability of publicly purchased vaccines, and other possible costs to providers. Referral of children needing vaccinations away from a child's "medical home" to a health department clinic also depends on insurance coverage and out-of-pocket costs to parents. These referrals cause missed opportunities and delays in timely vaccinations. Data from this topical module will provide information on these economic barriers to vaccination and their impact on vaccination coverage levels.

The Vaccines for Children (VFC) program is a federal entitlement program that provides publicly purchased vaccine for five groups of children: uninsured children; underinsured children if they go to a federally qualified health center; Medicaid-enrolled children; and American Indian or Alaska Native children. Nationally, approximately 32% of preschool children are eligible for VFC. In order to correctly distribute the VFC funds to the states, CDC needs to know the proportion of a state's preschool children that are entitled to VFC vaccine. This module will provide state-by-state estimates of this proportion, and these estimates will be combined with state program information to distribute funds.

Although VFC is a federal entitlement program, it is operated at the state level. Each state enrolls providers into the VFC program, and these providers administer VFC vaccine to eligible children. Since the NIS provider record check asks whether the provider is enrolled in VFC, this module will allow each state to determine the proportion of VFC-eligible children who receive their vaccines from VFC enrolled providers. This information will help guide states' provider enrollment efforts.

States purchase vaccine using one of three different policies: universal purchase (15 states), in which all vaccine is purchased by the state; augmented VFC (16

states), in which state funding is used to purchase vaccine for children whose commercial insurance does not cover the cost of vaccination so they can be vaccinated in their medical home; and standard purchase (17 states), in which vaccine for underinsured children is purchased only for health departments. This module will be used to determine whether the likelihood of referral to health department clinics is associated with vaccine purchase policy.

Parental Knowledge and Experiences. Factors that are believed to influence a child's vaccination up-to-date status include: (i) parental perception of vaccine safety; (ii) perception of vaccine efficacy; (iii) awareness of the recommended vaccination schedule; and (iv) perception of the need for vaccines, i.e., the severity and probability of exposure to vaccine-preventable diseases. Also, it is believed that there are important influences (such as doctors, vaccination campaigns directed toward families) on parental decisions to seek or avoid vaccines. Understanding these influences is important in developing new initiatives and in improving current vaccination programs.

The purposes of the survey questions in this topical module are to assess the parent/guardian's perceptions about vaccine safety and influences to seek vaccines. The relationship between these perceptions and a child's timely receipt of vaccines also will be evaluated. A better understanding of these factors will enable the National Immunization Program to address concerns and improve education with the overall goal to increase timely vaccination coverage among young children in the U.S.

Day Care/Breast Feeding/WIC Topical Module. Young children in day care facilities are at greater risk of disease because of the increased chance of exposure to vaccine preventable diseases. Because of this, some states require that children attending licensed day care facilities receive vaccinations. Parents of young children in day care may be informed by their day care facility about the importance of immunization, even if vaccines are not mandatory for the child's enrollment.

Among the objectives of the Year 2010 Healthy People goals⁵ is to increase the up-to-date vaccination rates among children attending childcare. Questions in this module will allow NIP to evaluate the extent that children attend day care and to analyze vaccination coverage as it relates to the child's attendance.

Breastfeeding is believed to provide children with antibodies and other factors that help protect against invasive forms of some diseases such as Haemophilus influenzae b. However, some parent may think breast

feeding lessens the need for timely vaccinations. Knowledge of the prevalence of breastfeeding in the U.S. and whether it is linked to vaccination status is important for maternal and child health programs to plan and improve campaigns to increase this healthy practice while maintaining high age-appropriate vaccine coverage.

The National Immunization Program routinely assesses whether the USDA's Women, Infants, Children (WIC) program encourages mothers to adequately vaccinate their children. The WIC program was established to provide education and access to nutritious diets to low income children and pregnant and lactating women at high risk for inadequate diets. This program serves about five and a half million infants and children. Almost half of the babies born in the U.S qualify for WIC. In 1997, WIC began an initiative to promote breast-feeding. From 1989 to 1995, the rates of mothers breastfeeding newborns in the hospital rose from 52% to 60% among all sociodemographic groups and from 34% to 47% among WIC recipients. The percentage of mothers still nursing their babies at 6 months of age increased from 18% to 22% and from 8% to 13% among WIC mothers. Despite the effort and relative success by the WIC program to increase breastfeeding, WIC supplies formula to mothers and may be creating an unintentional disincentive to breastfeeding. Analysis of data on day care, breastfeeding and WIC participation from the same sample of children allows the study of potentially complex relationships between up-to-date vaccination status and (i) child care attendance, (ii) WIC participation, and (iii) breastfeeding practices.

3. Methodological Issues

Independent Module Samples. The use of three topical modules on separate national subsamples imposes some limitations on the analysis of the data. It will be possible to examine the relationship between the questions within a module and demographic and socioeconomic characteristics collected in the NIS. More importantly, it will also be possible to study the relationship between questions in a module and vaccination up-to-date status for children with adequate provider data. It will, however, not be possible to examine the relationship between questions in different topical modules. This is not viewed as a major limitation, because the topical modules were intended to be self-contained. If there are specific questions that are required for analytic purposes in more than one module, those questions can be included in multiple modules.

Measurement Error Issues. Measurement errors can arise from several sources – the respondent, the interviewer, and the questionnaire, or response effects by mode of data collection.⁶ Several steps have been taken to address these issues. First, an extensive literature review of questionnaires covering the same topics was carried out. Second, the proposed questions were reviewed by nationally recognized experts in survey questionnaire design and construction. Third, cognitive testing by the Questionnaire Design Research Lab (QDRL) at the NCHS will be used to refine the questions so that they will be understandable and interpreted in the manner in which they were intended. Fourth, refined questions obtained from the NCHS QDRL will be pretested in an RDD field test of approximately 600 households to evaluate the questions' appropriateness and acceptability in the NIS interview setting. Also, in the field test, we plan to study the impact that the new topical modules have on break-off rates, item nonresponse rates, and the rate at which interviewed subjects provide consent to allow NIS study personnel to obtain vaccination histories from vaccination providers. Finally, for the purposes of providing assurances for the protections of human subjects, the Institutional Review Board (IRB) at the National Center for Health Statistics (NCHS) will have reviewed and approved the protocols and questions used in all pretests and final modules.

Specific measurement issues are being addressed in the design of each module. The Day Care Arrangements/Breastfeeding Practices/WIC Participation module begins with a series of questions to determine whether the selected child age 19-35 months ever participated in a child care arrangement including the situation where the child attended home daycare. Because vaccinations for young children are recommended primarily during the first and second years of life, the module goes on to ask about the specific child arrangements for each of these two time periods. The breast feeding questions ask whether the child was ever breast fed or fed breast milk. Given that the children are only 19-35 months old and that 85% of all the NIS respondents are the child's mother, we assume that the recall of events from birth will be accurate. This is an issue that we intend to investigate during the cognitive testing of the questions. Also, we want to ensure that respondents understand and interpret the terms for different types of day care consistently.

The Parental Knowledge and Experiences module begins with a question on how the parent knew when to take their child for his/her most recent immunization. The module also asks the parent's views on vaccine safety and who influences their decisions regarding

childhood vaccines. The goal of questions in this section is to assess attitudes and influences, yet not ask questions in a leading or alarming manner. During the cognitive testing of these questions, we will develop a better understanding of how respondents will interpret and react to the questions.

The Health Insurance Coverage and Access to Immunization module begins by determining whether the child has a “medical home.” The technical definition of a medical home uses terminology that may not be familiar to most respondents. The actual question therefore uses a description that respondents can understand regardless of the type of facility where the child receives pediatric care. The health insurance questions are based on questions developed by NCHS for use on other telephone surveys and make use of research on the accuracy of Medicaid reporting.⁷ Blumberg and Cynamon, as well as others who have assessed the accuracy of reporting health insurance coverage in national level surveys such as the NHIS, the Survey of Income and Program Participation, and the Current Population Survey⁸, identified issues of reporting error that need to be considered in the analysis of health insurance data reported in surveys. Specifically, comparisons to administrative data on health insurance coverage generally indicate that surveys may underreport health insurance coverage

The careful design of the questions in each module is intended to maximize the analytical usefulness of the information collected and to substantially reduce measurement errors. Cognitive testing and field testing are expected to shed light on major question wording and or questionnaire administration problems. To reduce the potential of response bias, response categories and a series of questions will be randomized so that any ordering effects are mitigated.

Survey Methods. Even though the topical modules will be questions “added” to the current NIS questionnaire, other changes may be required, and potential new methods and procedures added to ensure that the core NIS data are not adversely affected. We may decide to test several alternatives during the field test. Below are survey methods or procedures we envision will be reviewed, revised and or tested during the field test.

Advance letters mailed to households for which we have addresses will need to be examined and possibly updated to reflect the new subject matter that will be added to the NIS and the potential change in interview length for subsamples. There is evidence from previous tests on the use of advance letters in the NIS that a well-designed advance letter can improve response rates.⁹ We want to ensure that any change made to the advance

letter does not adversely affect the core NIS CASRO household response rate¹⁰ or other important survey statistics such as the vaccine coverage rate, the percent of cases that give consent to contact health care providers and the number of calls required to complete a case.

Similarly, the RDD household introduction/screener questions that ask about the existence of eligible children in the household may require minor modifications with the addition of topical modules. Experience we have gained in the NIS indicates that the survey introduction and screener length and content are important in gaining respondent cooperation. We will monitor this very closely during the field test and first quarter of full-scale implementation of the topical modules.

From a respondent point of view we would want to position the topical module questions in the NIS computer assisted telephone interview (CATI) questionnaire to make the interview flow in as pleasing and easily understandable manner as possible. Whether the new topical module questions affect the consent rate for contacting health care providers is an important issue since the vaccination coverage rates depend on obtaining adequate vaccination data from providers. Field testing of alternative positioning of the topical module questions or accepting higher levels of topical module unit nonresponse may be necessary. Other alternatives include making the CATI interview more flexible. For example, using a “hot key” to “jump” to the set of questions in which consent to contact the child’s vaccination provider(s) will be considered sooner in the survey.

In the Attitudes and Barriers topical module, it was decided that the household respondent would only be asked the set of questions once with references to the youngest age eligible child, even if multiple children were eligible in the household. This decision was made assuming that recall for the youngest child would be the most accurate and it also provides a more current read on attitudes and perceptions.

4. The Statistical Design for the Topical Modules

The design of the NIS includes 2 phases of sampling: a list-assisted RDD sampling phase of households followed by a mail survey of vaccination providers of eligible children in sampled households. The sampling plan of the RDD phase of the NIS is a stratified random sampling of 78 IAP areas, which comprise the 50 states and 28 large metropolitan areas including the District of Columbia. Each IAP area represents a stratum of the sampling design within which the NIS samples

independently. The target population for the NIS is all children living in the 78 IAP areas who are between 19 and 35 months of age at the time of the RDD screening interview. Each year, approximately 35,600 completed RDD interviews are obtained nationally in the NIS.

The split-sampling design will be implemented for 4 consecutive quarters and will proceed by randomizing each of the 35,600 completed NIS RDD interviews to receive only 1 of three topical modules: (1) the “Ability to Pay for Vaccinations” topical module, or (2) the “Parental Knowledge and Experiences” topical module, or (3) the “Day Care Arrangements / Breastfeeding Practices / WIC” topical module.

To obtain suitably precise estimates that satisfy the primary analytic objectives of each topical module, approximately 74.6% of the NIS sample respondents will receive the “Ability to Pay for Vaccinations” module, 13.2% will receive the “Parental Knowledge and Experiences” module and 13.2% will receive the “Day Care/Breastfeeding/WIC” module, fully utilizing the entire NIS sample. Each module will use an equal allocation design to the 78 IAP areas. The following subsections provides further details about how the statistical design of the split sampling plan achieves suitably precise estimates to satisfy the primary analytic goals of each module.

The Day Care/Breast Feeding Practices/WIC Module.

The estimated sample size for this module is based on the objective of producing suitably accurate national level coverage estimates for population subdomains that comprise at least 50% of the entire population. These subdomains include children enrolled in day care (54%), children who were breast fed (60%), and children who ever participated in the WIC program (51%). For these purposes, “suitably accurate” estimates are defined as being within 2.5 percentage points of the true but unknown national estimate with a probability of 0.95 when the true but unknown vaccination coverage rate is 80% within a subdomain that comprises 50% of the target population.

Among the 35,600 expected number of completed interviews obtained over the four quarter period in which the split sampling design will be implemented, approximately 4,700 (13.2%) completed household interviews will be randomly assigned to the “Day Care Arrangements / Breastfeeding Practices / WIC” topical module. Adequate provider data are generally obtained for approximately 67-69% of children who complete the NIS RDD survey each year. The number of children with adequate provider data for the Day Care/Breast Feeding Practices/WIC topical module is therefore expected to be around 3,245 over the four quarter

period. Assuming that the subdomain of interest comprises 50% of the target population, we expect to obtain information from this module along with provider reported vaccination histories from approximately 1,622 children who belong to the subdomain of interest. Using these specification and the average expected design effect of 1.52 for national level estimated vaccination rates in the NIS, the effective sample size¹¹ will be approximately 1,067 children in the subdomain of interest. Table 1 shows that with these design specifications, the half width of the 95% confidence interval will be approximately 2.5%.

The Parental Knowledge and Experiences Module.

Among the 35,600 expected number of completed interviews obtained over the four quarter period in which the split sampling design will be implemented, approximately 4,700 (13.2%) completed household interviews will be randomly assigned to the “Parental Knowledge and Experiences” topical module. Subdomains for this module are expected to range from 5% to 50% of sample households. Table 1 shows the expected 95% confidence interval half-widths for estimated national vaccination rates for different sized subdomains that will be studied in this module.

The Ability To Pay Module. The estimated sample size for this module is based on the assumption that accurate state level statistics are important. The 26,200 completed interviews for this module represent 74.6% of the 35,600 expected completed interviews annually and will be used to produce accurate state level estimates. State sample sizes will range from 335 to 1,675 completed interviews depending upon the number of IAP areas within a state (Table 2). A key estimate from the “Ability to Pay” module will be the percentage of children that are eligible for the Vaccine for Children (VFC) Program. If we assume 32% of the children 19-35 months of age in a state are eligible for the VFC program, then we expect confidence interval half-widths to be in the range of 3.3 to 5.4 percentage points, depending on the number of IAP areas in the state (Table 2). If we look further at the confidence interval half-width for the percentage of VFC eligible children who are up-to-date on immunizations and assume 80% are up-to-date, we expect a range of 6.6 to 10.4 percentage points for the 95% confidence interval half-width, depending on the number of IAP areas in the state (Table 2).

Smith et al.¹² provide a description of how sampling weights are developed for children sampled in the RDD portion of the NIS survey and have sufficient provider-verified vaccination history information to evaluate their vaccination status. A separate sampling weight will be determined for children sampled in a topical

module. These sampling weights will account for (i) selection into and completion of the RDD portion of the survey, (ii) the probability of being randomized to the topical module, (iii) nonresponse adjustment attributable to a break-off during the topical module, (iv) poststratification to known totals available from the NCHS Natality file, and (v) adjustment for provider-verified vaccination history nonresponse.

5. Conclusion

The introduction of the three topical modules described in the paper is expected to be the beginning of a long-term approach to enhance the programmatic value of the NIS and utilize the rapid turnaround capabilities of the NIS to address current event questions related to vaccination and the health of young children. We will be exploring avenues to enlist the guidance and ideas of other health care researchers to continue to maximize the utility of the NIS as a tool for NIP to monitor vaccination and develop new and improved programs for meeting the Healthy People 2010 goals.

Table 1: 95% Confidence Interval Half-Widths for Topical Modules Designed to Precise Yield National Level Coverage Estimates within a Specified Subdomain, Assuming a National Design Effect of 1.52. **REFERENCES:**

Subdomain Size	95% Confidence Interval Half-Width
5%	7.6%
10%	5.4%
20%	3.8%
30%	3.1%
40%	2.7%
50%	2.5%

Table 2: State Sample Sizes and 95% Confidence Interval Half-Widths for “Ability to Pay” Module

Number of IAP areas in state	Number of completed interviews over four quarters	95% confidence interval half-width for percentage of children eligible for VFC program	95% confidence interval half-width for 4:3:1 up-to-date rate for VFC eligible children
1	335	5.4%	10.4%
2	670	4.5%	8.5%
3	1,005	4.2%	8.0%
4	1,340	3.3%	6.7%
5	1,675	3.5%	6.6%

References

- Smith, P.J., Battaglia, M.P., Huggins, V., Ezzati-Rice, T., Rodén, A.-S., Khare, M., Wright, R.A. Overview of the sampling design and The statistical methods used in The national immunization survey. *American Journal of Preventive Medicine*, to appear.
- Smith, P.J., Rao, J.N.K., Daniels, D., Battaglia, M.P., Ezzati-Rice, T.M., Khare, M. Compensating for nonresponse bias in the national immunization survey using response propensities. NCHS Series 2 Report. Hyattsville, MD. To appear.
- Zell ER, Ezzati-Rice TM, Battaglia MP, and Wright RA, "National Immunization Survey: The Methodology of a Vaccination Surveillance System", *Public Health Reports*, Oxford University Press, Cary, NC, vol 115, pp 65-77, January/February 2000.
- Centers for Disease Control and Prevention. Reported vaccine-preventable diseases -United States, 1993, and the Childhood Immunization Initiative. *MMWR Morb Mortal Wkly Rep*. 1994;43:57-60.
- U.S. Department of Health and Human Services. *Healthy People 2010* (conference edition 2 volumes). Washington, D.C. 2000.
- Groves, R.M., *Survey Errors and Survey Costs*. 1989. John Wiley & Sons, New York.
- Blumberg, S. and Cynamon, M., Misreporting Medicaid Enrollment: Results of Three Studies Linking Telephone Surveys to State Administrative Records. 2000. Paper presented at the Seventh Conference on Health Survey Research Methods, September 24-27 1999, Williamsburg, VA. To appear.
- Bennefield, Robert, *Health Insurance Coverage 1997: Consumer Income*. Current Population Reports, U.S. Census Bureau: 1998, 60-202.
- Camburn, D., Lavrakas, P.J., Battaglia, M.P., Massey, J.T., and Wright, R.A. Using advance letters in random-digit-dialing surveys. In: *Proceedings of the Section on Survey Research Methods*; 1995, May 18-21, Ft. Lauderdale, FL. Alexandria, VA: American Statistical Association; 1995. P. 969-974.
- Massey, J. (1995). Estimating the response rate in a telephone survey with screening. (1995). *Proceedings of the Section on Survey Research Methods*, Alexandria, VA: American Statistical Association, 673-677. This paper is posted at <http://www.nisabt.org>
- Kish, L. *Survey Sampling*, John Wiley and Sons, Inc. NY: 1965.
- Smith, P.J., Rao, J.N.K., Daniels, D., Battaglia, M.P., Ezzati-Rice, T.M., and Khare, M. Compensating for nonresponse bias in the national immunization survey using response propensities. NCHS Series 2 Report, to appear. (2001).