National Immunization Survey

Guide to Quality Control Procedures

Centers for Disease Control and Prevention

National Immunization Program
and
National Center for Health Statistics

Prepared by Abt Associates Inc.
February 2002

Table of Contents

Appendices

Glossary of Terms and Acronyms

C-Tables	Detailed daily production statistics on household data collection for each IAP area.
Callback	The redialing of a telephone number on a different day and/or time from the previous call-attempt. The purpose of a callback is to resolve a sampled telephone number.
CATI	Computer Assisted Telephone Interviewing: a comprehensive computer system for telephone surveys that controls the administration of the questionnaire and skip-pattern logic, and saves the data collected directly into database files. The NIS CATI system also controls the sample delivery process.
CASRO	Council of American Survey Research Organizations: An organization for survey research professionals and organizations. Standard CASRO rates are used by survey researchers to compare survey results against other surveys.
CDC	Centers for Disease Control and Prevention
CLAF	Child-Level Analysis File: the file that integrates the data collected from a household with data collected from a child's providers.
GENESYS-ID	A contract software system that identifies business telephone numbers from directory listings and uses special hardware and software to detect nonworking numbers.
IAP area	Immunization Action Plan area. The NIP has 78 IAP areas, covering the 50 states and 28 metropolitan areas.
IHQ	Immunization History Questionnaire (mailed to providers)
NCHS	National Center for Health Statistics
NHIS	National Health Interview Survey
NIP	National Immunization Program
NIS	National Immunization Survey
PRCS	Provider Record Check Study
PUF	Public-Use File: name given to the final data files delivered to NCHS for public access.
Q-Based Report	A report used to manage sample release for the NIS. Generated daily, the report details the number of cases available to be called in each type of calling queue by time zone (i.e., Spanish Language, Not Yet Released [Fresh] Cases, Prior Refusals, Appointments, Other Language, etc.).
Q1, Q2, Q3, Q4	The quarters of a calendar year (e.g., Q1 comprises January, February, and March).

1. Overview

A survey ensures the validity of its data and the accuracy of the resulting estimates by careful attention to quality at each step, from the design of the sample and development of the instruments through sample preparation and data collection to the cleaning and editing of the data and the calculation of sampling weights and estimates. To accomplish the goals of the Centers for Disease Control and Prevention (CDC) in surveillance of childhood immunization in the United States, in each state, and in 28 urban areas, the National Immunization Survey (NIS) incorporates procedures throughout its operation, aimed at promoting and maintaining quality. The present guide describes those procedures and, in many instances, illustrates their application. It pertains to data collection for 1994 through 2000, though some procedures are the result of ongoing development and have not been in place for the entire period. Its organization largely follows the stages of the survey itself. The details of the National Immunization Survey are discussed further in the User's Guide for the Public-Use Data File (one for each year for which a public-use file has been released) and in the Annual NIS Methodology Reports (some of which are internal CDC documents).

Chapter 2 gives an introduction to the National Immunization Survey and its sample design and data-collection methodologies. Briefly, in each of 78 Immunization Action Plan areas (which together make up the U.S.) the NIS draws independent quarterly samples of telephone numbers and then uses random-digit dialing (RDD) to identify households that have one or more children 19 to 35 months of age. The telephone interview continues by collecting vaccination information for each such child and requesting consent to contact the providers of the child's vaccinations. In a second phase, a mail survey, the NIS asks the providers to report vaccination information from the child's medical record, which is generally more accurate and complete than the household's information. Chapter 2 also describes the roles of informed consent, confidentiality, and security in the NIS.

Chapter 3 surveys the quality control procedures in the NIS, placing them in the context of a focus on nonsampling errors. It describes the management structures that support quality assurance, summarizes a number of quality control research studies, discusses internal and external assessments, and lists the main indicators that are monitored on an ongoing basis.

Chapter 4 discusses the quality-related aspects of the components of collecting the household data: sample preparation, development and modification of the household questionnaire, mailing advance letters, sample management, management of the interviewers, the computer-assisted telephone interviewing system, processing consent forms, and the daily production reports.

Similarly, Chapter 5 covers the Provider Record Check Study, including the Immunization History Questionnaire, the comprehensive tracking system, the steps in collecting data from providers, data preparation, and data entry.

Chapter 6 describes the process involved in further preparing the household data and the provider data and then combining those two streams in preparation for estimation and analysis. Of particular importance are the provider edit program (which assembles each child's data and constructs a composite vaccination history) and the matching-sheet review (which investigates specific types of discrepancies in the resulting data).

Chapter 7 discusses steps related to estimation, with particular emphasis on sampling weights. Between the base sampling weight (initially assigned to each child with a completed household interview) and the final weight (for each child with adequate provider data) lie a number of adjustments that ensure the appropriate relation between data and population.

Chapter 8 summarizes a number of studies that evaluate the quality of the data and of the estimates of vaccination coverage.

Finally, Chapter 9 summarizes the ingredients and techniques that contribute to quality in the NIS. It includes a discussion of future steps and research that is already under way.

2. Introduction

Vaccinations are among the most effective public health interventions to prevent disease and death (1). In 1994 the Childhood Immunization Initiative (2) was established 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) improve vaccinations and their use; and 5) monitor vaccination coverage and occurrences of disease. Subsequently the Healthy People 2000 (3) and Healthy People 2010 (4) objectives established the goal of having at least 90% of two-year-old children fully vaccinated with the recommended schedule of vaccines (5). The National Immunization Program and the National Center for Health Statistics of the Centers for Disease Control and Prevention contracted with Abt Associates Inc. to implement the National Immunization Survey (NIS), to accomplish the mandate of monitoring vaccination coverage and marking progress toward achieving the Healthy People 2000 and 2010 goals.

The NIS collects vaccination information that allows coverage rates to be monitored at national, state, and local area levels (6). Additionally, the objectives of the NIS are to assist the CDC in allocating resources to states for the purpose of increasing coverage rates, to identify subpopulations and/or geographic areas in which rates are low, and to provide a data base for epidemiological research. In order for the NIS to meet all its intended uses, the quality of the collected information must be in line with these goals. The continuous quality program in the NIS aims to evaluate the quality of the data and implement procedures that enhance quality, to produce timely and accurate estimates of coverage rates for surveillance, and to prepare a public-use data file that ensures confidentiality of information collected in the NIS and avoids risk of disclosure under Section 308(d) of the Public Health Service Act (42 U.S.C. 242m).

2.1 Background

The target population for the NIS is children 19 to 35 months of age living in the United States at the time of the interview. Official coverage estimates reported from the NIS give rates of being up-to-date (UTD) with respect to the recommended numbers of doses of six vaccines (5). These vaccines and their recommended number of doses are: diphtheria and tetanus toxoids and pertussis vaccine (DTP) – 4 doses; poliovirus vaccine (polio) – 3 doses; measles-containing vaccine (MCV) – 1 dose; *Haemophilus influenzae* type b vaccine (Hib) – 3 doses; hepatitis B vaccine (Hep B) – 3 doses; and varicella zoster vaccine (varicella) – 1 dose. In addition to these vaccines, interest focuses on coverage rates for series of vaccines, including the 4:3:1:3 series (4 DTP, 3 polio, 1 MCV, and 3 Hib). Annualized vaccination coverage estimates are published on a semi-annual basis on the National Immunization Program's Web site and in *Morbidity and Mortality Weekly Report*. The time lag between the end of a reporting period and publication of official estimates is approximately 6 months. This timing enables the NIS to monitor the use of new vaccines that are added to the recommended schedule (for example, varicella vaccine was introduced in 1996 and promptly added to the NIS). The NIS also collects data on vaccinations with rotavirus vaccine and pneumococcal vaccine, but it does not currently produce coverage

estimates for these vaccines (rotavirus vaccine was added to the recommended schedule in September 1998 but suspended in July 1999 and removed in October 1999, and pneumococcal conjugate vaccine was added in 2001).

Beginning with the second quarter of 1994, the NIS has conducted independent quarterly surveys in 78 Immunization Action Plan (IAP) areas, consisting of the 50 states, the District of Columbia, and 27 other large urban areas. This design has made it possible to produce annualized estimates of vaccination coverage levels for the six vaccines within each of the 78 IAP areas with an acceptable degree of precision (a coefficient of variation of no more than 5 percent). Further, by using the same data collection methodology and survey instruments in all IAP areas since 1994, the NIS produces vaccination coverage estimates that are comparable among IAP areas and over time.

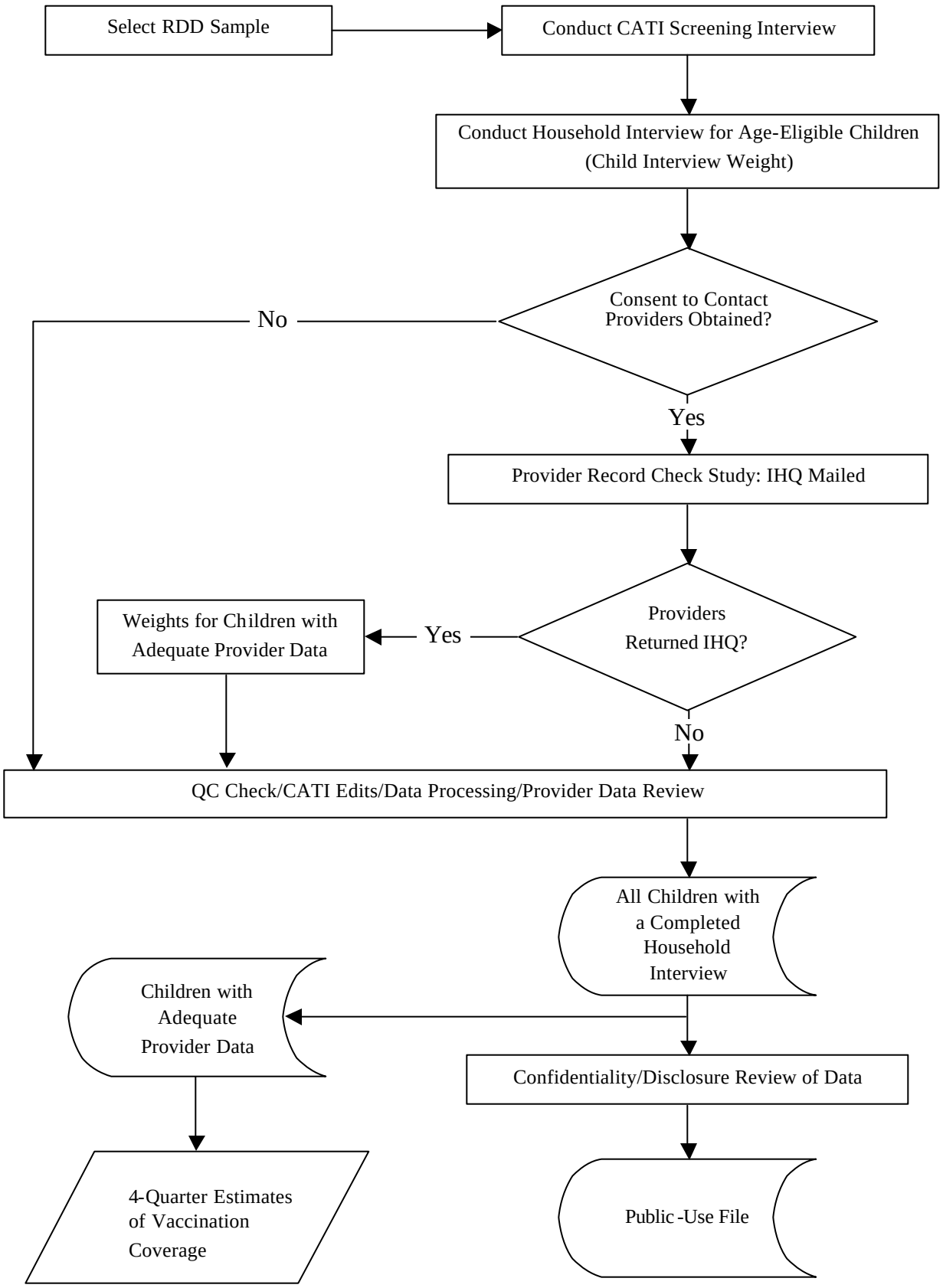
2.2 Sample Design and Data Collection Methodologies

The NIS collects immunization data from two sources—a telephone survey of households and a mail survey of immunization providers identified by household respondents. For the household survey, the NIS employs a list-assisted random-digit-dialing (RDD) sample design (7,8,9). Figure 2.1 shows major steps in NIS data collection (and quality control), culminating in production of estimates of vaccination coverage and an annual public-use file.

The number of completed telephone interviews needed in an IAP area varies each quarter. A total of approximately 35,500 children with a completed household interview, yielding approximately 22,000 children with usable immunization records from providers, are needed in four consecutive quarters to calculate coverage estimates within targeted levels of precision. The percentage of telephone interviews that result in acquisition of provider data varies among IAP areas. The targets for numbers of completed telephone interviews for each IAP area are set to compensate for this variation. Furthermore, if the number of completed telephone interviews in three consecutive quarters falls short of the goal, the target for the subsequent quarter is increased.

Using the current frame of active telephone exchanges, a sample of telephone numbers is drawn quarterly for each IAP area and divided into random subsamples called replicates. Before a replicate is loaded into the computer-assisted telephone interviewing (CATI) system, many nonresidential numbers are eliminated by comparing the sampled numbers with databases of listed residential and business numbers. An automated dialing system is used to identify as many nonworking numbers as possible. These numbers are removed from the set of numbers that is sent to the telephone center for interview. The remaining numbers are dialed and screened by telephone interviewers to determine the presence of eligible children.

Figure 2.1: Flowchart of Data Collection and File Production in the NIS



About 10 days before the sample telephone numbers are dialed, an advance letter is mailed to the sample numbers for which an address is obtained through a reverse-match service. The advance letter is used to increase participation and emphasize legitimacy of the survey.

When a household with an eligible child is located, a vaccination interview is conducted with the adult in the household who is most knowledgeable about the child's vaccination history. An interviewer obtains numbers of vaccination events (and, if possible, their dates) for each age-eligible child. Household respondents are encouraged to use written vaccination records ("shot cards") to enhance the accuracy of the reported data. Demographic information about respondents is also collected.

Finally, respondents are asked for names and addresses of immunization providers and for verbal permission to contact these providers to request immunization data from the child's medical records. An area-specific lookup database of previously identified providers is used to reduce data entry errors. The identified immunization providers are contacted by mail and asked to report vaccination histories from the children's medical records, along with some information on their practice. Providers that do not respond to the initial request are mailed a second questionnaire and, if necessary, telephoned in order to maximize provider participation in the survey.

The NIS uses state-of-the-art data collection and data processing systems with numerous built-in quality control functions to collect and process immunization histories from the households and the providers. The ongoing quality control (QC) procedures ensure the validity and accuracy of the vaccination coverage estimates. Detailed discussions of the history, sample design and quality control procedures of the NIS have been published (10,11,12,13).

2.3 Informed Consent, Confidentiality, and Security

During the household interview the parent or guardian reports demographic and medical data for the child and may also give consent for the NIS to contact the child's provider(s). Once the data have been collected, all personnel who have contact with them are responsible for preserving their confidentiality. This section describes the quality control procedures for maintaining the security and privacy of data as mandated by law and keeping the survey's obligations to people who provide their data.

All employees of CDC (NCHS and NIP) and Abt Associates who work with NIS data are required to sign an affidavit stating that they will not disclose information that would allow survey participants to be identified. (Unauthorized disclosure of the confidential information is punishable under Title 18, Section 1905 of the U.S. Code.) Additionally, all survey participants are informed about their rights under the Privacy Act of 1974, which mandates that their data are to be kept confidential. The confidentiality of respondents and their data is also assured by the Public Health Service Act. NCHS and Abt Associates take precautions at each stage of the data

collection and publication process to ensure that the confidentiality of participants' data is maintained. These include comprehensive security measures for facilities and computers.

2.3.1 Informed Consent

Informed consent is obtained from survey respondents to participate in the household survey and, separately, to contact the child's vaccination providers. The providers give consent by completing and returning the sample child's Immunization History Questionnaire.

Household Data Collection

The NIS household data collection begins with screening questions to identify households that have a child 19-35 months old—the focus of the main survey data collection. The following introduction informs potential household respondents of their rights, that the survey is authorized by law, and that their confidential information and identities will be protected.

> S3_INTRO This study is voluntary and is authorized by the U.S. Public Health Service Act. The information you give will be kept in strict confidence and will be summarized for research purposes only. It's all right to skip any questions you don't want to answer.

As described in Chapter 5, under Collection of Provider Data, during the telephone interview the NIS interviewer asks the parent or guardian of an eligible child for his or her permission to contact the child's vaccination providers and request vaccination information from the child's medical record. If consent is given, the telephone interviewer documents this fact. This documentation is supplied to the vaccination provider when the request is made to complete an Immunization History Questionnaire for a child. Care is taken in the interview to ensure that the person giving consent has the authority to do so. Then the parent or guardian is informed that vaccination records will be held in strict confidence. Finally, the interviewer asks the following question, after which he or she records the names of the child and parent or guardian, taking extra care to get the correct spelling.

> D7. Do we have your permission to contact the provider(s) named in this interview, give the provider(s) basic information that identifies your child(ren), and request that information relevant to your child(ren)'s immunization history be sent to the Centers for Disease Control and Prevention or its contractors for study purposes only?

Provider Data Collection

After a household interview has been completed and consent to contact immunization providers has been obtained, the next step is to contact the providers to request the child's immunization history from medical records. Providers are mailed a package containing a cover letter that

includes the following statement about the confidentiality of the data, the authorization for the data collection program, and the provider's rights:

> The information you give will be confidential, as specified by law in Section 308(d) of the Public Health Service Act. The Centers for Disease Control and Prevention, its contractors, and state and local immunization program staff use the immunization information for statistical purposes only. No information is released that could identify you, your practice or facility, the child, or the child's family. This study is authorized by Section 306 of the Public Health Service Act and The National Childhood Vaccine Injury Act of 1986. Your participation is voluntary.

The mailing package also includes the IHQ and a copy of the consent form, signed by either the telephone interviewer who spoke with the parent or guardian of the child or the interviewer's supervisor.

2.3.2 Confidentiality

Employee Affidavits

The NIS employs procedures at every stage of data collection and processing to ensure that identifying information is kept confidential and within the control of staff sworn to protect it. All CDC (NCHS and NIP) and Abt Associates staff working on the NIS are sworn to maintain respondents' privacy by keeping identifying data secure.

NIS staff, particularly data collection staff, are given instructions on protecting confidentiality based on NCHS standard procedures and on the specific features of Abt's physical facilities. All staff who have access to confidential information are required to sign an Affidavit of Non-Disclosure, a standard requirement for all permanent and temporary employees of the company. These staff also sign Affidavits of Non-Disclosure required by the CDC. The security training of all NIS staff includes viewing the NCHS video on computer security. Prior to each quarter of NIS data collection, confidentiality requirements are reviewed with telephone, locating, mailing, and data editing staff. The performance of and retention of staff are also reviewed at this time, with regard to individual observance of confidentiality and non-disclosure policies and procedures. Violation of the signed agreement is grounds for immediate dismissal.

Confidentiality of data collected in the NIS is mandated by federal law--the Privacy Act of 1974 (5 U.S.C. 552a) and Section 308(d) of the Public Health Service Act (42 U.S.C. 242m). Survey findings are released to the public only in summary form. The law requires that the names of study participants not be associated with any answers. Any willful and knowing disclosure in violation of the Privacy Act of 1974 would subject the violator to a fine of up to $5,000.

Under the Public Health Service Act, "Unauthorized disclosure of the confidential information is punishable under Title 18, Section 1905 of the U.S. Code." The following is excerpted from that section. In the case of the NIS, it applies to employees of CDC:

Whoever, being an officer or employee of the United States or any department or agency thereof, publishes, divulges, discloses or makes known in any manner or to any extent not authorized by law any information coming to him in the course of his employment or official duties or by reason of any examination or investigation made by, or return, report or record made to or filed with, such department or agency or officer or employee thereof, which information relates to trade secrets, processes, operations, style of work, or apparatus, or to the identity, confidential statistical data, amount or source of any income, profits, losses, or expenditures of any person, firm, partnership, corporation, or association; or permits any income return or copy thereof or any book containing any abstract or particular thereof to be seen or examined by any person except as provided by law, shall be fined not more than $5,000 or imprisoned not more than one year, or both; and shall be removed from office or employment.

Contractor employees are subject to the same provisions and penalties. The following is excerpted from the affidavit that every employee of Abt Associates who works on the NIS must sign. This form is signed after staff have viewed a video produced by NCHS that explains the law and how it supersedes all other laws that govern release of information, including the Freedom of Information Act (FOIA), court subpeonas, and even Presidential orders:

Assurance of Confidentiality

"In accordance with Section 308(d) of the Public Service Act (42 U.S.C. 242m), Abt Associates Inc. assures all respondents that the confidentiality of their responses to this information request will be maintained by Abt Associates and the Centers for Disease Control and Prevention (CDC) and that no information obtained in the course of this activity will be disclosed in a manner in which the individual or establishment is identifiable, unless the individual or establishment has consented to such disclosure, to anyone other than the authorized staff of CDC."

I have carefully read and understand the assurance which pertains to the confidential nature of all records to be handled in regard to this survey. As an employee of the contractor I understand that I am prohibited by law from disclosing any such confidential information which has been obtained under the terms of this contract to anyone other than authorized staff of CDC. I understand that any willful and knowing disclosure in violation of the Privacy Act of 1974 is a misdemeanor and would subject the violator to a fine of up to $5,000.

Disclosure Review

Prior to release, the contents of the public-use file undergo an extensive review by the NCHS Disclosure Review Board to protect the confidentiality of the survey participants. Additionally, other data products and reports are reviewed to ensure that cells in summary tables represent a sufficient number of sample cases to protect against disclosure. Cells that contain too few cases are suppressed or collapsed with other cells.

2.3.3 Security of Facilities

Security measures in the facilities where data are collected and stored contribute to protecting the confidentiality of respondents' information and reduce the risk of tampering. In each of the telephone data collection centers used for the NIS:

- Doors to the telephone center are always locked.
- Interviewers and other staff must show an ID badge for admittance to the telephone center.
- Doors to rooms containing computer network servers are always locked, and the rooms are accessible only to authorized personnel.

Because most of its studies require collecting and storing private information, Abt Associates devotes particular attention to issues of security. The building managers of all Abt facilities provide twenty-four-hour, seven-day-a-week security services with restricted access in the evenings and weekends based on tenant-approved access lists and positive identification from individuals.

Only authorized personnel are allowed access to confidential records, and only when it is appropriate for each individual — that is, only when his/her work requires it. For example, Abt maintains a separate, secure area to receive and process questionnaires returned for the NIS Provider Record Check Study. This area is accessible only to data preparation staff and appropriate project staff. Safeguards are in place to ensure control over survey returns at all times during transit between the Post Office and Abt's Data Preparation group through regular review of transport procedures with mail messengers and U.S. Postal Service personnel.

Provider surveys returned by facsimile transmission are received by a fax machine housed in the secure data preparation area and used only for this purpose. A high-quality printer located in the data preparation area is used to print all child identification labels for NIS mailings. Both the facsimile machine and high-quality printer are separate from the equipment used for Abt's corporate business and on other projects. Such separation of activities helps to ensure the confidentiality and security of data from NIS children and providers.

When not in use, NIS consent forms and survey returns are maintained in locked files, accessible with keys held only by the data preparation manager and a small number of supervisors assigned to the NIS. Abt's data preparation area also contains a locked supply room, accessible only to

authorized data preparation staff. Respondent names and addresses are kept separate from completed instruments, and individual respondents are identified only by Abt-generated identification numbers. Paper copies containing information that identifies individual respondents are kept in locked file cabinets. As noted above, only staff who need to know have access to these data, so that very few individuals are allowed to open these locked cabinets. Additional security measures are used to protect respondents' data residing on the computers in the telephone centers and offices of Abt Associates and NCHS. These measures are described in the next section.

2.3.4 Computer Security

The NIS data are stored on Local Area Networks (LANs) that are secured by login IDs assigned to individual project staff. There are no generic or temporary IDs. Each ID has a password that expires periodically. Every password is unique and has electronic intruder protection. The servers on these LANs are located in rooms that only IT personnel, certain office service personnel, and telephone center managers can enter. The servers are backed up to tape on a daily basis, and the tapes are shipped off-site.

Telephone interviewers have to log on to the system using a bar-coded ID badge. They must go through an additional layer of security by entering a user name and login code to enter the system for conducting interviews. All Abt Associates and CDC (NCHS and NIP) personnel must have network login accounts and passwords to access data files on their respective systems. These security measures ensure confidentiality by limiting access to respondent information to only those who have been trained and have signed confidentiality agreements (see Section 2.3.2). When someone terminates employment, that person's access rights and passwords are invalidated immediately, and the user name is removed from the computer system.

All data on site are kept under lock and key, and accessed only on an as-needed basis by individuals who have signed the security pledge and are authorized to work with the data. Access to the off-site storage facility is controlled through user names, passwords and account numbers. Data from this facility are returned only after the correct combination of user name, password, and account number has been supplied. Storage media are shipped to and from the storage facility using a bonded carrier.

Abt Associates limits access to data processing areas, allowing only authorized personnel in computer rooms and computer tape libraries. Locked tape files and storage areas are provided to all projects. Files containing respondent names and addresses are separately maintained with additional password protection so that data storage and analysis files contain only Abt-generated respondent identification numbers. In addition, individual data banks and files are protected by passwords and other techniques that prevent access by non-approved project staff. Access to areas where confidential data are maintained is restricted to authorized personnel.

In addition to the protection of privacy, data security encompasses backup procedures and other file management techniques, such as mirror-image disk drives, to ensure that files are not inadvertently lost or damaged. Published procedures are applied to identify important data files that are routinely backed up to tape cartridges or tape. Additional file protection is provided by procedures that prevent unauthorized changes or access to data files. Currently, all databases are password protected, with only data system administrators having write authority over files.

Security Awareness on NIS

All staff working on the NIS are aware of their responsibilities to follow sound security practices in the workplace. A pre-logon warning message is displayed when staff access the system, with a notification of appropriate computer use and a warning about the ramifications of unauthorized use. Ongoing communication and training emphasize the need for strict adherence to security, concentrating on password guidelines, virus scanning, data security, and reporting procedures for communicating problems.

Virus Scanning

Virus scanning is a critical component of computer system security and is automatically performed every day on every PC. McAfee VirusScan is the standard virus scan utility used on all PC workstations and the Local Area Network. All files received from outside sources are virus-scanned prior to use to ensure that they are virus-free.

Abt recognizes the importance of immediate reporting of any unit or network infection. Such reports go immediately to the Information Technology (IT) department. All public notifications of potential new viruses come or are forwarded to IT for verification; IT then alerts all company staff.

Reporting Procedures

Each Abt employee is required to report any suspicious and all confirmed violations to the Survey Group Programming Manager and to the relevant Project Director. It is the responsibility of these senior staff to review and evaluate each incident and to ensure that appropriate action is taken. Minor incidents may be resolved through a review of communication protocols and re-training. In the case of major violations, a report would be made to a Group Vice President and a Human Resources Representative and could result in termination of employment.

3. Quality Control Procedures

Two classes of error enter into all survey data—sampling and nonsampling errors—and both affect the accuracy of survey estimates. Sampling errors (or sampling variation) arise from chance mechanisms and fluctuations. The extent of sampling variation is mainly a consequence of sample design and sample size. All other types of survey error fall under the heading of nonsampling errors. They represent possible sources of bias, and hence are the primary focus of quality control efforts. Thus, examination of nonsampling errors is a major component of the NIS.

Nonsampling errors can occur in a variety of ways, including nonresponse, questionnaire content deficiencies, modes of data collection, inconsistencies in collected data, and errors arising in data editing and processing. The NIS mandate and its data collection methodology have evolved, necessitating a dynamic process for data quality control. This report provides an overall perspective on how the survey sponsors and collaborators have viewed quality control and adopted measures to ensure that the resulting data are of high quality and of maximum usefulness to the National Immunization Program and other researchers and policymakers.

3.1 The Quality Perspective in the NIS

The term "data quality" can have many definitions. As adopted by the National Immunization Survey and used throughout this document, high-quality data are defined as information that is timely, accurate, relevant, accessible, and cost-effective. Table 3.1 defines these terms, as used throughout this document.

Table 3.1
Elements of Quality in the NIS

Quality Element	Definition
Timeliness	Information is available within a period of time during which it remains useful.
Accuracy	Estimates of indicators of interest are close to the true value in the population.
Relevance	Data collection instruments and procedures are designed and implemented in accordance with clearly defined survey objectives.
Accessibility	Information is available to data users and other researchers and policymakers in a format that is easy to use and understand.
Cost-effectiveness	The information and accuracy required for all desired subgroups could not be obtained for less money.

The quality control (QC) procedures documented in this guide are directed at enhancing one or more aspects of these quality elements by promoting efforts to reduce errors that can affect each of them. As a framework for the discussions of specific procedures in subsequent chapters, Table 3.2 enumerates the main components of each phase of the NIS.

Table 3.2
Phases of the NIS and Their Major Components

Household Data Collection
Sample Preparation
Instrument Development
Mailing of Advance Letters
Sample Management
Training and Monitoring of Interviewers
Operation of the CATI System
Processing of Consent Forms
Re-interview Program
Production Reports

Provider Record Check Study
Instrument Development
Electronic Tracking System
Data Collection
Consent forms
Locating providers
Mailing the IHQ
Telephone prompting
Data Preparation
Data Entry

Preparation of Data Files
Household Data File
Cleaning, editing, and file construction
Automated back-coding of vaccines
Automated back-coding of race and ethnicity
Creation of composite variables

Preparation of Data Files (continued)
Provider Data
Unduplication
Editing
Provider Edit Program
Matching-Sheet Review
Construction of Analysis Data File

Weighting and Variance Estimation
Imputation for Item Non-Response
Base Sampling Weights
Adjustment for Multiple Telephone Lines in Household
Adjustment for Unit Nonresponse
Poststratification
Compensation for Nontelephone Children
Adjustment for Provider Nonresponse
Estimation
Variance Estimation

Evaluation
Comparison with External Sources
Sample Characteristics
Immunization Estimates
Coverage of Telephone Population
Nonsampling Errors
Noncoverage error
Nonresponse error
Measurement error

For the three categories of nonsampling errors (14) that fall within the quality control guidelines of the NIS, Table 3.3 lists specific types of error, with examples.

Table 3.3
Types of Nonsampling Errors

Category	Types of Error	Examples in the NIS
Coverage of frame	Incomplete coverage	Nontelephone households
	Overcoverage	Businesses included in frame
	Duplicate listings	Households with more than one telephone number
Nonresponse	Unit nonresponse	No consent to contact provider
	Item nonresponse	Inadequate provider addresses
		Refused or missed questions
Measurement	Inconsistency	Household versus provider reporting vaccination history
	Data editing & processing	Transcription errors
	Interviewer-related	Classification of responses
	Questionnaire-related	Screening text
	Respondent-related	Recall error/correct information

3.2 Organization of NIS Quality Management Oversight

The National Immunization Program (NIP) and the National Center for Health Statistics (NCHS), of the Centers for Disease Control and Prevention (CDC), and Abt Associates have designed and implemented the NIS. Physicians, epidemiologists, survey statisticians, sociologists, and information system designers evaluated the goals of the National Immunization Program and designed the National Immunization Survey to achieve those goals. Experts from these organizations form task groups to oversee various aspects of the NIS. Although several task forces have been involved in quality control (estimation task force, survey operations task force), the methodology task force has had the primary responsibility for the effort. This task force has both promoted and sponsored a wide array of experiments and research projects related to aspects of NIS quality, as well as many *ad hoc* analyses. Table 3.4 gives selected recent examples of quality-related studies performed under the NIS, many of which originated from the various task forces.

3.3 Quality Control Standards Affecting the NIS

The NIS adheres to all federal standards concerning data confidentiality, security measures, and data collection integrity. Because its funding is federal, the project is subject to all federal rules and regulations applicable to government-supported research. All standards and regulations in the Code of Federal Regulations concerning data collection funded through the government apply to the NIS. Further, the NIS is subject to the Privacy Act, directives from the NCHS Disclosure Review Board, and Federal Information Processing Standards with respect to data security.

Table 3.4
Examples of Recent Research Studies That Affect Quality Control in the NIS

Years	Description	Area of NIS	Area of Quality	Results
1997	Mail carrier experiment	Provider record check study	Variation in provider response rates by mailing carrier	Examined provider response rates based on whether the provider survey was sent using First Class mail versus second-day-delivery Federal Express. No noticeable difference was detected, and the lower-cost First Class mail was kept.
1997 - 1998	Behavior Coding Study	CATI Response/ Entry study	Assessed interaction between respondents and interviewers	By question, examined the number of questions that were interrupted, had a wording change when reading, required a probe, had an incorrect probe (when required), and had responses entered incorrectly. Found some problematic questions and incorrect procedures on the part of interviewers.
1998	NIS reinterviews	Household questionnaire and RDD	Reliability of household and child eligibility, information provided during NIS interview	Found good reliability (>90%) of non-working/nonresidential numbers, number of persons in household, number of age-eligible children in household, child's DOB, and sex of child; lowest reliability for availability of vaccination records, number of providers, and education of mother.
1999	Noncoverage of nontelephone children	Weighting and estimation	Vaccination coverage estimates improved by better weighting	Statistical weighting methods based on interruptions in telephone service were developed to compensate for the exclusion of nontelephone children and thus reduce bias.
1999	Provider questionnaire modification	Provider record check study (PRCS)	Update the IHQ	Added new vaccines and reorganized the IHQ to list single-antigen vaccines separately from combination vaccines.
1999	IHQ reviews by editors	PRCS	Inconsistencies and discrepancies in dates	Approximately 80% of errors were in dates recorded on original IHQ forms.
1999	Double data keying	PRCS	Errors in data-entry	Of 25,116 fields entered, 0.36% had errors.
1999	Discrepancies between household and provider data and within provider data	PRCS	Extensive matching-sheet review of household and provider data	Procedures were implemented to correct as many discrepancies as possible without recontacting any households or providers.
1999	Comparison with other published estimates	Household and PRCS	Comparison with area-based rather than telephone-based estimates	The NIS and the NHIS produce comparable estimates of vaccination coverage, despite their differences in methodologies and sample sizes.

3.4 Internal Assessment of Quality

Beyond the global mechanisms in place, the NIS has enhanced project-specific quality control through a continuous process of improvement and change. The survey adopted several systems that facilitate widespread dissemination of information and enable a large group of survey experts to examine the NIS data. The quality control system includes:

- ongoing reports in sample management;
- interviewer monitoring;
- continuous training and retraining of interviewers;
- information mining to develop policies that promote cost-effectiveness;
- extensive research into the best means of obtaining provider response; and
- use of groups for problem identification and problem solving.

The typical process first identifies an issue affecting quality or problems with data collection and raises it with appropriate task forces and the survey sponsor. Selected NIS team members would discuss the issue or problem and then might decide: 1) to examine the issue internally through a study or reassessment; 2) to bring in one or more outside experts to assess the issue and examine it; or 3) to organize an external board to examine some aspect of the NIS or to design a means to assess some aspect of the study. Many of the assessments have involved special data collection efforts and/or field tests.

3.5 External Assessments of Quality

As noted, one means through which the project has pursued data quality assessments is to have outside experts (under a confidentiality agreement) examine data quality issues or conduct an outside study. Additionally, two substantial efforts have produced independent assessments of the NIS. An independent firm was contracted to conduct an evaluation in 1996 on behalf of NIP, and the General Accounting Office (GAO) conducted a project assessment in 1996-1997.

In the first assessment the NIP asked the contractor to review the NIS and determine whether it was meeting NIP's programmatic needs. Further, this evaluation focused on whether less expensive data collection methods could be used and still meet NIP's information needs. The evaluation concluded that "the survey is a well designed response to a very difficult set of goals posed by CDC" (15). Four areas were identified for possible cost reduction: call scheduling, procedures for resolving whether randomly selected telephone numbers are working residential numbers, procedures for drawing the sample, and increasing the proportion of children for whom provider data are obtained. Steps have been taken to implement cost-saving changes in these areas. Several data analysis changes were also recommended, including the use of provider-adjusted weights for children with provider

data, separate weights for national analyses, the use of alternative methods to adjust for noncoverage of nontelephone households, variance estimates that better reflect the study's complexity, and use of imputation to compensate for missing provider data.

In response to a congressional request the GAO evaluation focused on the ability of the NIS to identify pockets of children in need of immunization in the U.S. Initially, the Department of Health and Human Services had told Congress that this was one of the NIS's uses; however, the CDC corrected this statement, and the GAO concluded that the survey did not and could not serve this purpose (16). The GAO concluded that the NIS was not designed to identify groups of children in need of immunization, although state differences had been identified and states had taken steps to address low immunization coverage. In terms of data quality, however, the design of the NIS reflected its purpose as originally outlined. Thus, although it identified some areas for improvement, the GAO review found that the NIS was a sound program.

3.6 Review Process

Within the NIS data collection program, several mechanisms exist to review and maintain data quality on a routine basis. For example, each data deliverable receives a technical review. In addition to the regular weekly calls on specific topics, biweekly management meetings examine and discuss all upcoming deliverables and identify potential quality problems. These regular discussions ensure that the major data quality indicators are examined on a daily, weekly, or monthly basis by data collection experts.

Early in the NIS, sets of major data quality indicators were defined for both the household data and the provider data. These indicators are arranged in tables that are produced weekly: the C-Tables for the household data and the Weekly Provider Record Check Study Summary Tables. The major items of information contained in these production and reporting tables are summarized in Table 3.5 and Table 3.6, respectively. Some items are reported for each replicate or for each IAP area. Items such as response rate, eligibility rate, and screening rate are all indicators of data quality and potential problems in the data collection. Daily versions of these tables are monitored to ensure that the quarter is progressing as expected. Many possible problems with either data collection or data quality can be identified through these tables.

Table 3.5
Key Indicators for the Household Survey

Total sample in all released replicates
Telephone numbers found to be nonworking or nonresidential through list-assisted and pre-CATI auto-dialing procedures
All telephone numbers initially loaded into the CATI system
Resolved sample
Households screened
Eligible households (%)
Incomplete interviews and refusals
Completed household interviews

Table 3.6
Key Indicators for the Provider Record Check Study

Children with eligible dates of birth
Primary children within complete households
Children with usable consent
Children with immunization information from any provider
Usable child/provider pairs
Original requests shipped
Reminder letters shipped
Telephone prompting calls attempted
Provider surveys returned
Surveys returned with immunization information

Additionally, calls and inquiries from states and other researchers act as a catalyst to examine other areas that may suggest potential data quality problems or to enhance current QC procedures. All such contacts are discussed in weekly meetings. These queries or questions often identify areas requiring closer scrutiny. In summary, the NIS has regular discussions with a diverse group of experts associated with the survey to ensure that areas that need improvement are monitored and discussed.

4. Collection of Household Data

The household data collection serves as the foundation for later stages, such as contacting providers, weighting the data, and estimating vaccination coverage. Thus it is critical that the data be of the highest quality. This section discusses the household data collection process and the quality assurance methods employed at each stage, from sample preparation through household interviews, including consent to contact providers for eligible children.

4.1 Sample Preparation

Given the task of obtaining 78 independent coverage assessments over a four-quarter period, extensive sample preparation steps are necessary to achieve this goal efficiently and cost-effectively. Each IAP area has a unique set of demographic characteristics that influence the sample needed to complete the work of the NIS. The proportion of possible telephone numbers that are currently in use, the subset of those that are assigned to households, and the percentage of those households that contain eligible children vary greatly from one area to another. Beyond the differences in demographics, people in some areas are more cooperative than others when they are called. The number of telephone numbers selected in each IAP area must accommodate this variation. The average number of selected telephone numbers required for one NIS telephone interview has ranged from 40 in one IAP area to 185 in another.

In the first quarters the NIS had little information on the practical consequences of the inherent differences among IAP areas. It was assumed that differences in response would be random (i.e., that all IAP areas would be comparable). As a result, equal sized samples were drawn for all IAP areas. Consequently, the work in some IAP areas was completed within the first half of the quarter, whereas other IAP areas required more than the full quarter to even approach their targeted number of completed interviews.

Those initial quarters of data collection demonstrated that the NIS would need to use more sophisticated procedures to achieve comparable results from each of the IAP areas. Using data on demographic variables and response rates from earlier quarters (with more weight on more-recent quarters), the size of the sample drawn for each IAP area is calculated each quarter. The goal of this process is to create samples that will yield equal numbers of completed interviews across IAP areas throughout the course of the quarter. This procedure guides case management (as discussed below) and helps to ensure that each case released to the telephone center will have an equal chance to be worked and completed. This minimizes any possible bias associated with time of release within the quarter.

The sample management task involves additional steps to enhance efficiency. First, the RDD sample is drawn only from "banks" of 100 consecutive telephone numbers that contain at least one directory-listed residential number (the "1+ working banks"). Then the selected numbers are processed to eliminate as many unproductive numbers as possible. The two

largest categories, in which many numbers can be automatically eliminated, are numbers that are not in service and numbers assigned to businesses. GENESYS-ID, a component of the GENESYS Sampling System (17) identifies businesses from directory listings and uses special hardware and software to detect nonworking numbers.

Most characteristics of the sample change slowly and consistently—for example, the number of households with eligible children may increase or decrease, but it is unlikely to change greatly in any quarter. One characteristic, however, is more volatile. With recent changes in the uses of telephones (e.g., cell phones, fax machines, and Internet access), demand for telephone numbers has grown dramatically, but unevenly, much faster in urban areas than in rural areas. In response to this demand, many new area codes have been introduced. This results in a dramatic increase in the number of unassigned telephone numbers. Without an automatic process for identifying and eliminating these numbers from the sample, much interviewer labor would be wasted calling out-of-service numbers.

To ensure a high level of efficiency, the NIS regularly considers alternatives in sample preparation. For example, an experiment in 1999 compared GENESYS-ID with GENESYS-ID Plus, a new product designed to remove a higher proportion of business and nonworking numbers from the sample before it reaches the interviewers. The results indicated that, by using GENESYS-ID Plus, the NIS would release roughly 15% fewer telephone numbers. The effect on other operational indicators was also favorable, and the new system was cost-effective. It was phased in during 2000.

4.2 Household Questionnaire

The original NIS survey instrument was developed in 1994. The screener was designed specifically for the NIS, but the immunization section was based on the Immunization Supplement to the National Health Interview Survey (18). The instrument was tested in the field and through in-depth interviews with respondents and focus groups with the interviewers. Complete details on the development of the instrument are given in the *1994 NIS Methodology Report* (19).

The household interviews are conducted primarily in English, but the questionnaire is available in Spanish if needed. The assistance of AT&T Language Line interpreters permits interviewing in a number of other languages. Around 15% of interviews are administered in Spanish, and some 1-2% are administered in other languages.

Project staff and CATI programmers meet regularly to discuss the CATI instrument and ways in which it might be improved. Changes to the instrument, such as wording changes, new questions, and special requests, are proposed and discussed to identify implementation issues that may require field testing and/or special monitoring. Once a minor change is approved, specifications are written for it, and it is programmed. It is then tested extensively by project staff both before and after it has been put into use. Major changes go through a similar process, but may require outside review (for example, by the NCHS Questionnaire

Development Research Laboratory or by an academic specialist) and/or a field test to ensure that the change has the desired effect (usually to improve response rates) and is not detrimental to data collection costs or survey estimates. All proposed and implemented changes are documented in internal memos and reports.

After changes to the instrument have been finalized, the new questions are translated into Spanish for the Spanish-language cases. The translation process occurs in several stages to preserve the quality of the instrument. First, a copy of the English instrument is sent to a Spanish-language translation expert, who translates the altered item(s) into Spanish. The translated version is then sent to another independent translator, who translates the item(s) back into English. Back-translation is a standard method for verifying that the intent of the text is conveyed accurately in the translation process. The use of two contractors ensures that each translation is done independently of the other. Project staff then carefully compare the back-translated version with the original English version. Any discrepancies are discussed with the contractor; and, if necessary, the item is sent back for re-translation. If the newly back-translated version matches the original English version, the Spanish instrument is sent to Spanish-language supervisors in the phone center, who review the new version and report to project staff on any problems.

4.3 Advance Letters

Letters sent to target households prior to telephone contact lend legitimacy to the study, increase participation, and help the respondents prepare for the interview. Approximately 350,000 letters (in English) explaining the purpose, importance, and legitimacy of the NIS are mailed to potential respondents each quarter. Names and addresses are obtained through an address matching service, which translates phone numbers into addresses through commercially available listings, including driver's license bureaus, telephone directories, and food delivery listings. Because advance letters make a difference in getting a completed interview, it is critical to control quality at each stage of the mailing process. First, experiments are frequently conducted to test changes in the content of the letter, with the aim of improving response rates. Second, quality is monitored at each stage of the mailing process.

The advance letter gives the recipient information about the survey, to help reduce refusals and the amount of time on the phone. Whenever the content of the letter is changed, a test is required to ensure that the new letter wording has a positive effect on participation.

Occasionally, experiments are conducted to compare the quality and efficiency of components of the mailing process. For example, one field test compared the telephone-number-to-address matching of the current vendor to a new company to see whether the second company could achieve a significantly higher match rate, resulting in more intended households receiving the advance letter. The second company was in fact able to match more households to addresses, resulting in a switch to this vendor.

The printing subcontractor for the advance letters checks about 1% of the work to ensure the quality of the letter and mailing procedures. Also, before the actual mailing, the contractor sends a proof of the letter to project staff. A code at the bottom of the letter allows them to check that the correct letter is being sent.

Finally, project staff monitor the quality of the mailing itself. "Dummy" names and addresses are given to the mailing subcontractor along with the rest of the addresses to be used in the study. The dummy addresses are selected from a pool of known addresses (i.e., project staff addresses). Recipients of these dummy letters report to project staff when they receive the letter. The time it takes for a letter to reach its destination is tracked (as the mailing dates are known), and the printing quality of the letter itself is checked. Information about the amount of time it takes for a letter to reach respondents enables the NIS to better determine when advance letters should be mailed, so that potential respondents will receive notice of the study shortly before they are called.

4.4 Sample Management

The sample for each IAP area is split into replicates, in order to release only as much sample as is needed to produce the desired number of completed interviews. Actual performance varies as some replicates yield more completed interviews and some fewer. The goal of sample management during the course of each quarter is to control the variation to the greatest extent possible and produce completed work across all IAP areas at approximately the same rate throughout the quarter.

The principal means of achieving this goal is selective case release on a daily basis. At the start of the quarter, cases are released in all IAP areas, and the rate of completing interviews in each area is carefully monitored. Each day, the number of cases projected to result from all sample currently released is calculated and compared to the target number of completed interviews. Additional sample is released in those IAP areas with the lowest projected completion rate. The estimated number of telephone numbers needed to complete the work in each IAP area is calculated, to prevent over-release near the end of the quarter.

A number of monitoring tools have been developed to support this process. The two primary sources are the key indicators in the C-Tables (summarized in Table 3.5), which provide detailed daily production statistics for each IAP area, and the Q-Based Report, which tracks the size of each of the case-delivery calling queues at the start of the day. Once a case is released, it goes into a calling queue that will determine how and when it will be delivered to the telephone center. For example, a case that has been released but not yet called is held in the Fresh Queue until an interviewer dials the case. Likewise, a case with a callback appointment is held in the Future Queue until its appointment time and then sent to an interviewer.

In addition to controlling daily case release, these systems monitor progress within and across IAP areas. This information is used to monitor the number and distribution of

telephone interviews. Current production is projected to estimate future production levels. When staffing in the telephone center is optimal, the work of a quarter will be completed within the quarter, whereas low projections may delay completion of the quarter, and high projections may lead to attrition of interviewers (often the best interviewers) because of insufficient workload.

4.5 Interviewer-Related Activities

Interviewers are fundamental to the success of the NIS in many ways. First, they are the link between the household and the survey; their ability to persuade respondents to participate is essential to obtaining reliable data. Second, their ability to collect the information and code it correctly is also critical to the survey's accuracy. Finally, the interviewers' ability to secure consent from respondents to contact their children's vaccination providers is necessary for obtaining the maximum amount of provider data. Because of the importance of the interviewers' role in the NIS, and many aspects of quality control are targeted toward the interviewing staff. Each quarter approximately 400 interviewers screen over 260,000 households and conduct interviews with approximately 9,000 households having age-eligible children. An NIS interview lasts approximately 20 minutes. On average, 5.6 calls are required to complete an interview.

The main activities familiarize the interviewers with the survey objectives and procedures, introducing them to standard methods for data collection and interviewing, and training them to interact effectively with the household respondents. Standardized interviewer procedures are important in minimizing measurement error caused by differences in the way interviewers perform their work. This effort is especially important in the NIS because interviews are conducted at three telephone centers. To reduce between-interviewer differences in interviews, Abt has implemented extensive training, monitoring, and performance evaluation procedures, discussed below.

4.5.1 Recruiting and Testing

Qualifications of Telephone Interviewers. Prior to attending Basic and NIS training, candidates for the Telephone Interviewer position are invited to an information session, which involves a series of tests and a presentation. The presentation covers the duties of and qualifications for the Telephone Interviewer position. The testing is described below:

Typing Test. Each candidate is given a 3-minute typing test. This test requires basic familiarity with the computer keyboard and a minimum typing speed of 15 words per minute.

Basic Skills Test. This aptitude test has 50 questions. The skills tested include word knowledge, sentence construction, and general information retrieval including reading comprehension. This test is timed at 20 minutes. A passing score indicates a reading level of 10th grade or above.

Oral Reading Test. Each candidate meets individually with a recruiter or a telephone center supervisor. During this short interview the candidate is asked to recite a paragraph aloud. Some medical terms are included to assess pronunciation ability. A five-point scale is used to determine the candidate's overall performance in each of the following areas: Voice quality, articulation, attitude, reading accuracy, and listening skills.

4.5.2 Training

New interviewers receive training on how the CATI system works and on general telephone center operations in a one-day "basic training" session. Once they have successfully completed this training, they attend twenty hours of NIS-specific training over a four-day period. Trainers in each telephone center use the same training materials, ensuring cross-site consistency.

During training, interviewers learn the purpose of the study, learn how to answer common respondent questions, and conduct several mock interviews. On the last day of training they are tested and evaluated. The trainee must pass the evaluation to be officially hired and allowed to start making phone calls. During the new interviewer's first two weeks, he/she is monitored more closely than other interviewers and given more feedback.

The NIS is conducted in English, Spanish, and other languages when necessary. Spanish-language interviewers are specially trained and are monitored by bilingual supervisors. To be considered for conducting Spanish interviews, an interviewer must already have attained a high level of proficiency. He or she must also be fluent in Spanish. Finally, the interviewer must pass written and oral exams covering the NIS in Spanish, as well as pass an evaluation using mock interviews.

Interviews are conducted in a number of languages other than Spanish or English with the aid of an AT&T Language Line interpreter. The interviewers who contact these households are selected on the basis of experience, particularly their in-depth knowledge of the project, skill in determining a respondent's language, and ability to maintain control of the interview. Once selected, these interviewers receive additional training on how to work with the AT&T Language Line, as well as more extensive monitoring and feedback from supervisors.

4.5.3 Supervisory Staff

Interviewers at all sites are directly managed by trained supervisory staff during all hours of production. All supervisors undergo extensive training pertaining to the survey instrument, monitoring procedures, and systems necessary to carry out daily production. Supervisors are responsible for a variety of tasks designed to promote quality of data collection in the phone centers.

At the beginning of each shift, supervisors distribute interviewer staff according to project and management needs. Additionally, supervisors control access to the production rooms and CATI stations and coordinate break and lunch periods.

During all hours of production at least one supervisor is available to assist interviewers with questions concerning the survey, speak with respondents, or troubleshoot problems with CATI stations or equipment. All supervisors not responsible for the direct supervision of production are assigned to monitor telephone interviews and provide feedback to the interviewers based on the results of those monitoring sessions (described in Section 4.5.4). Supervisors rotate between monitoring and production on a daily basis.

In addition to direct production supervision and monitoring, each supervisor is responsible for tracking the production statistics and quality of a specific set of interviewers. Using predetermined quality standards, supervisors discuss production rates with each interviewer during bi-weekly feedback sessions. At this time, supervisors offer suggestions for improvement, commend excellent performance, and follow up with any necessary disciplinary or quality issues.

4.5.4 Monitoring

The interviewer monitoring program is at the core of quality control for the household data collection. Supervisory staff monitor interviews to ensure that all interviewers perform according to project specifications. For this purpose a supervisor can listen to the telephone call and view exactly the same CATI screens that the interviewer sees. Monitoring provides the basis for much of the feedback given to interviewers to improve their skills. In addition, it provides valuable information for improvements in the survey instrument.

The criteria used in monitoring interviewers are standardized: supervisors and interviewers share the same knowledge about the project and about procedures for resolving problems regarding individual survey items and CATI screens. This standardization minimizes interviewer differences. As new problems arise, information on how to respond to them is disseminated immediately to all staff involved.

Data collection is monitored for accuracy by supervisors, who track the interviewer's performance on individual survey items and on the entire interview. Monitoring activities include examination of the case disposition (the outcome of the interview), refusals to individual items as well as refusal aversion techniques used by the interviewer, respondent interruptions, general interviewer-respondent interactions, and the interviewer's knowledge of the CATI system. Performance measures are then assigned to each interviewer.

An automated system, specially developed by Abt, selects interviewers for monitoring based on experience and recent monitoring scores. Priority for selection is given in the following order: 1) interviewers who have recently completed training, 2) interviewers with below-

average monitoring scores, 3) the length of time since last monitoring session, and 4) the need to monitor at least 5% of total NIS interviewing work.

Interviewers are monitored for 20-minute sessions and are not aware that they are being monitored. After a monitoring session is complete, the supervisor reviews the session with the interviewer and discusses the areas for improvement and the interviewer's strong points. Sometimes supervisors record the interviewing session and use the tape to illustrate areas in need of improvement.

Supervisor Alerts

At the end of each interviewing shift, the supervisor on duty summarizes interviewer activities that occurred during the shift. Key monitoring information in these reports includes the percentage of interviewer hours monitored and any CATI problems that occurred. These reports are reviewed by NIS staff. This system allows problematic trends to be detected and corrected in a timely fashion, thus ensuring the quality of the work conducted in the telephone center.

Refresher Training

Each quarter, interviewers receive refresher training. These training sessions introduce any recent changes, such as new questions, wording changes, or policy changes. They also serve as an ideal venue to review any problem areas in the survey. Identical refresher training presentations are used in each telephone center to minimize differences among sites.

4.5.5 CATI System

The CATI system is the backbone of the household data collection and assists interviewers and NIS staff with collecting high-quality information. Its main quality-related functions are: 1) delivering cases, 2) ensuring that respondents are asked correct questions by following skip patterns, 3) recording responses, 4) tracking the status of each case and returning it to the interviewing queue when needed, and 5) assigning final disposition (or status) codes.

Delivering Cases

Research has established a pattern of call attempts for the NIS that maximizes response rates while minimizing costs (20). Dennis and co-workers (21) demonstrated that the optimal call patterns place the first call attempt on a Sunday through Thursday evening. If no contact is made during that attempt, a second attempt should be made on the following afternoon; and if contact has still not been established, the third attempt should be made on the evening of the second dial attempt. Abt has also developed protocols for sending each type of case (e.g., refusals, unspecified callbacks, ring no answers) to the telephone centers. For example, after a respondent refuses for the first time, the case is put on hold for three days before another call attempt is made. Typically, by the fifth attempt 69% of interviews are completed.

Logic Checking

The CATI system is programmed to prevent data entry errors and ensure data quality. When entry errors or logical errors occur, such as shot dates prior to the reported date of birth, a warning screen alerts the interviewer. The system will not allow the interviewer to pass the screen until the inconsistency has been corrected.

Look-up Databases

In the household interview the respondent is asked to report the names and addresses of the child's vaccination providers. To record this information as accurately as possible, the interviewer has access to a database of medical providers that includes their names and mailing addresses. The information in this database comes from prior NIS interviews and subsequent experience. If a provider appears in the database, the interviewer selects that provider for that child. If the interviewer is unable to find the provider in the database, the interviewer enters the name and address information given by the respondent. This information is then reviewed by locating clerks when the provider mailout is being prepared. (Chapter 5 gives a complete description of this process.) Using the provider database reduces the time of the interview and increases the accuracy of the provider information collected.

CATI System Testing

Whenever the CATI system is changed, whether by altering a question or modifying the program, it is tested by at least two project staff. Once the testers are satisfied that the system is working properly, the change is finalized. The changed system has a final live test on the first morning that calls are made using the updated version, and then the system is used for actual interviewing. If the interviewers discover any problems, they fill out a CATI Problem Sheet and forward it to project staff and programmers for corrections.

CATI System Monitoring

Because the CATI system's proper functioning is critical to the household data collection, it is monitored on several levels using computer programs and telephone center staff. The first level of monitoring is a program called Station Monitor. This program maintains a record of who executed what CATI component (e.g., NIS Q1/2000, NIS Q2/2000) and at what time. Interviewers must use Station Monitor to enter the CATI system. Station Monitor also determines which queues a particular interviewer may access such as the refusal queue, the Spanish language queue, or new sample that has just been released to the telephone center.

The NIS status application is the second level of monitoring. While the phone centers are open and collecting data, this program monitors NIS CATI studies. It automatically alerts the programming staff when any CATI application breaks down or falls behind. The three CATI applications monitored by the status application are one application that saves the data collected during each phone call and two applications that work together to deliver cases from the calling queue to the interviewers. These two status applications turn on queues according to time zone and deliver cases to the CATI work stations.

As part of their routine responsibilities, supervisors monitor interviewers while they are calling respondents. During these monitoring sessions, the CATI system is also monitored by the telephone center supervisors, who report any problems that may have been missed by the computer programs. CATI system monitoring also takes account of comments from interviewers, who report problems on CATI problem sheets. These forms are reviewed by supervisors and project staff, and adjustments are made to the case by the programming staff.

When any of the computer programs encounters an error, the error information, along with any data, is written to an error log. The CATI programmers review this log on a regular basis to make sure the errors that occur are usual system processing errors (e.g., an interviewer could not log into the network on the first try) and are not anything detrimental (e.g., data loss). If a computer station freezes during an interview, a program saves the data to that computer's hard drive, and the data are recovered when the station is brought back on line by the programming staff. Furthermore, programming staff are available on call during all hours that the telephone centers are open, to address any CATI system problems that arise.

Data Backup Procedures

The NIS household survey uses dual servers. That is, for each call that is made, that same call and any information collected are updated on a separate server to ensure that no data are lost through equipment failure. Each night the production data are copied from the production drive to the back-up production drive, the back-up server reporting drive, and the back-up server zip storage drive to avoid any loss of data. Additionally, there is also a nightly incremental server-to-tape backup. A full backup is performed once weekly. Both the daily and weekly backup tapes are shipped to a secure, off-site storage facility.

Toll-Free Number

The NIS also has a toll-free number for respondents to call if they want to ask any additional questions, speak with a supervisor, or complete the survey in response to an advance letter or an answering-machine message left during the RDD dialing. The toll-free line is directly connected to the CATI system, permitting interviewers to bring up the case from the phone number or the case ID number (indicated on the advance letter). The telephone interviewers answering the toll-free number are refusal converters, interviewers who have received additional training to handle more advanced calls. Interviews and information can be given in either English or Spanish.

4.5.6 Processing of Consent Forms

At the end of an interview, when an interviewer receives consent to contact a provider over the phone, he/she records the information in the CATI system. The next morning (or Monday if the interview occurred on a weekend), data production clerks print the child and provider data from the interview and check for errors (e.g., missing data, spelling errors). If corrections need to be made, the form is given back to the interviewer for editing (this occurs for about 7% of forms). When it is returned, it is examined again by the clerical staff. When the form meets approval, a package is prepared for mailing to the immunization provider.

4.5.7 RDD Production Reports

Daily production reports focus on many aspects of household data collection. There are three main types of reports. First, the key indicator tables show the status of the entire sample by replicate and IAP area. The tables count each case according to its current disposition code at the close of the telephone center the previous night. These dispositions, in turn, are categorized into several indicators (e.g., refusals, nonworking numbers or businesses). These reports present information that covers the range of issues related to the household data collection, including information on purely production-related issues to higher-level outcome issues, and thus give a comprehensive overview of the data.

The Q-Based reports are also run nightly. They look solely at cases that are not at a final disposition code, and group the cases into queues, which show which kind of interviewer will get the case. For example, one can see exactly how many cases are at a non-contact disposition and have yet to be finalized. They differentiate by time zone, so that the sample management team can determine what kinds of cases to release. The Q-Based reports also assist in staffing, as telephone center management can see exactly which interviewers are needed at which times. For example, if the Eastern time zone starts to become backlogged with pending Spanish cases, management can assign more bilingual interviewers during evening shifts. These reports also identify programming changes that might be necessary, such as how to adjust case delivery by time zone, the ratio of case delivery by shift or by fresh versus non-contact cases.

The Statistics Reports monitor production by site, shift, and quarter. These reports are used by the telephone center management to track production at all three sites, and to see basic information such as dials and completes during each shift. This information helps configure staffing to handle outstanding work on the sample.

These three sets of reports are available every morning on the NIS intranet. This arrangement enables staff to use a centralized QC system on a routine basis to monitor several quality control mechanisms, including interviewer monitoring, production reports, and adverse events, and to provide rapid feedback. Because they can compare these production reports, NIS staff are able more quickly to detect abnormalities in operations or data.

5. Provider Record Check Study

The provider record check component was added to the NIS in 1995 in order to improve the accuracy of the immunization estimates from household reports. The 1994 National Immunization Provider Record Check Study (NIPRCS), a part of the National Health Interview Survey (NHIS), showed that providers reported more-accurate vaccination histories than could be obtained from the household. The NIS conducted feasibility studies in late 1994 and began collecting provider data for all children in 1995. Using the provider-reported vaccination data to determine vaccination coverage levels has been found to improve the quality of the NIS data (22).

5.1 Instrument

The Immunization History Questionnaire (IHQ) used in the NIS to collect the child's immunization history from providers is consistent with the questionnaire used in the NHIS/NIPRCS study (23). The questionnaire is modified from time to time as new vaccines are added to the recommended schedule, and changes are tested and adopted that make it easier for providers to complete the form accurately.

5.2 Household-Reported Provider Data

As described in the preceding chapter, the household interview asks the respondent for the names and addresses of the child's immunization providers. The respondent (or someone who can authorize the release of medical records if the respondent does not have that authority) is asked for consent to contact these providers about the child's immunization history.

5.3 Collection of Provider Data

The PRCS follows the Total Design Method (24) to maximize provider participation. Each identified provider (whom the respondent gave consent to contact) is sent an initial mailing and then a reminder postcard two weeks after the original mailing. Providers who have not responded after five weeks are sent a second questionnaire, and those who have not responded by seven weeks after the initial mailing are prompted by telephone. These steps are described below. A tracking system supports all phases of data collection.

5.3.1 Consent Forms

Consent to contact providers is obtained from the child's parent or legal guardian during the household interview. After the interview, the name of the consent giver and the child's identifying information are printed on a consent form. Each consent form is reviewed for accuracy and then signed by the telephone interviewer who obtained consent (or by their supervisor). These consent forms are kept on file, and a copy is sent to each provider, along

with the IHQ. The consent form is essential for release of data from the child's medical record.

5.3.2 Locating Providers

The names and addresses of vaccination providers for each child are loaded into the tracking system. Every provider address is reviewed for completeness and accuracy. The locating clerks have access to the database of previous NIS providers, various web searches, and other provider directories. After the clerks have reviewed and edited the addresses, a supervisor reviews the addresses and approves them as usable for mailing. On average, 93% of the provider names and addresses are approved for mailing. (The remainder do not have enough information to complete the mailing address.) This system aids in a timely mail-out and ensures an accurate match, resulting in the necessary provider information to determine the child's vaccination status.

5.3.3 Mailing of the IHQ

The initial packet mailed to the provider includes a cover letter, a labeled questionnaire, a copy of the signed consent form, a copy of a recent article in *Morbidity and Mortality Weekly Report (MMWR)*, and a return envelope. The labels for the questionnaire and for the mailing envelope are generated automatically. Mailing labels are sorted by ZIP code, so that requests for multiple children with the same provider can be sent in a single envelope. Mailing addresses are also rechecked for completeness as the packets are assembled. A quality control check of each packet confirms that all of the materials are included in the envelope and that the questionnaire label matches the mailing label.

Postcards

Reminder postcards are sent two weeks after the initial mailing. The postcard reminds providers to return the questionnaire, and thanks them if they have already done so. These postcards do not contain any information that identifies the child.

Reminder Mailing

Five weeks after the initial mailing, follow-up packets are sent to providers who have not returned the initial questionnaire (40% of providers in 1999). The follow-up packet contains a new cover letter and another questionnaire for the child. The QC procedures are the same as in the initial mailing. These procedures follow the Total Quality Management model, which advocates continued contact using varying means to maximize response.

5.3.4 Telephone Prompting of Providers

Providers who have not responded seven weeks after the initial mailout are prompted by telephone to return the questionnaire. The tracking system automatically creates a list of the cases that need prompting, and gives the prompting interviewers a means to track the progress of the case. Most prompting calls result in a request to send another questionnaire, which is faxed or mailed to the provider's office. In a few cases, the interviewer may

actually collect the vaccination information during this phone call. Approximately 34% of providers required a telephone prompt in 1999.

In addition to the telephone prompting, providers can call a toll-free number with questions about the Immunization History Questionnaire or the study. The provider can complete the questionnaire by calling the provider hotline. However, less than 1% of cases with immunization data are completed by telephone.

5.4 Data Preparation

5.4.1 Form Edit by Clerks

As the IHQs are returned, they are reviewed by editing clerks for completeness and legibility, as a means to ensure that the data are entered accurately. If a form is totally unreadable or a faxed form appears to be missing a page (the IHQ is a one-page, two-sided form), the provider is recontacted and asked to complete another form. The clerks also check the individual responses. If a date is not written clearly, the clerk rewrites the date. Shot dates are zero-filled (e.g., 1 is changed to 01) to minimize data entry errors. If the provider did not fill out the IHQ, but instead sent a copy of the child's vaccination record, a clerk transcribes the record onto the IHQ form. A second clerk then reviews the transcription. If the provider filled out the form and attached records, the two are compared for consistency.

All edited forms are reviewed by a senior editor. The senior editor sets aside forms with certain types of discrepancies (such as shot dates before the date of birth of the child), forms where the provider reported a different date of birth than the household, and forms with an unusually high number of shots or any other irregularity. These forms are then reviewed by the project management staff. One study (13) examined a sample of 172 IHQ forms (from Q3/1998 to Q2/1999) that had shot dates prior to the date of birth. Of the 218 individual errors found on those forms, 81% arose from dates received on the IHQ, 12% came from mistakes in transcription by the editors, and the remaining 7% were made in data entry (Section 5.5).

In 1999 a computer program was developed to give the senior editor access to the household data to help resolve discrepancies and inconsistencies in provider-reported data on IHQs. The household data are used to help clarify an illegible date on the form (e.g, 4 vs 9), resolve shot dates before the date of birth, and identify forms that have been filled out with data for the wrong child.

The editing process improves quality by reducing the likelihood of data entry errors, resolving contradictory information, and ensuring that child-provider matches are accurate.

5.4.2 Review by CDC Staff

Prior to Quarter 2 of 1999 a random sample of 10% of the returned forms were photocopied after the close of the quarter and sent to a physician at the National Immunization Program for review. This review focused on: 1) comparing the date of birth given by the provider to the household date of birth, 2) checking shot dates to see whether they follow the recommended schedule, 3) comparing shot dates with birth dates, and 4) verifying that data from attached immunization records were transcribed correctly. The NIP staff physician worked closely with the data preparation staff to monitor the error rate for editing IHQs, and to reduce it to 1% or lower.

This review was replaced in Quarter 2 of 1999 by the matching-sheet review process described in the next chapter.

5.4.3 Data Retrieval Calls

If a provider returns a form that is illegible or has missing information, a data preparation clerk calls the provider and asks that the form be resent. Starting in Q3/1999, additional provider callbacks were instituted for cases in which the provider returned records that appeared to be for the wrong child. This determination is made by comparing the date of birth given by the provider with the date given by the household, and comparing the shot card data, when available, with the provider shot data. If the information is inconsistent, a telephone interviewer calls the person who filled out the form, confirms the date of birth for the child, sends a new form to the provider, and asks the provider to complete and return the new form. In the first three quarters after this procedure was instituted, a total of 124 cases were called back, and new immunization data that matched the household report were obtained for 23 cases.

5.5 Data Entry

5.5.1 Data Verification

Data from the returned IHQs are entered in a data file by a data entry contractor. The forms are keyed and verified. The first data entry operator enters the data from the form, and then a different operator (the "verify operator") enters the data a second time. If there is a discrepancy between the first and second entry, the verify operator checks the hard copy form and corrects the entry as necessary. According to the data entry contractor, the NIS information from providers has a data entry accuracy rate of 99.5% using this double-data-entry procedure.

5.5.2 Quarterly Determination of Data Entry Contractor Keying Error Rate

Starting in 1999, a procedure for verifying the discrepancy rate for data entry was instituted. A random sample of between 200 and 400 forms is resubmitted for data keying each quarter.

In the first quarter in which this procedure was implemented, a sample of 195 forms yielded 91 discrepancies among the potential 25,116 fields keyed, a rate of 0.36% (13). Most of these discrepancies were in fields that are not edited, such as verbatim responses to Other-specify questions and phone numbers.

5.6 Summary

Overall, the process for obtaining provider shot data on children located through the household interview includes numerous procedures to ensure that the provider sends information for the correct child, the information appears plausible (based on logic and other information), and the data are accurately entered for analysis. The need to obtain a high response rate is also integral to the entire data collection process, and regular follow-up is incorporated into the model. Altogether, these procedures and systems enable the NIS to accurately collect copious amounts of information from a busy and not always responsive target population.

6. Preparation of Data Files

After data collection ends, the NIS combines the data collected from household respondents and from vaccination providers in a comprehensive analytical file. It is important that this file be accurate and well-documented and have a user-friendly format. This chapter details the steps taken to produce final data files for the NIS.

6.1 Household Data File

6.1.1 Quarterly Cleaning, Editing, and File Construction

Even though the CATI system makes numerous checks during data collection, a final editing process identifies any remaining data inconsistencies and takes steps to reduce or eliminate them. Once the CATI production files are passed to the data preparation stage, various household- and child-level files are produced by extracting specific fields from the CATI data.

A master look-up database, constructed from the questionnaire, contains information about each field, including allowable ranges of responses. The master database is maintained, reviewed, and updated each quarter. At the end of each quarter of data collection the raw data are matched against this master database, and a report details any data anomalies. This report is then reviewed by a senior project analyst for resolution.

Data quality checks are performed periodically throughout each quarter. Data files from the CATI system are transmitted weekly to the data editing system, so that analysts can check distributions of responses and review cross-tabulations for internal consistency. Any problematic questions or other data anomalies are immediately brought to the attention of the project staff, who work with CATI programmers and with telephone interviewers to identify and correct the sources of the problems. Data management staff, project staff, and data collection staff work closely to ensure that the data are thoroughly scrutinized and problems are resolved quickly.

Other logic checks on the data include assuring consistency between files, eliminating duplicate records, reviewing call comments by the telephone interviewers, and measuring the completeness of the vaccination history section of each record. A vaccination history is considered complete if the following data have been collected:

- All children have a date of birth that qualifies them for the NIS survey. That is, the child
  was 19 to 35 months old when the household gave the child's date of birth.

- If shot records were used during the interview, each shot listed has a valid date.

- If shot records were not available, but the respondent answered that the child had received
  a specific antigen, then the number of shots (or "All") was recorded for that antigen.

6.1.2 Automated Back-Coding of Vaccines

The objective of the household interview is to collect types and dates of vaccinations given to children in the target age range. At times during data collection, the respondent may not recognize a vaccine name or may be unable to decipher the name of a vaccination written on a child's medical records. In addition, providers often abbreviate the name of a vaccine, such as OPV for Oral Polio Vaccine. The interviewer records the names of all vaccinations mentioned in an open-ended text screen, referred to as the "other shot verbatim" section, to reduce the chance of error in this process and to ensure that complete information is collected.

The NIS maintains a cumulative database of all vaccines mentioned in the "other shot verbatim" section. This list includes alternate spellings, abbreviations, and other variations of vaccine names; for instance, "Varicela," "Varricella," "VZV," and "chicken pox" are all variations of "Varicella." During data editing these verbatim responses are collected and checked against the database. Any verbatim responses not found in the existing database are sent for review and coding at the National Immunization Program. New entries and appropriate vaccine codes are then added to the database each quarter.

Once coded for the proper antigen, the verbatim-response information can be used to complete the child's immunization history. The back-coding process ensures accurate and usable data files.

6.1.3 Automated Back-Coding of Race and Hispanic Ethnicity

Respondents are allowed to select one or more race or ethnicity categories (from lists in the questionnaire), or they may give verbatim responses in the "other" category. These responses are back-coded following OMB guidelines (25). During data editing a similar process as described for vaccine back-coding is applied to the verbatim responses on race and Hispanic ethnicity, creating a cumulative database. Experts at NIP review all new verbatim responses encountered each quarter. In this way, any verbatim response can be included in specific categories and used to create race and ethnicity variables for analysis.

6.1.4 Creation of Composite Variables for Household Data File

A number of composite (constructed) variables are created and included in the household interview file. Such variables assist users and data analysts by providing readily accessible measures of key items in the data, thereby serving many of the quality control objectives of the project. Providing these constructed variables as a standard part of the data file eliminates error and duplication of effort and makes it easier to use.

The composite variables include the child's up-to-date status on vaccinations, age of child at each shot (if a shot card was used), race of child, race of mother, Hispanic origin of child, Hispanic origin of mother, participation in WIC (the Special Supplemental Nutrition Program for Women, Infants, and Children), occurrence of chicken pox disease, household income, poverty status, firstborn status, education of mother, and mobility status. Project staff develop specifications for each analysis variable. These specifications are then approved by NIP and NCHS. The approved specifications are given to a programmer, and the resulting values of the analysis variable are confirmed by a project analyst. A quarterly review considers what new composite variables may be needed.

6.2 Initial Processing of Provider Data File

During the data collection phase of the Provider Record Check Study, no attempt is made to identify and eliminate duplicate records. Duplicate records often result when a reminder call prompts a mail or fax return of a form that is already in the mail. All forms received are checked in and sent for data entry. After entry the data file is unduplicated to remove such records. The process is structured in a way that maintains the integrity of the child-provider pairs and ensures that all pertinent information for each child is kept intact.

Provider data are keyed into an ASCII file. For a discussion of the quality control efforts applied to the data entry system, see Section 5.6. Using a predetermined layout, an analysis file is made from the ASCII file.

Each field is subjected to a detailed check for out-of-range values and sources of missing data, to reduce the potential for error in the data entry process. Out-of-range data often indicate incorrect vaccination dates. Problem cases are pulled and reviewed manually by a senior analyst. Problems are resolved using all available data sources: household file information, provider data, partnering agency experts, and recommended vaccination schedules. As in the NIS household file editing process, the analyst helps to identify problems or trends in the data that can lead to the reduction of error.

Any problems that cannot be resolved remain in the file until the household data and the provider data are combined.

6.3 Combining the Household and Provider Data

After the separate editing and processing described in Sections 6.1 and 6.2.2, the household data and the provider data are combined to produce the NIS data files. This step includes a number of checks on the consistency of the two sources. The creation of the NIS data files also involves combining provider data when more than one provider has reported vaccination history information for a child.

6.3.1 Provider Edit Program

A data file of IHQs received for the latest quarter is assembled and prepared for processing by the NIS Provider Edit Program. Before this program is run, all "other shot" verbatims recorded on the IHQ are automatically backcoded into their proper vaccine category (e.g., DTP). An extensive Provider Backcoding Translation Table (PBTT), containing all verbatim responses from Q2/1994 to the current quarter, is maintained for this purpose. This step ensures that all valid vaccinations are properly counted.

The Provider Edit Program is the key tool for assessing the quality of provider vaccination history data. The program begins by checking for duplicate IHQs for a child from the same vaccination provider. If a duplicate IHQ is found, the program retains the IHQ containing the largest number of vaccinations, because it generally contains the most complete vaccination history for the child. The Provider Edit Program then proceeds to assemble a child-level file of IHQ information.

The dates of birth reported on the IHQs are compared with the date of birth reported by the household respondent. (If the day component of a date is missing, but the month and year are present, the 15^{th} of the month is imputed.) An initial best date of birth is assigned. If there are no differences between the dates of birth on the IHQs, that date of birth is used as the best date of birth. If the IHQs contain different dates of birth, the one that agrees with the household-reported date of birth is used as the best date of birth. A flag is set to indicate a date-of-birth discrepancy between IHQs. If neither agrees with the household date of birth, the IHQ date of birth closest to the household date of birth is used as the best date of birth. A flag is set to indicate that the IHQ and household dates of birth differed. The program compares the initial best date of birth with the date of the first hepatitis B vaccination, typically given at birth. If the initial best date of birth is after this vaccination date, the program checks whether one of the other dates of birth reported for the child is on or before this vaccination date. If such a date of birth is found, it is used as the initial best date of birth.

The Provider Edit Program assembles the IHQ-level vaccination dates into a child-level data record. IHQs that contain no vaccinations are removed from the child-level file. In assembling a child-level record, the program checks for vaccination dates that are less than one month apart. These dates are flagged in the data file. Vaccination dates before the best date of birth are also identified. Initial counts of the number of vaccinations received by the child are created, and the program then calculates a disposition code for each child. The disposition code summarizes the quality of the vaccination data available for the child and determines whether the child should be classified as having adequate provider data for estimation purposes. "Adequate provider data" means that the available IHQ information is sufficient to determine the up-to-date status of the child.

For children with fewer than three provider-reported hepatitis B vaccinations, the program checks to see whether the "Given at Birth" box was checked for the first dose of hepatitis B. If it was checked and the date of the birth dose of hepatitis B was not reported, the program

assigns the date of the birth dose for this vaccine. The date is defined as given from birth (i.e., 0 days) to the date of birth plus 7 days. If the household used a vaccination record to report vaccination dates, these dates are examined to see whether the date of the birth dose can be taken from this record. If it is not reported, a value is imputed from the distribution of dates for the birth dose of hepatitis B in the current NIS quarter. This procedure was implemented in the first quarter of 2000. For Q1/2000-Q4/2000 a total of 361 children had the date of the birth dose of hepatitis B assigned using the above procedures.

As noted above, the Provider Edit Program makes extensive use of flags to identify children that may have potential discrepancies between the IHQ information if more than one provider responds, and between the IHQ and household information. The program uses over 200 edit flags. Flags are set for vaccination dates that may be incorrectly recorded, for date-of-birth discrepancies, and for indications that the IHQ may have been filled out for the wrong child. Key edit flags include date-of-birth differences between the household and provider, date-of-birth differences between providers, vaccination dates reported before the date of birth, reported vaccination dates that are close together, vaccinations reported on the IHQ out of date sequence, and transpositions of the day and month of a vaccination.

6.3.2 Matching-Sheet Review

The next step in the processing of the provider data for a quarter reviews matching sheets for children with specific discrepancies between household and provider data and/or between provider IHQs. The matching sheet (Appendix 1) displays all household and provider vaccination dates, all unique dates of visits to providers, and other key information such as each recorded date of birth (DOB) and the disposition code assigned to the child. The matching sheets allow an editor to compare the household and provider vaccination reports for a child, and to assess the accuracy and completeness of the provider-reported vaccination dates.

Eight matching-sheet review groups (Table 6.1) are formed in a sequential manner, so that each child with one or more discrepancies has only one matching sheet printed for review. Symbols on the matching sheets identify the specific potential problem with the provider-reported vaccination dates. For example, a single matching sheet for a child in Group 5 would contain at least one vaccination date before the best date of birth, but it could also contain vaccination dates that are close together.

Each matching sheet is reviewed by an editor, who indicates required edits to vaccination dates and/or the best date of birth. Other outcomes include that an IHQ was filled out for the wrong child or that, based on the best date of birth, the child is not eligible for the NIS. The review process is facilitated when the household used a “shot card,” because then household-reported vaccination dates can be compared with the provider-reported dates. Appendix 1 contains the detailed edit rules and guidelines for the matching-sheet review process. This process was implemented in the third quarter of 1998, and the number and definitions of the

Table 6.1
Definition of Matching-Sheet Review Groups and Number of Matching Sheets Reviewed for Q1/2000-Q4/2000 (23,416 children had provider data.)

Group Number	Description	Number of Matching Sheets Reviewed
1	Provider and household DOB are different, and sex of child and/or name of child are different.	33
2	Provider and household DOB are different.	535
3	Multiple providers have submitted different DOBs (includes providers who left the DOB blank).	371
4	Hepatitis B date is before the best DOB.	234
5	Other vaccinations are recorded before the best date of birth.	360
6	Vaccination dates are close together (if one provider, $\leq$14 days; if more than one, $\leq$30 days).	1,255
7	Raw vaccination dates indicated by provider(s) are not in time order.	1,501
8	Vaccination dates, other than hepatitis B, between zero and 37 days from best DOB, or either MCV or varicella given 6 months from best date of birth.	71
	Total	4,360

Note: For Group 5 the total of 360 children for Q1-Q4/2000 includes, for the first three quarters, vaccinations for MCV and varicella given before 9 months of age. In the fourth quarter this group was moved to Group 8 and changed to before 6 months of age. The review of children with vaccination dates between zero and 37 days of life was first used in the fourth quarter.

groups have changed over time. For example, Group 8 was added in Q4/2000. Also, CDC staff review a 10% sample of the matching sheets for each quarter.

The results of the matching-sheet reviews guide the editing of the provider data. The edits typically involve modifying the day, month, or year component of a vaccination date. The implementation of the edits involves using a template file that lists the existing vaccinations for the child. The required edits are entered into the template file. After implementing the edits, a check is run to see whether a child for whom edits have been made still falls into one of the eight groups. If so, the accuracy of the edits is checked. Khare et al. (13) found that a small percentage of children in the NIS receive one or more data edits to correct problems associated with the provider-reported vaccination histories.

For Q1-Q4/2000 the matching sheet review edited one or more vaccination dates for 1,638 children. Of the children in Groups 1, 2, and 3, 69 were identified as having a returned IHQ

that was filled out for a different (i.e., wrong) child; the data from each such IHQ were deleted. Another 122 children were determined not to be age-eligible; those children were removed from the household and provider data files. The procedure used to impute a missing date for the dose of hepatitis B given at birth assigned dates for 361 children.

The matching sheet review is not able to resolve all discrepancies. Also, providers are not recontacted to resolve discrepancies that may remain after the matching sheet review is completed. Among the 22,958 children with adequate provider data, no children had vaccination dates before the best date of birth, 86 (0.4%) had vaccination dates less than or equal to 14 days apart, and 38 (0.2%) had vaccination dates (other than for hepatitis B) from 0 to 37 days of life.

6.3.3 Construction of Variables from Provider Data

After completing all the edits, the Provider Edit Program calculates the number of doses received for each vaccine and creates both vaccination-date arrays and age-at-vaccination arrays. Additional up-to-date variables are created for various sentinel ages. These include variables indicating whether the child had received a specific number of doses of a vaccine by a certain age in months (e.g., 3 or more polio vaccinations by age 13 months). To accommodate the large and continually growing number of types of vaccinations collected in the NIS, vaccination-type indicator variables (Appendix 2) are also created from information recorded by providers on the Immunization History Questionnaire. For example, the vaccination-type indicator variable for the first dose of DTP indicates whether that dose was a DTP, DTaP, DT, DTP-Hib, or DTaP-Hib vaccination. Additional codes are included to cover the situation where the type of DTP or type of DTP-Hib vaccine is not specified by the provider. The vaccination-type indicator variables greatly reduce the number of vaccination-date and age-at-vaccination arrays that must be carried in the file without any loss of information. They also allow all data users to easily and consistently determine the specific type of vaccine given at each dose. The vaccination-type indicator variables were implemented in the first quarter of 2000. Other composite variables are also created at this stage. For example, a variable summarizes the type of facility that vaccinated the child.

6.3.4 Data File Construction

Three main data files are created for each 4-quarter time period.

The first file is the Internal Analytic Interview File with Provider Data. This file (for internal use only) contains all household and provider variables, including date-of-birth information and vaccination dates.

The second file is the Child-Level Analysis File (CLAF), an internal file that contains a subset of the variables included in the first file, including vaccination dates. In addition, the CLAF contains several vaccination variables (e.g., age in days at each vaccination), not included in the first data file, that NIP uses in analyses.

The third data file is the Public-Use File (PUF), which will ultimately be released for each calendar year of NIS data. A codebook, data user's guide, and other related documents accompany this file. The variables included in the PUF are a subset of the CLAF variables. Variables such as the date of birth and vaccination dates are not included in the PUF, in order to minimize disclosure risk.

7. Weighting and Variance Estimation

This chapter briefly summarizes the main steps in calculating the sampling weights, for each child with a completed household interview and for each child with adequate provider data, because appropriate and accurate weights are an essential part of the NIS data files. It also indicates how to estimate variances for estimates derived from NIS data. More-detailed discussions of these aspects appear in the annual *NIS Methodology Reports.*

Beginning with the second quarter of 1994, the NIS has conducted independent quarterly surveys in each of the 78 IAP areas. This design has made it possible to provide 4-quarter estimates of vaccination coverage levels for several antigens (diphtheria and tetanus toxoids and pertussis vaccine [DTP], poliovirus vaccine [polio], measles-containing vaccine [MCV], *Haemophilus influenzae* type b vaccine [Hib], hepatitis B vaccine [Hep B], and varicella vaccine) within each of the 78 IAP areas with an acceptable degree of precision (a coefficient of variation of no more than 5%).

The target sample size of completed telephone interviews in each IAP area is designed to yield approximately equal numbers of children with provider-reported vaccination histories in all 78 IAP areas. This is accomplished by assigning each IAP area to one of four categories according to its historical percentage of children with adequate provider data. In the categories where a lower percentage of children have adequate provider data, each IAP area has a larger target number of household interviews. In the categories where a higher percentage of children have adequate provider data, each IAP area has a smaller target number of household interviews. The approach of setting different targets for completed household interviews in the four categories makes it possible to come much closer to achieving equal numbers of children with adequate provider data across the 78 IAP areas than if equal numbers of household interviews were completed in all IAP areas.

The NIS makes a major effort to reduce nonsampling errors and to improve the quality of the estimates of vaccination coverage. A key concern is nonsampling errors arising from nonresponse and from the noncoverage of nontelephone households in the RDD survey. Nonresponse occurs mainly at two stages in the NIS. First, some households with age-eligible children do not complete the household interview. Second, for various reasons, some children with completed household interviews do not have adequate provider data. Noncoverage is a concern because about 10% of U.S. children reside in nontelephone households, with considerable variation among the 78 IAP areas.

The data are weighted to reduce potential biases and to allow population-based estimates of the percentage of children who are up-to-date on the key vaccinations and vaccination series. The process involves the calculation of two key weights. First, a child-level weight is assigned to each child for whom the NIS telephone interview is completed. Adjustments applied to those weights yield the second child-level weight, for each child with adequate provider data (i.e., other children do not have the second weight). Those children, with the

second weight, form the basis for the official estimates of vaccination coverage. The calculation of the weights and the development of vaccination coverage estimates are discussed in this chapter. Zell et al. (11) and Smith et al. (12) provide an overview of the calculation of weights in the NIS.

7.1 Imputation for Missing Information

Completed interviews may still not have data for some items. Before weights are calculated, imputation for item nonresponse takes place for the children in each individual quarter. For demographic and other variables used in the weight calculations (e.g., number of voice-use telephone lines in the household, maternal education, age category of child, and race/ethnicity of the mother), the hot-deck method (26) is used to impute missing values. The process uses family income to form imputation cells within IAP area. Item nonresponse rates in the NIS for demographic variables are typically very low. For example, age of mother is missing for 1.9% of the children in Q1-Q4/1999. Overall, in the public-use file for one year, a total of nearly 14,000 data items were imputed, but they constituted only 0.08% of all data items in the file.

7.2 Base Sampling Weights

For each individual quarter, a base sampling weight is assigned to the children for whom the telephone interview was completed. The IAP areas vary considerably on the total number of telephone numbers in the list-assisted sampling frame of working banks of 100 consecutive telephone numbers. The base sampling weight in an IAP area accounts for this variation by taking the ratio of the total number of telephone numbers in the sampling frame to the total number of sample telephone numbers in the replicates that were released for interviewing.

7.3 Adjustment for Multiple Telephone Lines in the Household

In a random-digit-dialing survey, telephone numbers are sampled. The actual interviews in the NIS, however, are conducted with households that have one or more children aged 19 to 35 months. Thus, a household with two telephone lines (i.e., two different telephone numbers) has a greater probability of being selected than a household with only one telephone line. To adjust for this unequal probability of selection, the base sampling weight is divided by the number of voice-use telephone lines in the household, up to a maximum of 3.

7.4 Adjustment for Unit Nonresponse in the RDD Survey

The RDD survey encounters unit nonresponse at three levels. First, for some sample telephone numbers, no determination is ever made as to whether they are nonworking, nonresidential, or residential (i.e., they remain unresolved). Second, for some sample numbers that are determined to be residential, no determination is ever made as to whether

the household contains one or more children age 19 to 35 months. Third, for some households with one or more children age 19 to 35 months, the immunization interview is never completed for the child. A three-level adjustment for unit nonresponse is used for each quarter. These adjustments use weighting cells within IAP area. In each weighting cell the weights of children for whom the telephone interview was completed are increased to compensate for unit nonresponse.

7.4.1 Formation of Weighting Cells

The sample telephone numbers in an IAP area are sorted into weighting cells based on: 1) residential directory-listed status of the telephone number and 2) demographic and socioeconomic characteristics of the exchange from which the telephone number was sampled. The demographic and socioeconomic characteristics are available with the GENESYS sampling frame. An example of such a characteristic is the percentage of persons aged 25+ years who have graduated from college.

For Q2/1994 to Q4/1994, the unit nonresponse adjustments were applied within IAP areas for categories formed by the residential directory-listed status and area code of the sample telephone number. Some IAP areas, however, contained only a single area code for a large geographic area. This situation reduces the effectiveness of area code for unit nonresponse adjustment. A different set of categories has been used since Q1/1995.

The new categories for each IAP area combine the residential directory-listed status of the sample telephone number and one or more characteristics of the telephone exchange (i.e., the combination of area code and central-office code). The characteristics of telephone exchanges examined in each IAP area were:

- Metropolitan Statistical Area (MSA) status,
- Percentage of households that are owner-occupied,
- Percentage of the adult population that are college graduates, and
- Percentage of the population that is white.

Variables that exhibited little variation were eliminated from further consideration for that IAP area. At least two telephone-exchange categories were selected for each IAP area. Thus, the minimum number of categories in an IAP area is four, because the two telephone-exchange categories are crossed with residential directory-listed status.

For each nonresponse-adjustment category in an IAP area, the actual number of children in the quarter for whom interviews were completed is determined. If the number of children is less than 10, a collapsing of categories takes place. The priority for collapsing is as follows:

1. Percentage of the adult population that are college graduates,
2. Percentage of households that are owner-occupied,
3. Percentage of the population that is white, and

4. Metropolitan Statistical Area status.

7.4.2 First-Level Adjustment for Unit Nonresponse

The nonresponse adjustment for unresolved telephone numbers uses secondary information from calling local telephone company business offices in order to estimate the proportion of unresolved telephone numbers that are residential numbers (27). Unresolved telephone numbers are divided into four categories:

- Noncontact numbers,
- Answering machines where the residential status of the number is unknown,
- Refusals at the introduction where residential status has not yet been determined, and
- Broken appointments at the introduction where household status was not determined.

In the second quarter of 1994 a large sample of unresolved numbers was drawn, and calls were made to local telephone company business offices to determine the status of those numbers. The sample was stratified into IAP area groupings based on telephone company service areas. The resulting proportions of residential numbers among unresolved numbers are used in the nonresponse adjustment.

7.4.3 Second-Level Adjustment for Unit Nonresponse

The sample of households that complete the screening interview in each weighting cell is used to adjust for households that did not complete the screening interview (so that it is unknown whether they contain any age-eligible children). The mean number of age-eligible children in the households that complete the screener in a weighting cell is used to estimate the number of age-eligible children in the households that do not complete the screener. This information is used to form a nonresponse adjustment that compensates for nonresponse at the screening stage of the NIS.

7.4.4 Third-Level Adjustment for Unit Nonresponse

The third adjustment for unit nonresponse compensates for households containing age-eligible children that do not complete the NIS interview. Within each weighting cell the total number of age-eligible children in the households that do not complete the interview is used to adjust the weights of the children for whom an interview is completed.

7.5 Poststratification and Compensation for Noncoverage of Children in Nontelephone Households

After calculating the nonresponse-adjusted base sampling weights for each quarter, the four quarters are combined for the remaining weight calculations, because the estimates from the NIS are for four-quarter time periods. (Because the base sampling weights use the total number of telephone numbers in the sampling frame, each child's weight is divided by 4

when the data from four quarters are combined.) Even with the adjustments for unit nonresponse, the NIS may still be subject to nonsampling error from undercoverage, unit-nonresponse bias, and other sources. Poststratification to Vital Statistics control totals is used to further reduce any potential bias in the telephone survey. Briefly, within each of a set of cells, poststratification applies a multiplicative adjustment to the sampling weight of each child, so that the total of the adjusted sampling weights equals the control total. The NCHS Natality file for the appropriate birth cohort is the starting point for the control totals. For each IAP area it contains information on maternal education, race/ethnicity of the mother, and the current age category of the child. The Natality file cannot be used directly to furnish population control totals, because some children die before they reach 19 months of age, some children are born outside the U.S., and some children move to a different IAP area from where the mother resided at the birth of the child. Information on infant mortality from Vital Statistics and on foreign births and geographic mobility from the most recent decennial census is used to adjust the Natality file control totals in each IAP area. Poststratification, with cells that combine maternal education, race/ethnicity of the mother, and age category of the child, is used in each IAP area to obtain a *simple-poststratification weight*, which further adjusts for potential nonsampling error in the RDD survey.

In the U.S. around 10% of young children reside in households without a telephone. This proportion, however, varies considerably among the 78 IAP areas. Bias from the exclusion of children in nontelephone households is a concern because the National Health Interview Survey indicates that they are less likely to be up-to-date on their vaccinations. Fortunately, the ratio of vaccination coverage for nontelephone to telephone children is available from the National Health Interview Survey, and telephone coverage rates are available from the Current Population Survey. Using this information, each poststratification cell is split into two subcells, for children who are 4:3:1:3 up-to-date versus not 4:3:1:3 up-to-date. The weights of the children are separately poststratified in each subcell. This approach, known as *modified poststratification*, is designed to reduce potential biases from the exclusion of nontelephone children from the NIS. Battaglia et al. (28) give details for this specific adjustment.

7.6 Children with Adequate Provider Data

The NIS attempts to obtain provider-reported vaccination histories for all children for whom the telephone interview was completed. Adequate vaccination histories are typically obtained for 65-70% of children. The remaining 30-35% of the children either do not have consent to contact providers, or have no provider-reported vaccination histories, or have histories that are too incomplete to determine whether the child is up-to-date. The IAP areas vary on the percentage of children who do not have adequate provider data. The NIS uses children with adequate provider data to form the vaccination coverage estimates for the IAP areas, the states, and the U.S. The methods used to adjust the modified-poststratification weights of children with adequate provider data are discussed next.

7.7 Adjustment for Provider Nonresponse

A response-propensity model is used to form weighting classes in each IAP area to help ensure the accuracy of vaccination coverage estimates by reducing the potential for bias from not including children without adequate vaccination histories. A national response-propensity model uses logistic regression to regress whether the child has adequate provider data on demographic, socioeconomic, and household-reported vaccination characteristics. For each child for whom the telephone interview was completed, the model yields a predicted probability of having adequate provider data.

The predicted probabilities are sorted in ascending order in each IAP area, and five equal-sized weighting classes are formed. Within each weighting class, the modified-poststratification weights of children with adequate provider data are divided by the weighted proportion of children in that weighting class with adequate provider data. These nonresponse-adjusted weights resulting from the response-propensity model are then raked to various control totals to produce the *final nonresponse-adjusted weights.* These weights help to compensate for potential biases from the exclusion of children without adequate provider data. Smith et al. (29,30) present the details of these methods.

7.8 Sampling Weights in the NIS Public-Use File

As noted above, two key weights are calculated for the NIS. The NIS PUF contains all children in a four-quarter time period for whom the telephone interview was completed. For these children, estimates of household characteristics are computed using the modified-poststratification weights. Examples of characteristics that can be estimated using the modified-poststratification weights include the proportion of children for whom vaccination records were used during the telephone interview, the proportion of children living below poverty, and the proportion of children who have had chicken pox.

The official estimates of vaccination coverage use the data from children with adequate provider data and the final nonresponse-adjusted weights. For these children the final nonresponse-adjusted weight is used to form estimates of vaccination coverage for the IAP areas, the states, and the U.S. For the children without adequate provider data, this weight is blank. As discussed below, to avoid the possibility of bias, it is important that the weights be used when estimates are computed from the NIS.

Estimation

An estimate of vaccination coverage for an IAP area can be expressed as a ratio. The numerator consists of the weighted number of children with adequate provider data who are up-to-date on a vaccine or vaccine series (e.g., they have four or more DTP vaccinations). The denominator consists of the weighted number of children with adequate provider data. The final nonresponse-adjusted weights must be used to obtain these weighted totals. Because the IAP areas form the strata of the NIS sample design, estimates for states

containing two or more IAP areas and national estimates can be obtained by summing the IAP-area numerators and denominators and then dividing. Statistical software packages such as SAS, SPSS, STATA, and S-PLUS can be used to obtain the weighted estimates of vaccination coverage (29,30,31).

7.9 Variance Estimation

The NIS sample design is a stratified one-stage cluster sample. The strata are the 78 IAP areas. The household is the primary sampling unit, and all children age 19 to 35 months in the household are eligible for the telephone interview. The children have unequal weights, arising from the sample design and the various weight adjustments used to account for nonresponse and for noncoverage of children living in households without telephones. Statistical software programs that assume simple random sampling will therefore most often compute standard errors that are too low.

Standard errors for the NIS can be obtained using the Taylor-series-approximation method, available in software such as SUDAAN, SAS, and STATA. The IAP area should be identified as the stratum variable, and the household should be identified as the primary sampling unit (PSU). The simplifying assumption that PSUs have been sampled with replacement allows these programs to calculate Taylor-series standard errors in a straightforward way.

Other variance estimation procedures are also applicable to the NIS. Specifically, the jackknife method using replicate weights and the bootstrap resampling method using replicate weights can also be used (via software such as SUDAAN and WESVAR) to obtain standard errors that fully reflect the impact of the weighting adjustments on standard errors. Smith et al. (31) discuss the calculation of standard errors in the NIS using the Taylor-series and jackknife methods.

8. Evaluation

Methodological evaluations to assess and improve the quality of the immunization coverage estimates are an integral part of the NIS program. The methodological areas covered include measuring and reducing:

- error in the survey from lower than expected eligibility rates
- noncoverage error
- household and provider nonresponse
- measurement error.

The NIS Methodology Task Force continually investigates and documents the levels of nonsampling error, and proposes alternative methods and procedures to reduce the effects of these errors. The Task Force also oversees work to improve the efficiency of the survey, such as investigating more-efficient calling patterns and optimizing the number of call attempts.

8.1 Steps to Evaluate Data Quality

8.1.1 Research and Evaluation Protocol

The NIS uses sound survey research methods. When seeking improvements in data quality and efficiency, a thorough research and evaluation protocol is employed. This is especially true where a suggested alternative or enhancement can have a marked impact on standard indicators such as household and provider response rates, eligibility rates, and the percentage of children who have usable provider data. The following protocol ensures that appropriate steps are taken to maintain data quality. These include ensuring that multiple approaches are considered, that effects are tested and measured, and that results for the NIS and for the survey research community are fully documented. Adherence to the research and evaluation protocol means that:

- Issues are identified, and proposals are developed that document alternatives under consideration.
- A thorough review of alternatives may solicit advice from experts in survey research methods and from subject-matter experts.
- Proposed studies are submitted to institutional review boards and confidentiality officers for review and approval.
- Cognitive testing with a recruited sample of respondents is often conducted to develop or finalize survey materials.
- Pretesting is conducted to evaluate and resolve operational and functional issues.

- Field testing with a larger, scientific sample is conducted to measure effects and test hypotheses.
- Test data and final data are evaluated, documented, and published.

8.1.2 Comparisons with External Data Sources

To help validate early results from the NIS, demographic and immunization estimates were compared to results from other sources such as:

- The National Health Interview Survey (NHIS)
- The Current Population Survey (CPS)
- 1990 decennial Census
- The National Immunization Provider Record Check Study (NIPRCS)
- Vital Statistics

Comparisons of NIS data with other methodologically sound studies were particularly important because the NIS contacts households via a random-digit-dialed survey. Substantial effort has been expended through the years to examine noncoverage of nontelephone households and develop methods that compensate for potential biases.

In the more recent years of the NIS, comparisons have been made to ensure that NIS results remain consistent with other studies. Results from these comparisons also added to knowledge about the reliability of other estimates (e.g., income and participation in the WIC Program). Estimates from the CPS, NHIS and Natality files are used in NIS weighting to bring sample estimates in line with demographic distributions from those sources (this has helped to reduce the mean squared error in the NIS).

NIS immunization rates also have been compared on a less formal basis with results from several state-sponsored immunization surveys.

Individual and Household Estimates
The characteristics examined between the NIS and other sources have included:

- Race of child
- Ethnicity of child
- Age of child
- Mother's education
- Mother's marital status
- Immunization rates for children 19 to 35 months of age

- Telephone coverage
- Eligibility rate
- WIC participation

Many comparisons have been made at the IAP-area level when the data were available at that level from the external source. The rest of this section summarizes recent comparisons with other surveys and with censuses.

Sociodemographic data

Table 8.1 shows a comparison of NIS data for the 4-quarter period Q1/1999-Q4/1999 with birth records from the Natality files for the appropriate birth cohorts. The Natality data are reasonably current and generally comparable. Weighted NIS percentages are about the same as the Natality data for race/ethnicity, education of mother, and age of child. Income is not available on the Natality files.

Table 8.1
Comparison of Demographic Characteristics of NIS Sample with Other Data Sources, NIS, Q1/1999 – Q4/1999—*U.S. Total*

Demographic Characteristic	NIS Unweighted (%)	NIS Weighted (%)[a]	Birth Records (%)
Race/ethnicity			
Hispanic	16	19	18
Non-Hispanic Black	17	15	15
Non-Hispanic Other	67	66	66
Education of mother			
12 years	44	54	55
13+ years	56	46	45
Age of child			
19-25 months	42	42	42
26-35 months	58	58	58

[a] Modified-poststratification weight used to form weighted NIS percentages.

Immunization data

National estimates of vaccination coverage for 1999 from the NIS and the National Health Interview Survey/National Immunization Provider Record Check Study (NHIS/NIPRCS) have been compared. As shown in Table 8.2, fairly large differences in vaccination coverage exist between the NIS and NHIS household reports, especially when the household reports from recall (i.e., no shot card). The NIS provider-based estimates of vaccination coverage,

Table 8.2
Comparison of Weighted Estimates of Vaccination Coverage from the 1999 NIS and the 1999 NHIS/NIPRCS[a]

		Percentage Up-to-Date				
	N	4:3:1:3[d]	4+ DTP	3+ Polio	1+ MCV	3+ Hib
Total						
NHIS, Household Report	1,793	69.8	80.0	85.9	91.5	80.9
NIS, Household Report	34,442	51.3	64.1	72.2	91.6	68.4
NHIS/NIPRCS, Best Values[b]	1,069	80.0	84.6	91.1	92.2	93.9
NIS, Provider Data	22,521	78.4	83.3	90.0	91.7	93.5
Shot Card[c]						
NHIS, Household Report	596	57.0	71.2	82.3	92.4	77.3
NIS, Household Report	16,829	56.6	70.4	85.6	90.5	79.4
NHIS/NIPRCS, Best Values[b]	470	87.2	91.8	98.1	95.5	96.8
NIS, Provider Data	11,964	81.0	86.5	91.3	93.6	94.4
Recall						
NHIS, Household Report	1,197	77.4	84.9	87.9	90.9	83.0
NIS, Household Report	17,613	41.0	55.9	55.2	93.1	51.2
NHIS/NIPRCS, Best Values[b]	599	76.8	81.3	87.9	90.8	92.6
NIS, Provider Data	10,557	75.6	80.1	87.8	89.8	92.6

a National Health Interview Survey/National Immunization Provider Record Check Study (a face-to-face household survey).
b Best Values reflect reconciliation of household and provider vaccination reports and supplement provider-reported vaccination dates with vaccination dates that appear in the household shot card report.
c "Shot Card" indicates those in the NHIS who reported from a shot card only and those in the NIS whose report on DTP was based only on a shot card.
d Series complete based on 4+ DTP, 3+ Polio, 1+ MCV, and 3+ Hib.

however, are generally close to the NHIS/NIPRCS "best value" estimates. The "best value" estimates reflect the results of an effort to reconcile discrepancies between the provider reports and household reports, and they also supplement the provider vaccination histories with vaccination dates that appear in the household shot card report. This additional step accounts for the somewhat higher estimates among NHIS/NIPRCS children for whom a shot card was used during the interview than among the corresponding NIS children. However, the estimates for the total sample are very close because a much lower proportion of NHIS interviews are conducted using a shot card, whereas roughly half of the NIS interviews are conducted using a shot card.

Additional analyses have compared NIS immunization estimates with NHIS/NIPRCS estimates for different estimation approaches (32). This study assessed whether reconciliation efforts between household and provider data and recontacting nonrespondents in the NIS (and use of the resulting "best values") would lead to greater precision in the estimates. The findings suggest that the full best-value approach for producing immunization

estimates in the NHIS/NIPRCS is not feasible for the NIS because of the large sample size and the need for timely estimates. However, the use of some of those procedures in the NIS to produce best values from the household and provider data (without reconciliation and recontacting nonrespondents) can produce reasonably accurate best-value shot dates. The analyses showed that, for 1999 NIS data, use of a limited best-value approach would increase the national estimate of 4:3:1:3 coverage by 3.5 percentage points, from 78.4% to 81.9%.

Telephone Coverage

For the percentage of telephone households with a child in the 19-35 month age range, Table 8.3 compares the value observed in the NIS (for 1999) with the expected value derived from Census data and adjusted for changes in the number of births and the number of households. The negative differences indicate that the NIS has some undercoverage of the population aged 19-35 months (the median difference is around –0.8 percentage point). However, as discussed in Chapter 7, the final weights incorporate a number of adjustments, aimed at reducing the potential bias from noncoverage of nontelephone households and from other coverage errors. The success of these weighting adjustments and the minimal level of bias are evident in Table 8.2, which shows that the NIS estimates of vaccination coverage are comparable to NHIS estimates at the national level. The NHIS directly surveys telephone and nontelephone households and obtains slightly higher response rates.

The eligibility rate in the NIS (the percentage of households with a child 19 to 35 months of age) is closely monitored. Maintaining the eligibility rate is important because it greatly affects the cost of the survey (screening) and the precision of the survey estimates.

8.2 Summary of Evaluation Results and Implementation

To reduce the effects of nonsampling errors such as noncoverage, nonresponse, and measurement errors, the NIS utilizes a variety of techniques during data collection, editing and cleaning of the data, and the estimation process. These have been described in the preceding chapters. Special studies also examine various types of nonsampling errors and develop strategies to minimize their impacts on the NIS estimates. Below is a summary of major research and evaluation efforts to investigate, measure, and identify ways to reduce the effect of nonsampling errors and to improve the quality of NIS data.

8.2.1 Noncoverage Error

Much of the work on noncoverage in the NIS has focused on nontelephone households. NIS staff have developed broadly useful techniques to compensate for noncoverage of nontelephone households. For example, they have used information from households that had interrupted telephone service (or at one time had telephone service) to adjust estimates to represent the portion of the population that never had telephone service (33,34). Their research continues and may lead to the adoption of this method to reduce noncoverage bias in future NIS estimates.

Table 8.3
Observed Percentage of Households with a Child 19-35 Months Old in the NIS Compared with the Expected Percentage, 1999

IAP Area	Observed Percent	Difference from Expected	IAP Area	Observed Percent	Difference from Expected
U.S. Total	3.71	-0.89	Orleans Parish	3.20	-1.07
Alabama			Maine	3.12	-1.15
Rest of State	3.80	-0.10	Maryland		
Jefferson County	3.59	-0.73	Rest of State	3.65	-1.27
Alaska	5.31	-1.45	City of Baltimore	3.08	-1.23
Arizona			Massachusetts		
Rest of State	3.60	-0.98	Rest of State	3.45	-0.91
Maricopa County	4.23	-0.56	City of Boston	3.16	-0.56
Arkansas	3.47	-0.60	Michigan		
California			Rest of State	4.11	-0.47
Rest of State	4.15	-1.04	City of Detroit	4.18	-1.34
Los Angeles County	4.08	-1.74	Minnesota	3.94	-0.76
Santa Clara County	4.20	-0.87	Mississippi	3.71	-0.68
San Diego County	3.89	-1.26	Missouri	3.30	-0.95
Colorado	3.81	-0.64	Montana	3.56	-1.07
Connecticut	3.42	-1.20	Nebraska	3.89	-0.46
Delaware	3.88	-0.79	Nevada	4.48	-0.03
District of Columbia	2.58	-0.44	New Hampshire	3.28	-1.70
Florida			New Jersey		
Rest of State	2.77	-0.80	Rest of State	3.40	-1.08
Duval County	3.84	-0.92	City of Newark	4.22	-0.14
Dade County	3.43	-1.02	New Mexico	4.02	-0.76
Georgia			New York		
Rest of State	4.15	-0.51	Rest of State	3.61	-0.99
Fulton/DeKalb Counties	3.37	-1.11	New York City	3.50	-0.60
Hawaii	4.22	-1.33	North Carolina	3.61	-0.57
Idaho	4.74	-0.18	North Dakota	3.92	-0.39
Illinois			Ohio		
Rest of State	3.63	-0.96	Rest of State	3.45	-0.81
City of Chicago	3.67	-0.92	Cuyahoga County	3.59	-0.44
Indiana			Franklin County	3.55	-1.13
Rest of State	3.88	-0.29	Oklahoma	3.30	-0.66
Marion County	3.85	-0.67	Oregon	3.58	-0.72
Iowa	3.64	-0.72	Pennsylvania		
Kansas	3.72	-0.71	Rest of State	3.69	-0.38
Kentucky	3.60	-0.24	Philadelphia County	3.15	-1.11
Louisiana			Rhode Island	3.33	-1.02
Rest of State	4.07	-0.78	South Carolina	3.50	-1.03

Table 8.3
Observed Percentage of Households with a Child 19-35 Months Old in the NIS Compared with the Expected Percentage, 1999

IAP Area	Observed Percent	Difference from Expected	IAP Area	Observed Percent	Difference from Expected
South Dakota	4.10	-0.40	Utah	7.01	-0.59
Tennessee			Vermont	3.56	-1.06
Rest of State	3.51	-0.05	Virginia	3.46	-0.95
Shelby County	3.82	-1.60	Washington		
Davidson County	3.42	-0.65	Rest of State	3.73	-1.10
Texas			King County	3.19	-0.99
Rest of State	4.19	-0.71	West Virginia	2.91	-0.07
Dallas County	4.35	-0.64	Wisconsin		
El Paso County	5.80	-0.38	Rest of State	3.53	-1.04
City of Houston	4.56	-0.25	Milwaukee County	4.00	-0.26
Bexar County	4.48	-0.86	Wyoming	3.82	-0.88

Earlier work on noncoverage problems associated with RDD surveys pertained to building a capability to distinguish business, nonworking and residential numbers. This distinction enabled creation of a better list frame from which to sample phone numbers, which improved the efficiency of call screening (35).

8.2.2 Nonresponse Error

In efforts to increase response rates, NIS staff have conducted research to improve the positive impact of advance letters to households (36). In this research three versions of the advance letter were mailed to subsamples of NIS households. The version of the letter that prominently described the purpose of the NIS had the highest cooperation rates, lowest refusal rates, and highest proportion of cases reaching a final disposition at the initial household contact. This version of the letter was later adopted for the full NIS sample.

Research has also been conducted on the effects of alternative question wording, alternative introductions and screeners, and alternative IHQ forms.

NIS staff conducted analyses to determine the patterns of telephone contact attempts that result in the highest contact rates and response rates (21). This research examined the outcomes of over 9.5 million call attempts from the NIS in 1996 and 1997, along with Census data for telephone exchanges. The research built on a previous study, whose results were used in the initial programming of the NIS call-scheduling system (37). Results of the latest study showed that 3-call patterns in which the first call was placed on Sunday-Thursday evening, the second call on the following afternoon, and the third call on the evening of the second call are optimal. Analyses of associated Census data resulted in variations of this pattern that could be incorporated into a revised call-scheduling program.

Abt Associates opened a new telephone center in the Pacific time zone to provide more calling capacity and enable additional evening calling.

Further research in this area focused on the number of call attempts, to determine the point at which there are diminishing returns in making additional calls (38). Preliminary results from this analysis suggest that a number of call attempts greater than 12 does not reduce bias enough to justify the cost of making those calls. Further work is planned in this area before actual changes to the NIS procedures are considered.

8.2.3 Measurement Error

Early research on the NIS indicated that the cross-sectional sample design, rather than a panel design, would produce more-accurate estimates of vaccination coverage (39). This research showed that a panel-conditioning effect would bias estimates of vaccination coverage through social desirability effects and because the survey measurement itself would influence immunization behavior.

Much research has shown that household-based data, even from those that used shot cards, were subject to large measurement errors (40,41,42,22). The measurement bias was greatest for households that did not use shot records to respond, but still large for households that did use shot records, because these records are often not current. Data from children's immunization providers were clearly of the highest quality and are now gathered as part of the regular NIS data collection. Because the NIS now uses data from providers, these measurement errors are minimized.

9. Summary

A principal goal of the National Immunization Survey (NIS) has been, and continues to be, the production of high-quality data on vaccination coverage for young children. This goal is accomplished through a variety of systematic procedures. Many of these processes have been in place since the start of the survey; others are evolving as new approaches are introduced and tested. Future goals of the quality assurance and quality control effort for the NIS focus on maintaining the high quality of the data and its usefulness in fulfilling the survey's objectives.

The NIS will continue these practices and, where it can, try to advance the quality and efficiency of the survey. Alternatives and enhancements based on careful research can be used to improve, for example, response rates of households and providers, eligibility rates, and the usability of data obtained from providers. Toward this objective, the NIS has developed a protocol to ensure that modifications are considered and made in a scientific manner and that, where appropriate, multiple approaches are considered. Every effort will be made to get timely approvals, to test recommendations, and to document results, so that not only will the NIS benefit from change, but also the research community.

Following is a list of research that is under way to enhance quality and dissemination of the NIS data:

- Additions/enhancements to the household questionnaire to meet current public health needs in the area of immunization
- Revision of the IHQ to improve quality of the provider-reported histories and information on new vaccines
- Improvements in the weighting adjustment and estimation methodology
- Evaluation to improve estimates of coverage rates by supplementing provider-reported immunization histories with the household-reported histories from a 'shot card'
- Wider, more timely distribution of the NIS estimates and public-use files to research professionals and policy-makers
- A comparison of the NIS sample in a selected state with the state's immunization registry

10. Bibliography

1. Centers for Disease Control. National Immunization Program. Evolution of vaccines. http://www.cdc.gov/nip/dev/evolution.htm Aug 1, 2001.
2. Centers for Disease Control. Reported vaccine-preventable diseases – United States, 1993, and the Childhood Immunization Initiative. Morbidity and Mortality Weekly Report (MMWR) 43:57-60. 1994.
3. National Center for Health Statistics. Healthy People 2000 review. Hyattsville, Maryland: Public Health Service. 1998-99.
4. U.S. Department of Health and Human Services. Healthy People 2010 (conference edition, 2 volumes). Washington. 2000.
5. Centers for Disease Control. Notice to readers: Recommended childhood immunization schedule – United States, 2000. Morbidity and Mortality Weekly Report (MMWR) 49:35-38, 47. 2000.
6. Centers for Disease Control. National, state, and urban area vaccination coverage levels among children aged 19-35 months — United States, 1999. Morbidity and Mortality Weekly Report (MMWR) 49:585-589. 2000.
7. Lepkowski JM. Telephone sampling methods in the United States. In: Groves RM, Biemer PP, Lyberg LE, et al., eds. Telephone survey methodology. New York: John Wiley & Sons, 73-98. 1988.
8. Tucker C, Casady RJ, Lepkowski J. A hierarchy of list-assisted stratified telephone sample design options. In: 1993 Proceedings of the Section on Survey Research Methods. Alexandria, VA: American Statistical Association. 982-987. 1993.
9. Brick JM, Waksberg J, Kulp D, Starer A. Bias in list-assisted telephone samples. Public Opinion Quarterly 59:218-235. 1995.
10. Ezzati-Rice TM, Zell ER, Battaglia MP, et al. The design of the National Immunization Survey. In: 1995 Proceedings of the Section on Survey Research Methods. Alexandria, VA: American Statistical Association. 668-672. 1995.
11. Zell ER. Ezzati-Rice TM, Battaglia MP, Wright RA. National Immunization Survey: The methodology of a vaccination surveillance system. Public Health Reports 115:65-77. 2000.
12. Smith PJ, Battaglia MP, Huggins VJ, et al. Overview of the sampling design and statistical methods used in the National Immunization Survey. American Journal of Preventive Medicine 20(4S):17-24. 2001.
13. Khare M, Battaglia MP, Huggins VJ, et al. Accuracy of vaccination dates reported by immunization providers in the National Immunization Survey. In: 2000 Proceedings of the Section on Survey Research Methods. Alexandria, VA: American Statistical Association. 665-670. 2000.
14. Lessler JT, Kalsbeek WD. Nonsampling error in surveys. New York: John Wiley & Sons. 1992.
15. Judkins, D. Directions for research on the National Immunization Survey. Research Triangle Institute, Contract No. 200-93-0697, Task 10. 1996.

16. General Accounting Office. CDC's National Immunization Survey: Methodological problems limit survey's utility. GAO/PEMD-96-16. 1996.
17. Battaglia MP, Starer A, Oberkofler J, et al. Pre-identification of nonworking and business telephone numbers in list-assisted random-digit-dialing samples. In: 1995 Proceedings of the Section on Survey Research Methods. Alexandria, VA: American Statistical Association. 957-962. 1995.
18. Centers for Disease Control. National Center for Health Statistics. Health measures in the new 1997 redesigned National Health Interview Survey (NHIS). http://www.cdc.gov/nchs/about/major/nhis/hisdesgn.htm
19. Abt Associates. The National Immunization Survey (NIS): 1994 annual methodology report. Chicago: Abt Associates Inc. 1995.
20. Massey JT, Wolter C, Wan SC, et al. Optimum calling patterns for random digit dialed telephone surveys. In: 1996 Proceedings of the Section on Survey Research Methods. Alexandria, VA: American Statistical Association. 485-490. 1996.
21. Dennis JM, Saulsberry C, Battaglia MP, et al. Analysis of call patterns in a large random-digit dialing survey: The National Immunization Survey. Paper presented at the International Conference on Survey Non-Response, October 28-31, Portland, OR. 1999. http://www.jpsm.umd.edu/icsn/papers/Index.htm
22. Zell ER, Ezzati-Rice TM, Hoaglin DC, et al. Adjusting for respondent bias on vaccination status in a telephone survey. In: 1995 Proceedings of the Section on Survey Research Methods. Alexandria, VA: American Statistical Association. 684-689. 1995.
23. Peak RR, Cadell DM. Overview of the National Immunization Provider Record Check Study. In: 1996 Proceedings of the Section on Survey Research Methods. Alexandria, VA: American Statistical Association. 332-334. 1996.
24. Dillman D. Mail and telephone surveys: The total design method. New York: John Wiley & Sons. 1978.
25. Office of Management and Budget. Standards for maintaining, collecting, and presenting federal data on race and ethnicity. http://www.whitehouse.gov/omb/fedreg/ombdir15.html Oct. 30, 1997.
26. Cox, BG. The weighted sequential hot deck imputation procedure. In: 1980 Proceedings of the Section on Survey Research Methods. Washington, DC: American Statistical Association, 721-726. 1980.
27. Shapiro GM, Battaglia MP, Camburn DP, et al. Calling local telephone company business offices to determine the residential status of a wide class of unresolved telephone numbers in a random-digit dialing sample. In: 1995 Proceedings of the Section on Survey Research Methods. Alexandria, VA: American Statistical Association. 975-980. 1995.
28. Battaglia MP, Malec DJ, Spencer BD, et al. Adjusting for noncoverage of nontelephone households in the National Immunization Survey. In: 1995 Proceedings of the Section on Survey Research Methods. Alexandria, VA: American Statistical Association. 678-683. 1995.

29. Smith PJ, Rao JNK, Battaglia MP, et al. Compensating for nonresponse bias in the National Immunization Survey using response propensities. In: 2000 Proceedings of the Section on Survey Research Methods. Alexandria, VA: American Statistical Association. 641-646. 2000.
30. Smith PJ, Rao JNK, Battaglia MP, et al. Compensating for provider nonresponse using response propensities to form adjustment cells: The National Immunization Survey. Vital and Health Statistics. Series 2, No. 133 (DHHS publication no. (PHS) 2001-1333). Hyattsville, MD: National Center for Health Statistics: 2001.
31. Smith PJ, Srinath KP, Battaglia MP, et al. Issues relating to the use of jackknife methods in the National Immunization Survey. In: 2000 Proceedings of the Section on Survey Research Methods. Alexandria, VA: American Statistical Association. 709-714. 2000.
32. Stokley S, Battaglia MP, Khare M, et al. Comparison of the NIS and NHIS/NIPRCS estimation methods. In: 2000 Proceedings of the Section on Survey Research Methods. Alexandria, VA: American Statistical Association. 721-725. 2000.
33. Frankel MR, Srinath KP, Battaglia MP, et al. Using data on interruptions in telephone service to reduce nontelephone bias in a random-digit-dialing survey. In: 2000 Proceedings of the Section on Survey Research Methods. Alexandria, VA: American Statistical Association. 647-652. 2000.
34. Frankel MR, Srinath, KP, Battaglia M, et al. Reducing nontelephone bias in RDD surveys. In: 1999 Proceedings of the Section on Survey Research Methods. Alexandria, VA: American Statistical Association. 934-939. 1999.
35. GENESYS Sampling Systems. GENESYS Sampling System: methodology. http://www.m-s-g.com/refernce/genmeth.htm August 1997.
36. Camburn DP, Lavrakas PJ, Battaglia MP, et al. Using advance letters in random-digit-dialing telephone surveys. In: 1995 Proceedings of the Section on Survey Research Methods. Alexandria, VA: American Statistical Association. 969-974. 1995.
37. Massey JT, Wolters CL, Wan SC. Analysis of calling patterns for the National Immunization Survey. Paper presented at the Joint Statistical Meetings, August 4-8, Chicago. 1996.
38. Srinath KP, Battaglia MP, Cardoni J, et al. Balancing cost and mean squared error in RDD telephone surveys: The National Immunization Survey. Paper presented at the annual meeting of the American Association for Public Opinion Research, May 18, Montreal. In: Proceedings of the 2001 Annual Meeting of the American Statistical Association (CD-ROM). Alexandria, VA: American Statistical Association. 2001.
39. Battaglia MP, Zell ER, Ching PLYH. Can participating in a panel sample introduce bias into trend estimates? In: 1996 Proceedings of the Section on Survey Research Methods. Alexandria, VA: American Statistical Association. 1010-1013. 1996.
40. Battaglia MP, Shapiro G, Zell ER. Substantial response bias may remain when records are used in a telephone survey. In: 1996 Proceedings of the Section on Survey Research Methods. Alexandria, VA: American Statistical Association. 452-455. 1996.
41. Battaglia MP, Ezzati-Rice TM, and Zell ER, Respondent characteristics associated with misreporting of vaccinations in a telephone survey. In: 1997 Proceedings of the Section on Survey Research Methods. Alexandria, VA: American Statistical Association. 976-

980. 1997.
42. Willis G, Brittingham A, Lee L, et al. Response errors in surveys of children's immunizations. Vital and Health Statistics. Series 6, No. 8 (DHHS publication no. (PHS) 99-1083). Hyattsville, MD: National Center for Health Statistics. 1999.

Appendix 1

A Matching Sheet and the Matching-Sheet Review Process

Example of a Matching Sheet (Artificial Data)

```
                                                MULTIPLE PROVIDER / HH MATCH
                          Note: ! - marks dates before BEST_DOB or MMR before 9month, ? - marks too close dates within provider
                                                * - too close dates between providers

CASEIDX: XXXXXXXXX   Matching Sheet Group: 4
 LAST DIGIT CASEIDXY1: Y  LAST DIGIT CASEIDXY2: Y   Batch1: YYY   Batch2: YYY
DATE OF INTERVIEW: 08/26/1998     RDD DOB: 10/02/1996     PROVIDER1 DOB : 10/02/1996   PROVIDER2 DOB : 10/02/1996   BEST DOB : 10/02/1996
RESULTING PROVIDER DOB: 10/02/1996   PREPRINTED DOB1: 10/02/1996   PREPRINTED DOB2: 10/02/1996
FIRST VISIT1: 06/04/1998     LAST VISIT1: 06/04/1998    FIRST VISIT2:                LAST VISIT2:

SOURCE: HISTORY
43133 UTD HH:       Yes
        PROVIDER: Yes
ADJUDICATION GROUP:  MULTIPLE PROVIDER

REVISED NIS DISPCODE: 2
NUMBER OF PROVIDERS WITH IMMUNIZATION INFORMATION: 2
NUMBER OF PROVIDERS ON ROSTER                   : 2

PROVIDER FACILITY TYPE(S): 1. Public Health Department-operated clinic
                           2. Private Practice

          DATE1          OP1       DATE2       OP2       DATE3       OP3       DATE4       OP4       DATE5       OP5            Total Shots        Total Shots
                                                                                                                                Including After
                                                                                                                                    Interview

DTP HH                                                                                                                          All                All
 PROV#1  12/14/1996             02/15/1997      X     04/19/1997      X                                                          3                  3
 PROV#2  04/14/1996!            02/15/1997            04/19/1997            09/15/1997                                           4                  4

POL HH                                                                                                                          All                All
 PROV#1  02/15/1997      X      04/19/1997      X                                                                              2                  2
 PROV#2  02/15/1997             04/19/1997            09/15/1997                                                             3                  3

MCV HH                                                                                                                          All                All
 PROV#1  04/10/1997!     X                                                                                                      1                  1
 PROV#2  04/10/1997!                                                                                                            1                  1

HIB HH                                                                                                                          All                All
 PROV#1  12/14/1996             02/15/1997      X     04/19/1997 *    X                                                          3                  3
  HIB RAW     12/14/1996        02/15/1997            04/19/1997
  DTP/HIB RAW
  HEP-HIB RAW
 PROV#2  02/15/1997             04/14/1997 ?          04/19/1997 ?          09/15/1997                                           4                  4
  HIB RAW     04/14/1997        02/15/1997            04/19/1997            09/15/1997
  DTP/HIB RAW
  HEP-HIB RAW

HEP HH                                                                                                                          All                All
 PROV#1  10/02/1996      X      01/05/1997      X     04/19/1997      X                                                          3                  3
 PROV#2  02/10/1996!            05/11/1996!           04/19/1997                                                             3                  3
 HEPB AT BIRTH CHECKED BY PROV#1:
 HEPB AT BIRTH CHECKED BY PROV#2:

VRC HH                                                                                                                          All                All
 PROV#1                                                                                                                         0                  0
 PROV#2  04/15/1998                                                                                                             1                  1

               DATE1       DATE2       DATE3       DATE4       DATE5       DATE6       DATE7       DATE8       DATE9       DATE10

VISITS HH
VISITS PROV#1 10/02/1996  12/14/1996  01/05/1997  02/15/1997  04/10/1997  04/19/1997
       PROV#2 02/10/1996  04/14/1996  05/11/1996  02/15/1997  04/10/1997  04/14/1997  04/19/1997  09/15/1997  04/15/1998
```

NIS Edit Rules and Guidelines for Household/Provider Matching Sheets

April 12, 2000

Revised October 5, 2000

I. Date of Birth Differences between Provider and Household

NO SHOT CARD

1. If dates of birth differ by *exactly* one, two or three years, and provider shot dates are consistent with provider DOB, then accept provider DOB as best DOB and mark "OK" on matching sheet. If provider shot dates are not consistent with provider DOB and shot dates are consistent with household DOB, then edit best DOB to be household DOB. Otherwise, mark as provider form filled out for wrong child ("WC").

2. If dates of birth agree on year *and* month but differ on day *or* agree on year *and* day but differ on month *or* the month and day are transposed, and provider shot dates are consistent with provider DOB, then accept provider DOB as best DOB and mark "OK" on matching sheet. If provider shot dates are not consistent with provider DOB and shot dates are consistent with household DOB, then edit best DOB to be household DOB. Otherwise, mark as "WC".

3. If dates of birth differ by less than one month, and provider shot dates are consistent with provider DOB, then accept provider DOB as best DOB and mark "OK" on matching sheet. If provider shot dates are not consistent with provider DOB and shot dates are consistent with household DOB, then edit best DOB to be household DOB. Otherwise, mark as "WC".

4. If dates of birth differ in any other way, give strong consideration to marking matching sheet with "WC" (e.g., the provider and household DOBs are totally different). However, if DOBs differ by somewhat more than one month or in some pattern that looks like it could be an error in recording the DOB (e.g., 05/24/97 vs. 05/23/96), and provider shot dates are consistent with provider DOB, then consider accepting provider DOB as best DOB and mark "OK" on matching sheet. If provider shot dates are not consistent with provider DOB and shot dates are consistent with household DOB, then edit best DOB to be household DOB.

5. If the provider left the DOB blank, and provider shot dates are consistent with household DOB, then accept household DOB as best DOB and mark "OK" on matching sheet. Otherwise, mark as "WC".

SHOT CARD

6. If dates of birth differ by *exactly* one, two or three years, and provider shot dates and household shot dates are *not* similar, mark matching sheet with "WC". If provider and household shot dates are similar and shot dates are consistent with provider DOB, then accept provider DOB as best DOB and mark "OK" on

matching sheet. If provider shot dates are not consistent with provider date of birth and shot dates are consistent with household DOB, then edit best DOB to be household DOB.

7. If dates of birth agree on year *and* month but differ on day *or* agree on year *and* day but differ on month *or* the month and day are transposed, and provider shot dates and household shot dates are *not* similar, mark matching sheet with "WC". If provider and household shot dates are similar and shot dates are consistent with provider DOB, then accept provider DOB as best DOB and mark "OK" on matching sheet. If shot dates are not consistent with provider DOB and shot dates are consistent with household DOB, then edit best DOB to be household DOB. Otherwise, mark as "WC".

8. If dates of birth differ by less than one month, and provider shot dates and household shot dates are *not* similar, mark matching sheet with "WC". If provider and household shot dates are similar and shot dates are consistent with provider DOB, then accept provider DOB as best DOB and mark "OK" on matching sheet. If shot dates are not consistent with provider DOB and shot dates are consistent with household DOB, then edit best DOB to be household DOB. Otherwise, mark as "WC".

9. If dates of birth differ in any other way, and provider shot dates and household shot dates are *not* similar, mark matching sheet with "WC". If provider and household shot dates are similar and shot dates are consistent with provider DOB, then accept provider DOB as best DOB and mark "OK" on matching sheet. If shot dates are not consistent with provider DOB and shot dates are consistent with household DOB, then edit best DOB to be household DOB. Otherwise, mark as "WC".

10. If the provider left the DOB blank, and provider shot dates and household shot dates are *not* similar, mark matching sheet with "WC". If provider and household shot dates are similar and shot dates are consistent with household DOB, then accept household DOB as best DOB and mark "OK" on matching sheet. Otherwise, mark as "WC".

II. Shot Dates before the Best Date of Birth

1. If there is a difference between the provider DOB and the household DOB, first follow the rules given in Section I.

2. If step 1 does not resolve the problem, **for shot card children**, compare the household and provider shot dates to identify the correct date for the shot before the best DOB. If the household fails to list this shot date in the shot card, then look at the other visits in the household and provider shot dates to see whether a similar visit date exists. Examples include the same day and month but a different year and a transposition of the day and month. If the shot date was recorded out of sequence (e.g., the first DTP box has a date of 3/4/97 and the second DTP box has a date of 4/28/96), review the raw shot date listing on the matching sheet to establish when the shot was given relative to its actual position in the original shot grid. If no such similarity exists, look at the other shot dates for the vaccine in question and the immunization schedule, try to establish the most likely correct date for the shot.

3. If step 1 does not resolve the problem, **for non-shot-card children**, look at the other visits in the provider shot dates to see whether a similar visit date exists. Examples include the same day and month but a different year and a transposition of the day and month. If the shot date was recorded out of sequence (e.g., the first DTP box has a date of 3/4/97 and the second DTP box as a date of 4/28/96), review the raw shot date listing on the matching sheet to establish when the shot was given relative to its actual position in the original shot grid. If no such similarity exists, look at the other shot dates for the vaccine in question and the immunization schedule, try to establish the most likely correct date for the shot.

4. Special additional conditions exist for Hepatitis B. For Hepatitis B shot dates before the best DOB, look to see whether the shot date is a few days before the best DOB. If yes, change the shot date to the best DOB, provided there is no other birth dose with a different date. If another birth dose already exists, then use that as the correct shot date. If these procedures do not resolve the problem, look to see whether the "given at birth" box is checked. If yes, change the shot date to the best DOB, provided there is no other birth dose with a different date. Be careful: the Hepatitis B shot date before the best DOB is not always the birth dose (e.g., it could be the final dose).

III. Shot Dates That Are Close Together

1. If a shot card was used or if two or more providers responded with vaccination information, it is generally possible to examine the shot dates that are close together and determine which one was recorded incorrectly. One must, however, be on guard against incorrect shot cards. It is possible that the child's first provider recorded the shot date correctly because they actually administered the shot. Then the child's current provider records the date incorrectly and also creates a shot card for the child with the wrong date also on the shot card.

2. If the above procedure does not revolve the problem or if the child has no shot card and only one provider responded with vaccination information, two other variables should be consulted to determine which provider actually administered the shot. The "OP" indicator next to a shot date means that the provider did not administer that shot. The first and last visit dates can also be consulted to determine which provider was serving the child at the time the shot was given. In general, the provider who administered the shot is more likely to record the date correctly.

3. It is also very useful to consult the other vaccine types to see whether they show which shot date is incorrect. Often, the first, second and third doses of DTP, Polio and Hib are given at the same visit.

4. Also consult the listing of unique visit dates at the bottom of the matching sheet to see whether one of those shot dates involves a date component transposition error.

5. Special caution needs to be used for the first and second Hepatitis B shots. It appears that some providers record the birth dose of Hepatitis B even if they did not administer that dose. Within 2-4 weeks the child may visit the provider and receive another Hepatitis B shot.

6. If the above rules and guidelines do not resolve the close shot dates, make no edits to the shot dates.

IV. MCV shots before 9 months of age

1. First compare the shot date recorded to the other shots dates. If it matches with a shot date given for another vaccine and is relatively close to nine months after best DOB, then write OK next to it and go to the next matching sheet. However, if the date for the MCV vaccine is deemed to be too close to the best DOB, check between other shot dates to see if the month and day may have been transposed, i.e. **05/11/1997** and **11/05/1997** or if the wrong year was recorded, i.e **05/11/1996** and **05/11/1997.**

V. Non-Consistent shot dates (shot card)

1. If you find that none, or very few, of the shot dates reported from the household and the provider are in agreement, then look first at the reported **DOB's**. If there is disagreement between the household and provider DOB as well consider marking the matching sheet with a WC for wrong child. If the reported DOB's are the same and the shot dates given by the provider are consistent with the DOB, do not add any household dates to make the shot record complete. If the reported DOB's are the same and the shots dates given by the household are consistent with the DOB and the provider dates are not, consider inserting the household dates for the provider dates.

Appendix 2

Vaccination-Type Arrays

VACCINATION-TYPE INDICATOR VARIABLES FOR USE WITH VACCCINATION-DATE ARRAYS AND AGE -AT-VACCINATION ARRAYS

Vaccination-Type Indicator Variable Description	Vaccination Type Code	Specific Type of Vaccination Recorded on Immunization History Questionnaire
DTP (›DTP/DT-containing vaccine:)	01	DT
	02	DTP
	03	DTP - unknown type
	04	DTaP
	05	DTP/Hib
	06	DTP/Hib - unknown type
	07	DTaP/Hib
POLIO (›Polio-containing vaccine:)	20	OPV
	21	IPV
	22	Polio - unknown type
MCV (›Measles-containing vaccine:)	30	MMR
	31	Measles only
	32	Measles/Mumps
	33	Measles/Rubella

HIB (›Hib-containing vaccine:)	40	Pedvax Hib
	41	Other Hib
	42	Hib - unknown type
	05	DTP/Hib
	06	DTP/Hib - unknown type
	07	DTaP/Hib
	43	Hep B - Hib
HEP B (›Hep B-containing vaccine:)	60	Hep B only
	43	Hep B - Hib