
Vital and Health Statistics

Design and Operation of the National Hospital Discharge Survey: 1988 Redesign

Series 1:
Programs and Collection
Procedures
No. 39

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Centers for Disease Control and Prevention
National Center for Health Statistics

Hyattsville, Maryland
December 2000
DHHS Publication No. (PHS) 2001-1315

Copyright Information

All material appearing in this report is in the public domain and may be reproduced or copied without permission; citation as to source, however, is appreciated.

Suggested Citation

Dennison C, Pokras R. Design and operation of the National Hospital Discharge Survey: 1988 redesign. *Vital Health Stat* 1(39). 2000.

Library of Congress-in-Publication Data

Dennison, Charles F.

Design and operation of the National Hospital Discharge Survey / [by Charles Dennison and Robert Pokras].

p. cm.—(Vital and health statistics. Series 1, Programs and collection procedures ; no. 39) (DHHS publication ; no. (PHS) 2001-1315)

"October 2000."

Includes bibliographical references.

ISBN 0-8406-0565-X

1. National Hospital Discharge Survey (U.S.) 2. Hospital utilization—United States—Statistics. 3. Medical care surveys—United States 4. Health surveys—United States. I. Pokras, Robert. II. National Center for Health Statistics (U.S.) III. Title IV. Series. V. Series: DHHS publication ; no. (PHS) 2001-1315.

RA409.U44 no. 39

[RA407.3]

362.1'07'23 s—dc21

[362.1'1'0973]

00-064743

For sale by the U.S. Government Printing Office
Superintendent of Documents
Mail Stop: SSOP
Washington, DC 20402-9328
Printed on acid-free paper.

National Center for Health Statistics

Edward J. Sondik, Ph.D., *Director*

Jack R. Anderson, *Deputy Director*

Jack R. Anderson, *Acting Associate Director for International Statistics*

Jennifer H. Madans, Ph.D., *Associate Director for Science*

Lester R. Curtin, Ph.D., *Acting Associate Director for Research and Methodology*

Jennifer H. Madans, Ph.D., *Acting Associate Director for Analysis, Epidemiology, and Health Promotion*

P. Douglas Williams, *Acting Associate Director for Data Standards, Program Development, and Extramural Programs*

Edward L. Hunter, *Associate Director for Planning, Budget, and Legislation*

Jennifer H. Madans, Ph.D., *Acting Associate Director for Vital and Health Statistics Systems*

Douglas L. Zinn, *Acting Associate Director for Management*

Charles J. Rothwell, *Associate Director for Information Technology and Services*

Division of Health Care Statistics

Linda Demlo, Ph.D., *Director*

Thomas McLemore, *Deputy Director*

Irma Arispe, Ph.D., *Associate Director for Science*

Robert Pokras, *Chief, Hospital Care Statistics Branch*

Contents

Abstract	1
Introduction	1
Background	1
Sample Design.....	2
First Stage Sampling—Primary Sampling Units	2
Second Stage Sampling—Hospitals	2
Third Stage Sampling—Discharges	3
Confidentiality	4
Survey Operations	4
Manual Data Collection and Processing	4
Automated Data Collection and Processing.....	6
Estimation Procedures	9
Inflation by Reciprocals of Probabilities of Selection	9
Adjustment for Nonresponse	9
Population Weighting Ratio Adjustment	10
Reliability of Estimates	10
Estimation of Standard Errors.....	10
Standard Error Approximations	10
Relative Standard Error for Estimates of Percents	11
Nonsampling Errors.....	12
Data Dissemination	12
References	12

Text Tables

A. Definition of noncertainty hospital specialty-size groups used as secondary strata in the National Hospital Discharge Survey sample design, 1988 sample design	3
B. Number of hospitals in the National Hospital Discharge Survey sample, number of in-scope and responding sample hospitals, response rates, and number of abstracts collected: United States, 1988–97	3
C. Estimated parameters for approximate relative standard error equations by selected characteristics: National Hospital Discharge Survey, 1997	11

Appendixes

I. Legislative Authorization	14
II. Induction Letter	15
III. American Hospital Association Endorsement Letter	16
IV. Confidentiality Letter	17
V. Hospital Interview Questionnaire (HDS-11)	18
VI. Medical Abstract—National Hospital Discharge Survey	29

VII.	Sample Selection Sheet (HDS-5)	31
VIII.	Statement of Agreement	32
IX.	Abstract Service Agreement	33
X.	Transmittal Notice	36
XI.	Automated File Layout	37
XII.	Information Request Letter	41
XIII.	Definitions of Terms Related to Survey	42

Abstract

The National Hospital Discharge Survey (NHDS), a national probability sample survey of discharges from non-Federal hospitals, began in 1965 and has been conducted annually since then. The original design of NHDS was in place through 1987. This report provides information about the survey design, instruments, data collection procedures, and survey methodology used for NHDS since the implementation of its redesign in 1988.

Keywords: inpatients, hospital, hospital discharges, National Hospital Discharge Survey

Design and Operation of the National Hospital Discharge Survey: 1988 Redesign

by Charles Dennison and Robert Pokras, M.A., Division of Health Care Statistics

Introduction

In 1965 the National Center for Health Statistics (NCHS) initiated the National Hospital Discharge Survey (NHDS) to collect and disseminate data on inpatient utilization of short-stay non-Federal hospitals in the United States and District of Columbia. The importance of measuring inpatient care made NHDS the first survey of medical care delivery conducted by the NCHS (1,2) and inpatient care continues to play a major role in health care delivery (3). The NHDS has been conducted annually since its inception.

Authority for the NHDS derived from the National Health Survey Act of 1956 (Public Law 84-652) and it has continued under various acts of Congress, principally under Section 306(b)(1)(F) ([appendix I](#)) of the Public Health Service Act (42 USC 242k) and its amendments to collect data concerning the public's use of health care services.

The NHDS was conducted under the same design from 1965 through 1987. A previous report describes the development and design of the NHDS during this period (4). A redesign for the survey was implemented in 1988. This report describes the design and operation of the NHDS, which began with the 1988 survey year and major changes to NHDS resulting from the new design. These changes include the use of a 3-stage sample design, the use

of automated data, and the ability to use SUDAAN to estimate variances.

Background

The impetus to redesign the NHDS was to select a new independent national probability sample of hospitals. The initial sample of hospitals for NHDS was used for the 23-year period 1965-87. The redesign focused on improving the efficiency and analytic capability of the survey by 1) linking the NHDS to the design of NCHS's National Health Interview Survey (NHIS), 2) incorporating hospital discharge data available in electronic form, and 3) allowing the use of existing statistical software to estimate variances.

The redesign included linkage with the design of NHIS (5), the principal source of information on the health of the civilian noninstitutionalized population of the United States. Before 1988 the NHDS used a two-stage sample design (a sample of hospitals and a sample of records within hospitals); the redesign included a third stage prior to sampling hospitals, the selection of geographic primary sampling units (PSU's). The redesign used a sub-sample of PSU's selected for the 1985-94 NHIS. This modification was intended to enhance the analytic capabilities of both surveys and to reduce the field costs of the surveys by conducting them in the same geographic areas.

The redesign also took into account the increasing availability of discharge data in electronic (automated) form. Before the redesign, the NHDS depended almost entirely on manual sampling and abstraction of information from medical records. The exception to this was that automated data were obtained for about 75 hospitals annually during the period 1985–87. With the implementation of the new design in 1988, information on the characteristics of a majority of discharges in NHDS were obtained from existing hospital discharge data bases. However, a majority of hospitals provide data by means of manual sampling and abstraction. This seeming contradiction is explained by the fact that, as part of the redesign, more records are sampled from hospitals providing automated data than are sampled from hospitals in which data are manually abstracted from medical records. The target sample size is 250 discharges from manual hospitals and 2,000 discharges from automated system hospitals.

Also, the new redesign allows the use of available statistical software to estimate variances for estimates. Before 1988 NCHS used a program written at NCHS specifically for the NHDS to calculate variances. The redesign allows the use of SESUDAAN and SUDAAN (6, 7) for this purpose. This is more efficient for data dissemination and it allows researchers to more readily generate variances for specific areas of interest.

There are a number of less significant differences between the original and new designs that have been documented (8,9). These include a different source as the operational definition of the hospital universe and minor differences in the static design of the survey. The changes introduced do not compromise the ability to conduct trend analysis; however, the addition of clustering introduced by including a third stage of sampling (PSU's) reduced the precision of estimates.

Sample Design

The NHDS is based on a national probability sample of discharges from noninstitutional hospitals exclusive of Federal, military, and Department of Veterans Affairs hospitals, located in the 50 States and the District of Columbia. Only short-stay hospitals (hospitals with an average length of stay for all patients of less than 30 days) or those whose speciality is general (medical or surgical) or children's general regardless of length of stay are included in the survey. Also, a hospital must have six or more beds staffed for patient use to be in scope for the survey. The 1988 NHDS sampling frame consisted of hospitals that were listed in the April 1987 SMG Hospital Market Database (10) and that began to accept inpatients by August 1987.

Given fixed costs, the NHDS is designed to provide estimates of inpatient hospital utilization based on the following priority of objectives: national aggregate statistics, national trend statistics, and aggregate statistics for the four major Census Regions of the United States. The NHDS uses a modified three-stage probability design, the stages are: 1) primary sampling units (PSU's); 2) hospitals within PSU's; and 3) discharges within hospitals. The modification was that the largest PSU's and hospitals were selected with certainty.

First Stage Sampling— Primary Sampling Units

The first stage of sampling consisted of 112 primary sampling units (PSU's) that comprised a probability subsample of PSU's used in the 1985–94 NHIS. PSU's are counties, groups of counties, county equivalents (such as parishes or independent cities), or towns and townships (for some PSU's in New England). The NHDS sample included with certainty 26 PSU's with the largest populations. In addition, the sample included one-half of the next 26 largest PSU's, and one PSU from each of the 73 PSU's strata formed from

the remaining PSU's in the NHIS sample design. Those 73 PSU's strata were defined within four geographic regions and metropolitan statistical area (MSA) or non-MSA status (MSA was a metropolitan statistical area defined by the U.S. Office of Management and Budget on the basis of the 1980 census). PSU's were assigned to strata by using 1980 Census of Population data and a computer program that minimized the between-PSU variances for the NHIS stratification variables. From the 73 strata thus formed, the PSU's were selected with probability proportional to the projected 1985 population. A more detailed description of the NHIS PSU sample design has been published (5).

Second Stage Sampling—Hospitals

The second stage of sampling consisted of a systematic random sample of noncertainty hospitals selected from the sample PSU's with probability proportional to their annual number of discharges. To assure distribution of the sample across PSU's and to maximize the potential for automated data collection, the noncertainty hospitals were stratified by region, PSU, and in the 12 largest PSU's, by data collection type (whether the hospital subscribed to a commercial abstracting service). Within the strata, the hospitals were ordered by PSU, whether the hospital participated in the 1987 NHDS and by hospital size and speciality as defined in [table A](#). Finally, hospitals were arrayed within speciality by their annual number of discharges. The sampling rates were such that at least three hospitals were selected from every PSU containing three or more eligible hospitals. In PSU's with fewer than three hospitals, all hospitals in the PSU were selected.

The sample design resulted in the selection of 542 hospitals for the NHDS in 1988. Of these, 11 hospitals were out of scope and 109 refused to participate or did not provide enough data to be considered responding. This resulted in 422 hospitals participating in the survey, a response rate of 79.5 percent. Through a concerted effort to increase the hospital response rate, this was raised to

Table A. Definition of noncertainty hospital specialty-size groups used as secondary strata in the National Hospital Discharge Survey 1988 sample design

Hospital group	Bed size	Type of service
Group 1	6–999 beds	Selected specialties ¹
Group 2	6–174 beds	General (medical and surgical) and other specialties ²
Group 3	175–349 beds	General (medical and surgical) and other specialties ²
Group 4	350–999 beds	General (medical and surgical) and other specialties ²

¹Includes psychiatry, tuberculosis and other respiratory disease, rehabilitation, chronic disease, mental retardation, alcoholism and other chemical dependency, and children's psychiatry.

²"Other specialties" include obstetrics and gynecology; eye, ear, nose and throat; orthopedics; other specialty; children's general; children's tuberculosis and other respiratory disease; children's eye, ear, nose, and throat; children's rehabilitation; children's orthopedics; children's chronic disease; and children's other specialty.

91.3 percent in 1990 and it did not fall below 90 percent through 1997. [Table B](#) provides information on hospital participation for the NHDS from 1988 through 1997.

NCHS conducts several activities to keep the sample of hospitals in the NHDS current with the changing universe of hospitals in the United States: hospitals in the NHDS sample that no longer meet eligibility requirements for the survey are removed; methods were developed and implemented for hospitals that merge; and NCHS samples a "birth panel" of hospitals every 3 years from all new hospitals that came into existence since the previous set of birth panel hospitals were selected. Research has shown that adding hospitals from birth panels was effective in ensuring that estimates from the survey reflect the changing universe of hospitals (11).

Third Stage Sampling—Discharges

At the third stage, a sample of discharges from each hospital is selected by a systematic random sampling technique. For hospitals using the manual system of data collection, sampling rates for discharges are designed to generate samples of approximately equal numbers regardless of hospital size. This characteristic of the design allows efficiencies in field operations by ensuring that the number of medical records sampled and abstracted in the field can be completed in a single day, thereby avoiding repeated trips to a hospital by census personnel (see Manual Data Collection and Processing). To achieve this operational goal, discharges are sampled in inverse proportion to hospital size

with a target of 20–25 discharges sampled per month.

For hospitals whose data are collected via the automated system, discharges are sampled in one of two methods depending on whether NCHS receives a sample or a census file of discharges for the hospital. Some automated sources prefer to provide a sample of discharges. For these sources NCHS provides instructions and sampling specifications to use terminal digits of medical records to sample discharges as described above for manual data collection. For automated data that contains all discharges for hospitals, NCHS samples discharges after sorting by the first two digits of the *International Classification of Diseases, Clinical Modification, Ninth Revision*, or ICD–9–CM code (12) of the first-listed diagnosis, patient age group at time of admission (under 1 year, 1–14 years, 15–44 years, 45–64 years, 65–74 years, 75–84 years, 85 years and over, and age unknown), sex, and date of discharge. Discharges are sampled by starting with a randomly selected discharge and taking every *k*th discharge thereafter.

The third-stage sampling rate is determined by the hospital's sampling stratum and the data collection method (manual or automated) used to collect data from the hospital. One percent and 5 percent of discharges in the certainty hospitals are selected under the manual and automated systems. Except for certainty hospitals, the target sample size is 250 discharges annually from manual system hospitals, from the automated system hospitals that had fewer than 4,000 discharges annually according to the 1987 sampling frame data, and from automated hospitals that provided sample data. Samples of 2,000 discharges are targeted for each of the remaining noncertainty automated system hospitals that provide files of all discharges to NCHS. The sample design resulted in the selection of 300,464 discharges for the NHDS in 1997. [Table B](#) provides information on the number of sampled discharges included in the NHDS from 1988 through 1997.

Table B. Number of hospitals in the National Hospital Discharge Survey sample, number of in-scope and responding sample hospitals, response rates, and number of abstracts collected: United States, 1988–97

Year	Sample hospitals	In-scope hospitals				Hospital response rate	Discharges sampled
		Response type		Total respondents	Percent		
		Total	Manual				
1988	542	531	255	167	422	79.5	250,243
1989	542	526	256	152	408	77.6	233,493
1990	542	519	299	176	475	91.3	265,556
1991	¹ 528	521	323	161	484	92.9	274,311
1992	528	514	311	183	494	96.1	274,273
1993	528	513	303	163	466	90.8	235,411
1994	¹ 525	512	298	180	478	93.4	276,533
1995	525	508	299	167	466	91.7	262,809
1996	525	507	299	181	480	94.7	282,008
1997	¹ 513	501	284	187	474	94.6	300,464

¹Changes reflect triannual updates to the NHDS hospital universe.

Confidentiality

Participation in surveys conducted by the National Center for Health Statistics (NCHS) is voluntary, and information on individuals and/or facilities is confidential. For the NHDS, assurance of confidentiality was provided to all hospitals according to Section 308(d) of the Public Health Service Act (42 U.S.C. 242m), which states that:

“No information, if an establishment or person supplying the information or described in it is identifiable, obtained in the course of activities undertaken or supported under section ...306,... may be used for any purpose other than the purpose for which it was supplied unless such establishment or person has consented (as determined under regulations of the Secretary) to its use for such other purpose and in the case of information obtained in the course of health statistical or epidemiological activities under section ...306, such information may not be published or released in other form if the particular establishment or person supplying the information or described in it is identifiable unless such establishment or person has consented (as determined under regulations of the Secretary) to its publication or release in other form.”

Strict procedures are utilized to prevent disclosure of confidential data in survey operations and data dissemination. Names or other identifying information for individual patients are not obtained from the sampled hospitals. Data received in electronic form are maintained at the CDC Atlanta computer processing center. Under provisions of the DHHS guidelines for storage of health records, all original data collected for the survey are transferred to the Federal Records Center located in Atlanta, Georgia. These records are retained for 7 years before they are destroyed. To prevent identification of sampled hospitals or

patients from publicly available research files, variables such as medical record number, date of birth, admission and discharge dates, and patient ZIP Code are not released.

Survey Operations

This section describes the survey operations for manual and automated data collection procedures. With minor exceptions the survey forms, manuals, and operating procedures used for manual data collection during the initial design were continued with the redesigned survey in 1988. As described in the Introduction, there were three major factors that influenced the redesign of the NHDS, one of which was the use of existing discharge data in electronic (automated) form. Data for the NHDS were collected manually from 1965 through 1984. For the 3 years, 1985–87, NCHS purchased discharge data for about 75 hospitals annually from a contractor. Beginning in 1988 NCHS began purchasing automated data for approximately 170 hospitals annually.

Manual Data Collection and Processing

The Bureau of the Census has served as the agent for data collected manually in the NHDS since its inception in 1965. The Census Bureau is composed of a headquarters and 12 regional offices. In collaboration with NCHS, Census Headquarters staff developed the survey operation procedures; wrote, printed, and distributed all field manuals and forms for the NHDS; made modifications to these documents as variables were introduced or changed; and, generally oversaw operations of the manual data collection. Regional office staff are responsible for daily operations of the survey, training new field staff, and supplying survey forms and materials to hospital staff.

The major elements of the NHDS field operations consist of inducting sampled hospitals into the survey, sampling discharges, and abstracting

data from medical records. These functions are described below.

Hospital Induction

Hospital induction is the process of getting sampled hospitals to participate in NHDS. As a voluntary survey, NCHS, through the Census, must obtain written permission from each sampled hospital to abstract data from its medical records. While hospital participation is the primary purpose of induction, other important activities are performed.

Hospital induction begins when the Census regional office mails an introductory letter ([appendix II](#)) provided by the NCHS to the hospital administrator or chief executive officer. Enclosures with this letter include an information packet containing a fact sheet about the NHDS, a publication list and a description of data products available from the survey, an endorsement letter from the American Hospital Association ([appendix III](#)), a recent NHDS publication, and a letter that addressed confidentiality ([appendix IV](#)).

Approximately 5 days after mailing the introductory letter, Census staff call the hospital administrator to arrange an appointment for an induction interview. Since part of the interview relates to information about medical records and the medical records department, a representative from the medical records department (usually the Medical Records Administrator) is usually present at this meeting. The functions of the initial interview are to explain the purpose of the NHDS, to gain the hospital's cooperation in the survey, to describe the data collection methods available for survey participation (described below), to obtain information about the hospital by completing the Hospital Interview Questionnaire ([appendix V](#)), to set up sampling and data collection procedures, and to distribute manuals and forms.

During induction, a Hospital Interview Questionnaire is completed. Information about the facility such as ownership, service type, bed size, and average length of stay is collected to verify the hospital's eligibility for the survey. Information about the medical

records department, where and how records are kept and stored, and where to locate the survey variables in the medical records is collected to assist in survey operations. In addition, reimbursement rates are established for hospital personnel who work on survey activities. The level of involvement of hospital personnel in the NHDS is determined by the data collection method. Methods for manual data collection are described in the next section.

There was a large scale induction of hospitals in 1988 as a result of selecting a new sample of hospitals in the redesign. Although less frequent, inductions also take place as nonresponding hospitals agree to participate in the survey. Nonresponding hospitals are contacted annually in an effort to gain their participation. Also, hospitals that are added to the NHDS sample as part of the universe update (see Statistical Design) go through the induction process.

Manual Data Collection Methods

There were three options for manual data collection: primary, alternate, and printout. In the primary method, hospital personnel in the medical records department sample discharges for inclusion in the survey and abstract data from the sampled medical records onto the survey form ([appendix VI](#)). As the name implies, this is the preferred method because hospital staff are more familiar with their medical records and data than are Census personnel. In the primary method hospital staff are given field manuals, detailed instructions, and trained by Census personnel to perform the sampling and abstracting. In 1997 about 40 percent of all manual hospitals in the NHDS provided data using the primary method.

The alternate method is used when hospitals are unwilling or unable to have their personnel perform the sampling and abstracting. In this method, the Census Field Representative selects the sampled discharges, arranges with the hospital medical records staff to have

the medical records pulled for abstracting, and the Field Representative abstracts the records. Every effort is made to select experienced Census Field Representatives who have an understanding of medical terminology and are familiar with working with medical records professionals. Thirty percent of manual hospitals participated in this mode of manual data collection for the 1997 NHDS.

Hospitals also provide data in the “manual” mode by supplying computerized printouts of information for sampled discharges. This is considered manual data collection because patient information on the printouts required manual data entry (see Medical Coding and Data Entry). Hospitals using this mode provide definitions for coded variables (that is, 1 = male; 2 = female) to ensure correct data entry. Computerization of inpatient data has made printouts more common in the NHDS; this method of data collection accounted for about 8 percent of manual data collection in 1988 rising to about 30 percent in 1997.

Sampling Discharges

Discharges within hospitals are selected using systematic random sampling with the sampling intervals based on the statistical design of the survey and specified by NCHS. This is usually accomplished using daily or monthly discharge lists and selecting discharges with specified terminal digits of medical record numbers. In some cases an admission number, billing number, or other patient number is used. If patient specific numbers useful for sampling are not available, discharges are selected by using a random starting number and then physically counting through a discharge list and selecting every kth discharge thereafter. In a few hospitals discharges are sampled using the terminal digits of the medical record numbers on in-house computer files.

Information about sampled discharges is recorded on a Sample Listing Sheet ([appendix VII](#)). A Sample Listing Sheet is completed for each month’s sample of records. It includes information such as medical record numbers and dates of discharges that

allowed medical records to be identified and pulled for abstracting. Sampling schemes that require a physical count of discharges (as described above) are kept continuous by tracking on the Sample Listing Sheet where the sample selection process stopped in the previous period.

Medical Record Abstracting

Using the Sample Listing Sheets, completed medical records are pulled for abstracting. Records that are not complete or not available are located at a later time for abstracting. The NHDS was designed to sample approximately 20–25 discharges per month per hospital that allow Census personnel to sample and abstract 2 months of medical records in a single day.

Primarily using the medical record face sheet and discharge summary, patient information is abstracted onto the survey form ([appendix VI](#)). Variables collected in the NHDS conform with the Uniform Hospital Discharge Data Set (UHDDS) (13,14). These variables include birth date or age; sex; race; ethnicity; marital status; admission, discharge, and surgery dates; discharge status; patient ZIP Code; expected source(s) of payment; medical record number; and information on diagnoses and procedures. The NHDS survey methods instruct that, where possible, narrative information for diagnoses and procedures be collected, but in many instances only *International Classification of Diseases, Ninth Revision, Clinical Modification* or ICD–9–CM (12) codes are available. In 1997 about 30 percent of all manually abstracted records included only ICD–9–CM codes for diagnoses and procedures.

Additional variables are entered on the abstract form for survey operations. These included the hospital number (a number assigned to the hospital for processing by NCHS) and the HDS number (a number assigned to the sampled record for processing by NCHS).

Medical abstracts and Sample Listing Sheets are sent to the Census Regional Offices where forms for each month are logged in and reviewed for completeness. This allows Regional

Office staff to track survey activities and ensure that forms were completed correctly. Monthly shipments of abstracts and Sampling Listing Sheets from the Regional Offices are sent to the NCHS processing center in Research Triangle Park, North Carolina, for medical coding and data entry.

Medical Coding and Data Entry

For the 1965–97 survey years all materials sent from the Census Regional Offices were processed by the Division of Data Processing at the NCHS facility in Research Triangle Park, North Carolina. Beginning with the 1998 NHDS survey year these activities are being done by a contractor. Sample Listing Sheets and abstracts received by NCHS are recorded and tracked through a receipt and control process to monitor data flow and completeness. During this process, batches consisting of approximately 1,000 abstracts are formed and identified to track and conduct medical coding, data entry, and quality control.

Trained medical coding personnel code diagnoses and procedures using the ICD–9–CM. A minimum of one and a maximum of seven diagnostic codes are assigned for each sample abstract. If an abstract included surgical and/or diagnostic procedures, a maximum of four codes are assigned.

Beginning in 1986 annual modifications or addenda have been made to the ICD–9–CM. These addenda, which go into effect on October 1 of each year, added, deleted, or modified existing codes. However, because the NHDS is conducted for calendar years, the addenda are not incorporated into the manual coding process until the next survey year (beginning January 1, 3 months after the effective date). This results in consistent coding for all medical data for the calendar survey year.

There is quality control for medical coding and data entry. This procedure is based on a 5-percent sample for each batch (group of 1,000 abstracts) in which abstracts are independently coded by a second coder, with discrepancies adjudicated by a chief coder. Batches for

which the quality control sample have an error rate greater than 5 percent for medical data or 1 percent for nonmedical data are rejected and recoded. The overall error rate for records manually coded for the 1997 data year was 1.0 percent for final diagnoses, 0.7 for surgical and diagnostic procedures, and 0.2 percent for nonmedical data entry.

Automated Data Collection and Processing

Of the factors that influenced the redesign of the NHDS, the use of discharge data in computerized form has had the greatest impact on the ongoing operations of the survey. Because hospital participation is voluntary, NCHS cannot require hospitals or their intermediaries who supplied discharge data in machine readable form to comply with standard file and variable formats. And, with limited survey resources, NCHS cannot afford to have automated data sent to NCHS in a defined format. As a result, in many instances unique procedures for processing electronic files were developed, and these procedures have been subject to regular change as the sources of data changed. What became known as “automated” data collection in NHDS has been a dynamic process. This section provides background to and a general description of NHDS automated data collection and its processing.

Background

The concept of collecting data on health encounters through existing electronic systems was formalized in the 1970’s as part of the NCHS’s Cooperative Health Statistics System (CHSS). The CHSS operated under the legislative authority of the Health Services Research, Health Statistics, and Medical Libraries Act of 1974 (Public Law 93–353) (15). A major emphasis of the CHSS was national data collection through a “bottom-up” cooperative arrangement, a system in which data would be collected once, processed initially at the State level, and submitted to the Federal level in machine readable

form to provide national data. During the 1970’s, under the CHSS, NCHS funded hospital discharge demonstration projects in 12 States. This activity ended at NCHS in February 1979, by a decision of the Secretary of Health, Education, and Welfare (16).

A general evaluation study of the NHDS conducted in the mid-1970’s (17) recommended that the NHDS collect data from three frames: 1) hospitals covered by the CHSS, 2) non-CHSS hospitals that were members of a private abstracting system, and 3) all other hospitals for which no special data source was available. In order to implement a plan that incorporated automated data collection, two major issues needed to be addressed: the availability and quality of data from these sources. The evaluation study concluded that all the NHDS data items, with the exception of marital status, were widely available from automated sources.

Studies conducted in the late 1970’s on the reliability of hospital discharge abstract service data (18) and on the reliability of the NHDS data (19) showed that data collected from hospitals for discharges by abstract services met and in some cases exceeded standards of data quality traditionally held by NCHS. More recently there has been evidence of improvement in the general quality of hospital discharge data files as this information became important for determining reimbursement by Medicare and other payers (20–22).

NCHS funded two studies on the operational aspects of collecting data in electronic form from abstract services. The first study, initiated in 1980 (23), focused on the suitability of purchasing and using data from discharge abstract services in the NHDS. One of the findings from this study was that the purchase of data could greatly reduce, but not eliminate, the need for direct data collection. In 1981 approximately 60 percent of hospitals in the United States subscribed to discharge abstract service systems that serviced 10 or more hospitals. Without covering the universe of hospitals, these systems could not provide a representative sample of all hospitals in the United States.

Under the second study, NCHS initiated a demonstration project in 1984 (24) to develop procedures for acquiring, processing, and validating the quality of existing discharge data in electronic form. Under this contract automated data were collected from commercial abstract services for 76 hospitals that participated in the NHDS. Estimates from NHDS based on data collected by both means (manual and automated) produced comparable results (25). The study also demonstrated that abstracting service organizations could comply with data delivery schedules necessary for NHDS. Based on the results of this study, NCHS began to use data in automated form in the 1985 NHDS.

For the 3 years, 1985–87, NCHS acquired existing automated inpatient data for some NHDS hospitals under a contract with the Center for Health Policy Studies (CHPS). CHPS provided NCHS with a census file for all discharges for 78, 77 and 74 hospitals, respectively, for each of these years and NCHS sampled records from these files based on the sample design for manual data collection.

Beginning in 1988 purchasing and processing of data in automated form was conducted by NCHS. Automated data collection increased to include about 170 hospitals annually (table B) and the design specifications provided for a greater number of sample discharges from hospitals that provided data in machine readable form.

Data Sources

In 1987 NCHS requested client lists from 22 major hospital discharge abstracting services. These lists were used to identify the abstract service for hospitals sampled in the NHDS. Of the 542 hospitals sampled for the 1988 NHDS, 240 or 44.4 percent, used at least one of these services.

Changes in computer technology and the transition of automated data from commercial services to State-based systems resulted in a dynamic and evolving system for processing automated data. The mix of these sources changed dramatically from 1988 to 1997: by 1997, most of the 22

commercial abstract processing organizations used for the 1988 NHDS had either closed or transformed to a State-based system (hospital associations, State departments of health and State contractors). State-based systems accounted for approximately 95 percent of all machine readable data used in the 1997 NHDS. In 1997 the NHDS collected automated data from 20 different automated data systems and 12 individual hospitals directly. Of the 20 systems, 16 were State systems, 2 were corporate owned, and 2 were commercial abstract companies.

With some exceptions, automated data sources other than individual hospitals (abstract services, hospital associations, and State organizations) were bound by contracts with their client hospitals or by State laws to obtain permission from hospitals before they could release data to NCHS. This required NCHS to go through a two-step process in order to obtain data from these sources. First, NCHS had to secure permission from each hospital to allow the automated data source to release the hospital's data to NCHS. Second, NCHS had to negotiate and establish a formal relationship with each automated data source about issues related to purchasing data files. For purchasing data from individual hospitals, elements of the process were similar, but the process was conducted with a single entity—the hospital.

Data Release Agreements—Hospitals

To obtain data release agreements from hospitals, NCHS sends an introductory letter and packet of information to each hospital to solicit their participation in the survey. The packet contains a release agreement (appendix VIII), information about the NCHS and NHDS, an endorsement letter from the American Hospital Association (appendix III), selected endorsements from State hospital associations, and recent publications from the National Center for Health Statistics using data from the NHDS.

An initial telephone call by the NCHS staff is made to explain the release request and obtain the name of

the hospital administrator. If the release form is not returned within 2 weeks, the hospital receives a follow-up telephone call. The Census Bureau is advised of all hospitals that were late or refused the request, and they performed a final follow up with the hospital to obtain a release form. At this stage a meeting with the hospital administrator is usually required to obtain the hospital's cooperation. For individual hospitals that supply automated data directly to NCHS, this is the only agreement that is needed. For hospitals that agree to have a third party release their data to NCHS, a copy of the release agreement between NCHS and the hospital are forwarded to the organization that maintains the hospital's discharge data.

Data Release Agreements—Automated Discharge Data Sources

NCHS established standard guidelines and formal agreements (appendix IX) for purchasing hospital discharge data from medical record abstract service organizations. These agreements delineated unit record costs, startup programming costs, delivery schedules, tape formatting instructions, record layouts, field and data element definitions, and assurances of confidentiality. Confidentiality issues in these agreements work to each party's benefit: to ensure hospitals that NCHS would not release identifying information, and to protect NCHS from the automated data sources divulging the identity of hospitals that participated in the NHDS. Whenever possible releases were negotiated for multiple years or were open ended to reduce the administrative resources needed to renegotiate annually.

Beginning in 1993 NCHS incorporated nonconfidential data files from State discharge data sources. Hospital release agreements were not required from these sources because the data were not confidential and the data were for research and statistical use. For these files NCHS receives some variables in a modified form to ensure confidentiality. For example, exact admission and discharge dates are replaced with length of stay, and exact

birth date is replaced with the patient's age. Data were obtained for 110 hospitals from nonconfidential sources for the 1997 NHDS.

File Processing

Automated files of discharges and accompanying documentation are sent directly to NCHS. Individual hospitals also provide statistical information about the volume of discharges per month on a transmittal notice ([appendix X](#)) that is used in the estimation process.

NCHS developed a recommended file layout with variable structures ([appendix XI](#)) and processing procedures for submission of automated data. However, as a voluntary survey, NCHS does not have authority to require the submission of data, its structure, or the form of its submission. To keep the hospital response rate as high as possible, NCHS is flexible in its receipt of automated data. As a result, electronic files are sent to NCHS on various media (cartridge tapes, reel tapes, CD-ROM, diskettes, and as E-mail attachments); in ASCII or EBCDIC; in various layouts and blocking factors; for varying time frames (quarterly, semiannually, and annually); and often with unique variable structures. In addition, automated data files are accepted as a census or sample of discharges. Census files contain all discharges from hospitals and sampling is performed by NCHS (see *Statistical Design*). Sample files consist of discharges that sampled by the automated source prior to submission to NCHS based on sampling specifications provided by NCHS (see *Statistical Design*). In all, NCHS receives approximately 2,000,000 discharges in automated form each year from about 35 different sources on approximately 60 physical files.

The variety of file formats and data structures precludes a uniform system for processing. NCHS developed and regularly modifies a process to restructure variables to common definitions, reformat files to a common layout, and to evaluate the quality of all automated data. Evolution of the file processing system resulted in many adjustments to individual data files

depending on the problems encountered. Some activities or functions were used once for a single file while others have been used across years. The general characteristics of this process are described below. The goal was to produce files containing data of acceptable quality in a single format with identical variable definitions.

Each automated file is copied to a mainframe computer and compared with the previous year submission from that data source for changes in internal and external characteristics. If a file fails to copy correctly, further investigation is conducted using a program to display the encoding structure, density, internal labeling, blocking, and record length of the file, and also to detect the existence of extraneous data. If a file is physically unusable, NCHS requested a replacement and the process began anew.

After a file is determined to be physically usable, a specific program is used for each file that is not in the recommended NCHS format to reformat all data elements to a uniform format with uniform data structures. At this point automated files are examined to detect and remove duplicate records within hospitals.

Next, all ICD-9-CM codes on the file are subjected to an annually updated program that converts the ICD-9-CM addenda effective on October 1 of the survey year back to valid codes before October 1 of the survey year. This process was required to make fiscal year ICD-9-CM codes compatible with calendar year codes for the NHDS. Also, there were instances in which hospitals had not properly instituted the addenda from previous years and submitted files that contained codes that were no longer in use. A computer program was developed that changed all outdated codes from previous years to valid current-year codes. The first process, changing the addenda to NHDS survey year codes, affected thousands of codes annually because it covered codes from all automated discharges for the last quarter of the survey year; the second process involved very few codes and none in some years.

All records are checked for invalid or missing variables, inaccuracies,

outliers, and a detailed review is conducted for most variables for each hospital. For diagnoses and procedures codes, the review process generated a list of invalid ICD-9-CM codes, first-listed and all-listed diagnoses, the 25 most frequently occurring diagnoses and procedures by age group, and the distribution of discharges for all patients by chapters of the ICD-9-CM. For all other variables the review involves a comparison of distributions for the current year with data from previous years. This review detected potential problems. If the problems cannot be fixed at NCHS, a corrected file is requested.

The final product of these processing operations are files in a standard format with identical variable structures, valid survey-year ICD-9-CM codes, and data of acceptable quality. Using these files, data from hospitals that provided a census of discharges are merged to form a single file for sampling (see *Statistical Design*).

To avoid disruption of the flow of automated data for the NHDS, NCHS needs timely information about hospitals that changed or discontinued their abstract service. Early knowledge of these changes provides time to obtain the necessary data release agreements with these hospitals and their new abstract service in time to process the data. To learn about these changes, NCHS initiated semiannual contacts with each automated hospital as recommended (24). In its initial form this process was envisioned to be conducted by sending postcards to hospitals. The hospitals would complete the questions and mail the self-addressed, stamped postcard to NCHS. However, using actual postcards could not be accomplished administratively, so letters are used ([appendix XII](#)).

This activity has evolved to provide a measure of the completeness of data being collected for each automated hospital. In 1997 it was conducted semiannually to get information directly from automated hospitals to validate their total number of discharges, numbers of newborns, and total hospital beds.

Editing

Extensive computerized edits are applied to each record. The computer edits are augmented by a manual review of selected records rejected by the edit program. The edits include validity and range checks for nonmedical variables, and consistency checks for dates of admission, discharge and birth. The medical data (diagnoses and procedures) are verified against a list of all valid ICD-9-CM codes. Invalid ICD-9-CM codes are removed. The edit performs a series of consistency tests from decision logic tables for sex-specific and age-specific ICD-9-CM codes. When the sex or age of a patient is incompatible with the recorded medical information, priority is given to the medical information.

With two exceptions, the order of diagnoses and procedures for sampled discharges is preserved to reflect the order on the medical record face sheet or in the automated file. One exception is for women admitted for delivery for whom a code of V27 from the supplemental classification is assigned as the first-listed code in order to provide an estimate of all deliveries. In the other exception, whenever an acute myocardial infarction is present with another circulatory diagnosis and is other than the first-listed diagnosis, it is reordered to the first position.

The computerized edit process identifies some records for manual review: records with ICD-9-CM codes that are not valid as first-listed or single-listed codes; all records with lengths of stay longer than 100 days; records that indicate a delivery with a length of stay over 30 days; and all records with age equal to or greater than 100 years. A nosologist at NCHS examines these records and makes changes as appropriate to eliminate gross inconsistencies. In 1997, 276 records were identified for manual review.

Imputation

Before 1996 missing values of age and sex were imputed by using random assignment with female being assigned first for sex and a decision-logic table

was used to impute age. Beginning with the 1996 NHDS a hot deck method for missing values of age and sex has been used that maintains the known age or sex distribution of records within the same 3-digit level of first-listed ICD-9-CM diagnostic code.

Estimation Procedures

Statistics from the NHDS are derived by a multistage estimation procedure that produces essentially unbiased national estimates and has three basic components: inflation by reciprocals of the probabilities of sample selection, adjustment for nonresponse, and population weighting ratio adjustments. The second and third components are made separately by admission types—that is, for discharges of newborn infants (whose hospital stay began with their own birth) and for discharges other than newborn infants.

Inflation by Reciprocals of Probabilities of Selection

The first two probabilities of the three-stage sample design for the NHDS are known and fixed (see Statistical Design). However, the number of discharges for individual months is not always constant or an integral multiple of the sampling fraction. For these reasons, in the estimation process, this probability is the ratio of the actual number of discharges per month to the number of sampled discharges for the month.

Adjustment for Nonresponse

NHDS estimates are adjusted to account for nonresponse at two levels: hospitals and discharges within hospitals. Within hospital adjustments are made for nonresponding months and within months for nonresponding discharges.

Hospital nonresponse occurs when an in-scope (NHDS-eligible) sample hospital does not respond for at least

one-half of the months during which it was in scope. In this case, the weights of discharges from hospitals similar to the nonrespondent hospitals are inflated to account for discharges represented by the nonrespondent hospitals. For this purpose, hospitals are judged to be similar if they are in the same region, hospital specialty-size group, and if possible, the same sampling stratum (that is, the same abstracting status group if the nonrespondent hospital was in the 12 largest PSU's and in the same PSU, otherwise).

The adjustments for this nonresponse are made separately for admission types—that is, for discharges of newborn infants and for all other discharges. The adjustment consists of a ratio for which the numerator is the weighted number of discharges of the admission type in all similar sample hospitals (regardless of response status) and the denominator is the weighted total of discharges of that admission type from the respondent hospitals similar to the nonrespondent hospitals. Data on the number of discharges for each admission type for each hospital come from either the hospitals or the April SMG Hospital Market Database for the survey year.

At the discharge level, responding hospitals have a month(s) of nonresponse when the number of discharges received for the month was less than one-half of the expected number of discharges. For a hospital's month(s) of nonresponse, the weights of discharges in the hospital's respondent months are inflated by ratios that varied with discharge groups defined by the ICD-9-CM diagnostic classes of those discharges' first-listed diagnoses. The adjustment ratio for each partially respondent hospital and each discharge group is calculated using only data from sample hospitals that were NHDS eligible and respondent for all 12 months of the data year. The ratio has as its numerator the weighted sum of discharges in that discharge group for all months in which the partially respondent hospital was in scope. The ratio has as its denominator the weighted sum of discharges in that discharge group that occurred in the months when the

partially respondent hospital did respond to the NHDS.

Discharge nonresponse also occurs when NCHS fails to collect all of the discharge abstracts expected for a month (the number expected was the product of the hospital's total discharges each month and the discharge sampling rate assigned to the hospital). In each month when the hospital was respondent (at least one-half of the expected abstracts were collected), the weights of abstracts collected for the month are inflated to account for the missing abstracts.

Population Weighting Ratio Adjustment

The final adjustment consists of a population weighting ratio adjustment that is applied separately for newborns and for other than newborns. These adjustments are made within each of 16 noncertainty hospital groups defined by region and hospital specialty-size classes to adjust for oversampling or undersampling of discharges reported in the sampling frame for the data year. For discharges other than newborn infants, the adjustment is a multiplicative factor that has as its numerator the number of admissions reported for the year at sampling frame hospitals within each region-specialty-size group and as its denominator the estimated number of those admissions for that same hospital group. The adjustment for discharges of newborn infants is similar, but numbers of births are used in place of admissions. The ratio numerators are based on the figures obtained annually from the SMG Hospital Market Database and the ratio denominators are obtained through a simple inflation of the SMG figures for the NHDS sample hospitals. NCHS performs a review for outliers found by comparing the SMG Hospital Market Database file for the current and previous years. The object of this review is to allow detection of potential systematic problems and to correct the file prior to weighting.

Reliability of Estimates

Because the statistics from the survey are based on a sample, they may be different from the figures that would have been obtained if a complete census had been taken using the same forms, definitions, instructions, and procedures. However, the probability design of the NHDS permits the calculation of sampling errors. The standard error of a statistic is primarily a measure of sampling variability that occurs by chance because only a sample rather than the entire population is surveyed. The standard error, as calculated for the NHDS, also reflects part of the variation that arises in the measurement process, but does not include any systematic bias that may be in the data. The relative standard error (RSE) of an estimate is obtained by dividing the standard error by the estimate itself and when multiplied by 100 is expressed as a percent of the estimate. Generally, in tables of NCHS published data reports, an asterisk (*) is used to indicate the relative unreliability of estimates based on fewer than 30 records or records that have a relative standard error greater than 30 percent.

In repeated samples using the same forms and procedures, the chances are about 68 in 100 that an estimate from the sample would differ from a complete census by less than the standard error. The chances are about 95 in 100 that the difference would be less than twice the standard error, and about 99 in 100 that it would be less than 2.5 times as large.

Estimation of Standard Errors

Beginning with the 1988 NHDS, estimates of sampling variability for the NHDS statistics presented in NCHS publications have been computed using software that produces error estimates for statistics from complex sample

surveys. The software employs a first-order Taylor Series approximation of the deviation of estimates from their expected values. A description of this software and the approach it uses has been published (6,7).

Standard Error Approximations

Parameters for calculating approximate relative standard errors for aggregate estimates are presented in [table C](#). To derive error estimates that would be applicable to a wide variety of statistics, numerous estimates and their variances are produced. A least-squares method is then used to produce best-fit curves, based on the empirically determined relationship between the size of an estimate X and its relative variance. The relative standard error of an estimate X [RSE(X)] is the square root of the relative variance and may be calculated from the formula:

$$RSE(X) = \sqrt{a + b/X}$$

with a and b provided in [table C](#). When multiplied by 100, the RSE(X) is expressed as a percent of X .

For example, in 1997 the estimated number of discharges from short-stay hospitals for females with a first-listed diagnosis of atherosclerotic heart disease (ICD-9-CM code 414.0) was 384,000. Using the applicable constants from [table C](#) for estimates by sex produces:

$$RSE(384,000) =$$

$$\sqrt{.00127 + (325.984/384,000)}$$

$$RSE(384,000) = .046$$

When multiplied by 100, the relative standard error for the estimate of interest becomes 4.6 percent. The standard error of the estimate is obtained by multiplying the relative standard error by the estimate itself:

$$SE(384,000) = 384,000 * .046 = 17,664$$

The standard error can be employed to generate confidence intervals for statistical testing. In this example, the 95 percent confidence interval for the

Table C. Estimated parameters for approximate relative standard error equations, by selected characteristics: National Hospital Discharge Survey, 1997

	First-listed diagnosis		Days of care		All-listed diagnosis		All-listed procedures	
	<i>a</i>	<i>b</i>	<i>a</i>	<i>b</i>	<i>a</i>	<i>b</i>	<i>a</i>	<i>b</i>
Total	0.00132	313.983	0.00242	1,036.029	0.00271	428.987	0.00288	323.247
Sex								
Male	0.00148	335.128	0.00299	1,309.507	0.00214	307.917	0.00268	380.213
Female	0.00127	325.984	0.00252	1,060.275	0.00127	311.394	0.00197	299.725
Age group								
Under 15 years	0.01470	181.262	0.02393	346.675	0.01617	223.921	0.02639	196.719
15–44 years	0.00137	294.357	0.00285	934.043	0.00151	309.115	0.00219	305.574
45–64 years	0.00138	301.320	0.00284	1,248.476	0.00290	370.245	0.00235	298.267
65 years and over	0.00147	343.779	0.00275	1,866.761	0.00144	361.717	0.00239	292.252
Region								
Northeast	0.00384	195.564	0.00863	495.447	0.00592	239.709	0.00695	219.126
Midwest	0.00609	191.492	0.00866	481.071	0.00853	180.424	0.00870	163.190
South	0.00369	320.084	0.00572	1,581.301	0.00307	341.005	0.00498	274.548
West	0.00513	338.516	0.01008	926.285	0.00508	378.827	0.00663	288.835
Race								
White	0.00300	314.704	0.00480	1,103.096	0.00357	314.530	0.00463	369.830
Black	0.00529	248.048	0.00822	875.877	0.00519	236.801	0.00662	225.180
All other	0.01770	200.033	0.02524	581.314	0.01615	224.604	0.02070	211.065
Race not stated	0.01973	196.517	0.02387	394.423	0.02002	187.166	0.02126	142.769
Expected source of payment								
Worker's compensation	0.00721	320.711	0.01413	896.903	0.01093	337.073	0.01143	270.432
Medicare	0.00171	325.698	0.00301	1,822.158	0.00157	372.529	0.00264	297.577
Medicaid	0.00438	293.366	0.00709	685.261	0.00379	287.541	0.00540	282.342
Payment not stated	0.00962	313.407	0.01966	1,484.182	0.01219	310.942	0.01814	248.591
Other government payments	0.00181	278.864	0.00345	617.794	0.00199	299.815	0.00265	289.092
Private insurance	0.00373	265.383	0.00994	854.606	0.00415	281.681	0.00689	227.293
Self-pay	0.02771	94.095	0.03240	244.803	0.03212	127.947	0.03297	94.813
No charge/other	0.01828	394.867	0.02548	1,488.967	0.01931	386.315	0.02059	388.354

estimate of female inpatients with a first-listed diagnosis of atherosclerotic heart disease is:

$$(384,000 - 2 * 17,664) <-> (384,000 + 2 * 17,664) \\ 348,700 <-> 419,300$$

Relative Standard Error for Estimates of Percents

Approximate relative standard errors for estimates of percents may be calculated from [table C](#) also. The relative standard error for a percent, $100p$ ($0 < p < 1$) may be calculated using the formula:

$$RSE(p) = \sqrt{b * (1-p) / (p * X)}$$

where $100p$ is the percent of interest, X is the base of the percent, and b is the parameter b in the formula for

approximating the $RSE(X)$. The values for b are given in [table C](#). When multiplied by 100, the $RSE(p)$ is expressed as a percent of the estimate, p .

For example, in 1997 the estimated number of discharges from short-stay hospitals that were female was 18,647,000. This is 60.3 percent of the estimated 30,914,000 discharges for that year. Using the applicable constants from [table C](#) for estimates by sex produces:

$$RSE(.603) =$$

$$\sqrt{325.984 * (1 - .603) / (.603 * 30914000)}$$

$$RSE(.603) = .002635$$

When multiplied by 100, the relative standard error for the estimate of interest becomes .2635 percent. From this the standard error is obtained by

multiplying the relative standard error by the estimate itself:

$$SE(.603) = .603 * 0.002635 = .0016$$

The standard error can be employed to generate confidence intervals for statistical testing. In this example, the 95 percent confidence interval for the estimate of the percent of female inpatients is:

$$(.603 - 2 * .0016) <-> (.603 + 2 * .0016) \\ .5998 <-> .6062$$

or, equivalently, 59.98 percent <-> 60.62 percent

The SUDAAN software can be used to compute specific standard errors for the NHDS estimates from which relative standard errors may be derived. The files needed to use SUDAAN have not been publicly released because they contain information that is confidential (however, recently these files have become available for analysis, see Data

Dissemination). The generalized procedure for approximating relative standard errors for the NHDS described above is widely applicable for a variety of statistics and cost efficient. As a result, standard errors computed from the generalized curves should be interpreted as approximate rather than exact for any specific estimate.

Nonsampling Errors

Estimates from the NHDS are subject to nonsampling as well as sampling errors, as are estimates from any survey. For the NHDS nonsampling errors may include those due to sampling frame errors, hospital nonresponse, missing abstracts, and errors introduced at the time of data collection or electronic data entry. Although the magnitude of the nonsampling errors cannot be computed, these errors were kept to a minimum by procedures built into the operation of the survey.

Errors resulting from the exclusion of in-scope hospitals from the sampling frame are believed to be small because the hospitals excluded were hospitals that opened after the frame was constructed and, hence, they tend to have few discharges relative to hospitals that were in the frame. Other nonsampling errors were kept to a minimum by training the data collectors in sampling and data abstraction, quality control procedures, and edit checks discussed in the data processing section. Because the survey results are subject to sampling and nonsampling errors, the total error is larger than the error due to sampling variability alone.

Data Dissemination

Data are available for all years of the survey, but the mechanisms for data release have changed to reflect the needs of data users and new technology. This section summarizes the current methods of data release for the NHDS.

Summary data from NHDS have been published annually for three levels of statistical detail: the *Advance Data*

From Vital and Health Statistics (26) utilizes a brief format for reporting of inpatient statistics as soon as possible after the survey has been completed; more detailed summary statistics are available from Series 13 *Vital and Health Statistics* (27); and discharge estimates for individual ICD-9-CM diagnostic and procedure codes have been published in Series 13, *Vital and Health Statistics* (28). Series 13 *Vital and Health Statistics* reports that combined ambulatory surgery and surgery performed on hospital inpatients for 1994, 1995, and 1996 are available (29-31). In addition, there are reports comparing hospital use in the United States and other countries under Series 5, *Comparative International Vital and Health Statistics Reports*; trend reports under Series 13, *Vital and Health Statistics*; and, methodological reports using hospital survey data published in Series 2, *Vital and Health Statistics* (32). Reports on topics of special interest have been published in *Advance Data From Vital and Health Statistics*, Series 13 *Vital and Health Statistics* reports, and in published journal articles and papers presented at professional meetings (33-34). As resources permit, special tabulations and analyses are provided to data requestors inside and outside the Federal Government.

For each survey year, a public-use data tape and documentation are prepared for distribution through the National Technical Information Service (NTIS). NCHS has also produced a multiyear data file (1979-97), which is available on data tape and CD-ROM. For analysts interested in going beyond what is available in the public use files, NCHS has established a Research Data Center that allows the use of SUDAAN and linking NHDS data with other sources of information.

A Catalog of Publications and a Catalog of Public-Use Data Tapes from the National Center for Health Statistics are available via the Internet (at: www.cdc.gov/nchswww/products/products.htm) or from NCHS (Data Dissemination Branch, National Center for Health Statistics, 6525 Belcrest Road, Room 1064, Hyattsville, Maryland 20782; Telephone (301) 436-8500). Many of the publications,

data files and file documentation mentioned above are available at this Internet site.

References

1. Origin, Program, and Operation of the U.S. National Health Survey. National Center for Health Statistics. *Vital Health Stat* 1(1). 1965.
2. Woolsey TD. "Hospital Statistics from the U.S. National Health Survey." Presentation for Annual Conference for Executives of Allied Hospital Association, Grand Hotel, Mackinac Island, Michigan. July 1963.
3. U. S. Bureau of the Census. Statistical abstract of the United States: 1997. 117th ed. Washington: 1997.
4. Simmons WR. Development of the design of the NCHS Hospital Discharge Survey. National Center for Health Statistics. *Vital Health Stat* 2 (39). 1970.
5. Massey JT, Moore TF, Parsons VL, Tadros W. Design and estimation for the National Health Interview Survey, 1985-94. National Center for Health Statistics. *Vital Health Stat* 2(110). 1989.
6. Shah BV. SESUDAAN: Standard errors program for computing of standardized rates from sample survey data. Research Triangle Park, North Carolina. Research Triangle Institute. 1981.
7. Shah BV, Barnwell BG, Hunt PN, LaVange LM. SUDAAN user's manual, release 5.50. Research Triangle Park, North Carolina: Research Triangle Institute. 1991.
8. Haupt BJ, Kozak LJ. Estimates from two survey designs: National Hospital Discharge Survey. National Center for Health Statistics. *Vital Health Stat* 13(111). 1992.
9. Shimizu I, Cole C. A comparison of precision between two- and three-stage designs for National Hospital Discharge Survey. *Proc Amer Statist Ass. Survey Research Methods Section*. 1991.
10. SMG Hospital Marketing Group, Inc. Hospital market database. Healthcare Information Specialists. Chicago, Illinois: 1987.
11. Moien M. The use of hospital discharge survey data to produce trend estimates of hospital utilization. *Proc Amer Statist Ass. Social Statistics Section*. 1978.
12. National Center for Health Statistics. International Classification of Diseases,

- 9th Revision, clinical modification. Washington: Public Health Service. 1980.
13. Report of the U.S. Committee on Vital and Health Statistics. Uniform hospital abstract: minimum basic data set. National Center for Health Statistics. Vital Health Stat. 4 (14). 1975.
 14. Department of Health and Human Services. Health information policy council: 1984 revision of the uniform hospital discharge data set. Office of the Secretary. Federal Register; vol 50 no 147. 1985.
 15. Task Force on Definitions. The cooperative health statistics system: Its mission and program. National Center for Health Statistics. Vital Health Stat. 4 (19). 1977.
 16. Final of the Panel to Evaluate the Cooperative Health Statistics System. Directions for the 80's. National Center for Health Statistics. Hyattsville, Maryland: 1980.
 17. Lissler JT. Hospital discharge survey evaluation study. Contract NCHS 255U-1178. Research Triangle Park, North Carolina: Research Triangle Institute. 1977.
 18. Institute of Medicine. Reliability of hospital discharge abstracts. Washington: National Academy of Sciences. 1977.
 19. Institute of Medicine. Reliability of National Hospital Discharge Survey data. Washington: National Academy of Sciences. 1980.
 20. Green J, Wintfeld N. How accurate are hospital discharge data for evaluating effectiveness of care? *Med care* 31:719-31. 1993.
 21. Fisher ES, Whaley FS, Krushat, et al. Accuracy of Medicare's hospital claims data: Progress has been made, but problems remain. *Am J Public Health* 82:243-8. 1992.
 22. Hsia DC, Ahern CA, Ritchie, et al. Medicare reimbursement accuracy under the prospective payment system, 1985 to 1988. *JAMA* 268:896. 1992.
 23. Duggar BC, et al. Alternative sources for producing national statistics on hospital utilization. Contract NCHS 223-81-2057. JRB Associates. 1981.
 24. Miller H, et al. Optional methods for collecting data for the National Hospital Discharge Survey. Contract NCHS 282-84-2107. Center for Health Policy Studies. 1984.
 25. Tompkins L. Comparison of manually abstracted and automated abstract service data in the 1984 NHDS. Unpublished paper prepared by the Office Research and Methodology for review within NCHS. 1989.
 26. Graves GJ, Owings MF. 1996 Summary: National Hospital Discharge Survey. Advance data from vital and health statistics; no 301. Hyattsville, Maryland: National Center for Health Statistics. 1998.
 27. Graves GJ, Kozak LJ. National Hospital Discharge Survey: Annual summary, 1996. National Center for Health Statistics. Vital Health Stat 13(140). 1999.
 28. Graves GJ, Kozak LJ. Detailed diagnoses and procedures, National Hospital Discharge Survey 1996. National Center for Health Statistics. Vital Health Stat 13(138). 1997.
 29. Pokras R, Kozak LJ, McCarthy E. Ambulatory and inpatient procedures in the United States, 1994. National Center for Health Statistics. Vital Health Stat 13(132). 1997.
 30. Kozak LJ, Owings MF. Ambulatory and inpatient procedures in the United States, 1995. National Center for Health Statistics. Vital Health Stat 13(135). 1998.
 31. Kozak LJ, Owings MF. Ambulatory and inpatient procedures in the United States, 1996. National Center for Health Statistics. Vital Health Stat 13(139). 1998.
 32. Harris KW, Hoffman KL. Quality control in the Hospital Discharge Survey. National Center for Health Statistics. Vital Health Stat 2(68). 1975.
 33. Pappas G, Hadden WC, Kozak LJ, Fisher GF. Potentially avoidable hospitalizations: Inequalities in rates between U.S. socioeconomic groups. *Am J Public Health* 87(5):811-6. 1997.
 34. Wilcox SW, Koonin LM, Pokras R, et al. Hysterectomy in the United States, 1988-1990. *Obstet Gynecol* 83(4):549-55. 1994.

Appendix I.

Legislative Authorization

Public Health Service Act
Section 306(a) & (b)

NATIONAL CENTER FOR HEALTH STATISTICS

Section 306. [242k](a) There is established in the Department of Health and Human Services the National Center for Health Statistics (hereinafter in this section referred to as the “Center”) which shall be under the direction of a Director who shall be appointed by the Secretary and supervised by the Assistant Secretary for Health (or such officer of the Department as may be designated by the Secretary as the principal adviser to him for health programs).

(b) In carrying out section 304(a), the Secretary, acting through the Center-

(1) shall collect statistics on-

(A) the extent and nature of illness and disability of the population of the United States (or any groupings of people included in the population), including life expectancy, the incidence of various acute and chronic illnesses, and infant and maternal morbidity and mortality,

(B) the impact of illness and disability of the population on the economy of the United States and on other aspects of the well-being of its population (or of such groupings),

(C) environmental, social, and other health hazards,

(D) determinants of health,

(E) health resources, including physicians, dentists, nurses, and other health professionals by specialty and type of practice and supply of services by hospitals, extended care facilities, home health agencies, and other health institutions,

(F) utilization of health care, including utilization of (i) ambulatory health services by specialties and type of practice of health professionals providing such service, and (ii) services of hospitals, extended care facilities, home health agencies, and other institutions,

(G) health care costs and financing, including the trends in health care prices and costs, the sources of payments for health care services, and Federal, State, and local governmental expenditures for health care services, and

(H) family formation, growth, and dissolution;

(2) shall undertake and support (by grant or contract) research, demonstrations, and evaluations respecting new or improved methods for obtaining current data on the matters referred to in a paragraph (1);

(3) may undertake and support (by grant or contract) epidemiologic research, demonstrations, and evaluations on the matters referred to in paragraph (1); and“

(4) may collect, furnish, tabulate, and analyze statistics, and prepare studies, on matters referred to in paragraph (1) upon request of public and nonprofit entities under arrangements under which the entities will pay the cost of the service provided. Amounts appropriated to the Secretary from payments made under arrangements made under paragraph (4) shall be available to the Secretary for obligation until expended.

Appendix II

Induction Letter



DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
Centers for Disease Control and Prevention

National Center for Health Statistics
6525 Belcrest Road
Hyattsville, Maryland 20782

Dear

For more than 30 years, the National Center for Health Statistics of the Centers for Disease Control and Prevention has administered the National Hospital Discharge Survey, the Nation's primary resource for monitoring the utilization of short-term, acute care hospitals. This voluntary survey has served as a valuable source of information for policymaking, health care research, academic education and various applications within the hospital industry. Because of its importance, the survey has had the continuing support of the American Hospital Association and the American Health Information Management Association.

The purpose of this letter is to request your participation in this survey. The National Center for Health Statistics will make every effort to facilitate the collection of discharge data from your hospital. An equitable basis for compensation will be arranged. All information will be held in strict confidence and will be used for statistical purposes only, as is required by Section 308(d) of the Public Health Service Act. All published summaries will be presented in such a way that no individual hospital or patient can be identified. Your participation is voluntary and there is no penalty for not participating in the survey; however, the success of this survey rests on the willingness of health professionals like you to provide current medical information.

We would like very much to discuss matters relating to the participation of your hospital in the National Hospital Discharge Survey. Therefore, within the next several days, a representative of the Bureau of the Census, acting as an agent of the National Center for Health Statistics, will telephone you to arrange for an appointment. This meeting should take less than an hour of your time. We have enclosed a packet that provides valuable information on the survey. In it you will find a description of the survey and copies of some of our latest published reports.

Your cooperation in this survey will be very much appreciated.

Sincerely yours,

Robert Pokras, Chief
Hospital Care Statistics Branch
Division of Health Care Statistics

Enclosures

Appendix III

American Hospital Association Endorsement Letter

American Hospital Association



840 North Lake Shore Drive
Chicago, Illinois 60611
Telephone 312.280.6000
Cable Address AMHOSP

January 1991

TO: Chief Executive Officer

SUBJECT: National Center for Health Statistics
National Hospital Discharge Survey

The National Center for Health Statistics has been conducting the National Hospital Discharge Survey since 1965. The objective of the survey is to produce statistics which are representative of patients discharged from all non-Federal short-stay hospitals. The results from this important survey are used to determine hospital utilization by age and sex of patients, their lengths of stay, and their diagnoses, surgical and non-surgical procedures. Diagnostic data are also classified by DRG. These data are used by hospitals, government agencies, and health care researchers to monitor and improve hospital care.

The American Hospital Association has reviewed the goals and methods of the survey and we endorse the study. As one of the hospitals selected nationwide, we encourage your hospital to participate.

Thank you for your cooperation.

Peter D. Kralovec
Director
Hospital Data Center

Appendix IV

Confidentiality Letter



DEPARTMENT OF HEALTH AND HUMAN SERVICES
OFFICE OF THE SECRETARY
ROCKVILLE, MARYLAND 20857

OCT 16 1981

OFFICE OF THE GENERAL COUNSEL

To Whom It May Concern:

With reference to participation in the National Hospital Discharge Survey conducted by the Census Bureau for the National Center for Health Statistics (NCHS), the question of potential liability for violations of patient privacy has been raised by a few hospitals. We wish to assure those concerned that, in our opinion, there would be no basis for hospital liability arising out of the disclosures required of participants in the Hospital Discharge Survey.

Our conclusion is based on the fact that the patient information to be provided by hospitals does not contain any patient names, addresses, social security number, or other information which would permit identification of a patient with reference to other publicly available information. Participating hospitals supply a medical record number for each patient so that those conducting the survey may discuss particular records with the hospital from which the records came, if they need to verify or add to the data received on certain response forms. Since the information to be disclosed will not be in identifiable form, there is no reasonably foreseeable harm to individual patients in the disclosures and accordingly no liability would lie against the hospital. 1/

In addition, it is our understanding that the Census Bureau physically safeguards the data and its research staff with access to the data are bound by statutory confidentiality restrictions 13 USC §§ 8 and 9 governing the Bureau, and 42 USC 242m governing NCHS activities.

1/ See Park v. Consolidated Mutual Insurance Co., 423 F.2d 41 (3rd Cir. 1970) for case requiring a showing of loss as well as improper disclosure for court to give a remedy. See Hospital Law Manual, Volume for Attorneys "Medical Records" ¶3-22, pp. 58 and 60 citing Prosser Law of Torts §117 (4th ed. 1971) for discussion of the nature of the harm required to establish invasion of privacy. See also general discussion of tort liability in 74 Am. Jur. 2d. "Torts" §§ 6 and 10 (1974).

We also note that even where disclosures of information from confidential records are identifiable, various courts have upheld reasonable disclosures of otherwise privileged and protected information. See, for example: Whalen v. Roe, 429 U.S. 589 (1977) (This case upholds disclosures for a purpose in which there is a substantial public interest and which disclosures are circumscribed by statutory confidentiality protections.); and Dupont v. Finklea, 442 F.Supp. 821 (1977). (A Federal District Court found no violation of privacy rights where there was no showing that upon disclosure protected medical records would be used improperly.) See also 61 Am. Jur. 2d "Limitation of nondisclosure privilege because of public policy considerations" §170 (1981).

In conclusion, we believe hospitals may participate in this survey without concern about liability. Hospitals are encouraged to discuss the reasoning in this communication with their local legal counsel.

Susan Greene Merewitz

Susan Greene Merewitz
Senior Attorney
Office of the General Counsel
Public Health Division

Part A – ADMINISTRATIVE OFFICE – Continued

5. We have the hospital name and address as (Read label on cover page). Is this correct?

- Yes (6) No – Correct label, then 6

6. Is this a _____ hospital?
(Ownership)

- Yes (7a) No – What is the form of ownership?
- Government Nonprofit
 Proprietary Other – Specify Church related

7a. Is this a _____ hospital?
(Service)

- Yes No – Which one of the following best describes the hospital's service?
- Children's Eye, ear, nose, and throat
 Orthopedic Alcohol or drug dependency
 Maternity Other – Specify General (7b)

Ask for "General" hospitals only, then 8

b. Are any services excluded such as obstetrics or pediatrics?

- Yes – What services? No

8. How many hospital beds are maintained; that is, staffed for inpatient use, excluding "newborn" bassinets?

Current number

Number in 19 ____

Compare with number in 19 _____

- Same Different – Specify the bed size as of January 1 and the date and amount of each subsequent change

9a. How many discharges were there last year (19 ____) from the entire hospital, including deaths and newborns?

Discharges

Estimate

b. How many live births were there last year (19 ____)?

Live births

Estimate

Notes

Part A — ADMINISTRATIVE OFFICE — Continued

10a. What was the average length of stay for all patients last year? Days
 Estimate

b. Does this hospital contain more than one unit?

- Yes (10c) No (11)

c. Specify the name of each unit for which separate medical records are kept.

Name of unit (a)	Number of beds (b)	Number of discharges (c)	Average length of stay (d)	In-scope unit (e)

11. When a patient moves from one unit, section, or service of this hospital to another unit, section, or service, is it always recorded as a transfer or are there circumstances under which it would be recorded as a discharge and readmission?

- Always transfer (12) Sometimes discharge and readmission — Record details, then 12
-

12. Are patients ever transferred between this hospital and any other hospital without being discharged?

- Yes — Explain, then 13 No (13)
-

13. Is computerization of medical records in effect or planned?

- Yes — Describe, then 14 No (14)
-

14. Does this hospital subscribe to a private abstracting service?

- Yes — Specify name No (15)
-

CHECK ITEM

Refer to hospital key and abstract service list.

- Core hospital (Code 1) — Go to question 15
- Noncore hospital — Abstract service on list — Stop and complete Abstract Service Questionnaire
- Noncore hospital — Abstract service not on list — Go to question 15

Part A — ADMINISTRATIVE OFFICE — Continued

15. Take out the Memorandum of Agreement Form.

The National Center for Health Statistics pays _____ per completed abstract.

Is this acceptable to you?

- Yes — Fill agreement as appropriate No — Negotiate acceptable rate and fill agreement as appropriate

Reimbursement Rates — Offer in order of priority:

- | | |
|---|-----------------------------|
| 1. Primary — \$1.00 per abstract (uniform) | } To be approved
by NCHS |
| 2. Primary — \$1.01 up to \$3.00 per abstract | |
| 3. Alternate — \$20.00 per 100 abstracts (uniform) | |
| 4. Alternate — \$21.00 up to \$100.00 per 100 abstracts | |

ASK THE ADMINISTRATOR TO SIGN THE MEMORANDUM OF AGREEMENT.

Notes

16. HOSPITAL PERSONNEL PRESENT DURING INTERVIEW

Name	Title

The rest of my questions concern your medical records department and record keeping practices. I can go over these now or with your Director of Medical Records, whichever you prefer.

(If leaving, thank the Administrator for his/her cooperation — Go to Part B.)

Part B — MEDICAL RECORDS DEPARTMENT

1. *(Read if necessary)* — The National Hospital Discharge Survey is designed to provide national statistics on a continuing basis for hospitalizations in short-stay hospitals. It involves the collection of a limited amount of information from the records of a sample of discharged patients. Data abstracted from medical records are items usually found on the face sheet or discharge summary of the medical record. These are the sampling and abstracting forms which will be used. *(Show HDS-5 and HDS-1)*

For this survey, we need to determine the information needed to locate the medical records for the sample cases.

2a. Is the medical record numbering system used in this hospital serial, unit, or some other system?

Serial Unit Serial-unit

Other — Describe _____

b. Does the numbering system include outpatients?

Yes — How can they be identified? — Explain, then 2c ↘ No (2c)

c. Does this numbering system include patients who are admitted and discharged on the same day, such as for ambulatory surgery, diagnostic testing, dialysis, and so forth?

Yes — Which types of services are included? How can they be identified from other inpatients? *(Explain, then 3)* ↘ No — What listing is maintained for patients who do not stay overnight? *(Explain, then 2d)* ↘

d. Is there any way to distinguish between the various reasons for these visits, such as for ambulatory surgery, diagnostic testing, dialysis, or routine outpatient visits?

Yes — How? _____
 No (3a) _____

Notes

Part B – MEDICAL RECORDS DEPARTMENT – Continued

3a. Does a newborn infant get a medical record number?

- Yes (3b) No(3c)

b. Does a newborn infant receive a different number from that of the mother?

- Yes (3d) No (3c)

c. Are there any circumstances in which a (number/different number) is assigned to a newborn, such as if the baby is transferred to another unit in the hospital or is discharged at a time different from the mother's discharge, or is born before the mother is admitted to the hospital?

- Yes – Give details, then 3d No (3d)

d. Is the newborn's record filed with the mother's record?

- Yes (3e) No (4)

e. Is there a face sheet for the newborn separate from that of the mother?

- Yes No

Notes

Part B – MEDICAL RECORDS DEPARTMENT – Continued

4. Do all inpatient units of the hospital use the same face sheet?

- Yes — Ask for two samples No — What face sheets are used? — Ask for two samples of each and explain where used ↴

Transcribe captions exactly as they appear on face sheet.

5a. Under which captions are discharge (final) diagnoses listed?

b. Under which captions are complications listed?

c. Under which captions are surgical and diagnostic procedures listed?

6. Please state the best source in the medical record to obtain the data for the following items:
(Mark or specify as appropriate)

Item	Face Sheet	Discharge summary	Other(s) (Specify)	Item not available
Name of patient				
Medical record number				
Date of admission				
Date of discharge				
Residence ZIP Code				
Date of birth				
Age				
Sex				
Marital status (Explain codes)				
Race				
Ethnicity (Hispanic/Nonhispanic data)				
Expected sources of payment				
Discharge status/disposition				
Final diagnoses and procedures				
Dates of procedures				

Part B – MEDICAL RECORDS DEPARTMENT – Continued

7. In order to select the sample, we will need to work with a list of inpatients of this hospital who were discharged during a given month. This list must include date of discharge, medical record number, and name, if needed to locate records.

Is there ONE discharge listing for all discharges from all short-stay parts of this hospital, including inpatient deaths and newborns (for *each in-scope unit in A-10*)?

- Yes – What is the list called? No – What discharge lists are kept?
 Complete columns below. Complete columns below for EACH list.

Name of list (a)	Mark "X" for each "Yes"									How long kept? (k)	Where kept? (l)	Is there any other discharge list? (Complete all columns for each "Yes") (m)	
	Does it show – ?				Does it use – ?		Does it include – ?					Yes	No
	Date of admission (b)	Date of discharge (c)	MR No. (d)	Name (e)	Sequential or serial (f)	Unit (g)	Out-patient (h)	Same day (i)	Only overnight patient (j)				

8. Are changes ever made to the discharge list(s) other than spelling?

- Yes – What kind of changes? No

Notes

Part B – MEDICAL RECORDS DEPARTMENT – Continued

11a. Is there a specific time period for completing or signing the medical record?

Yes – *Specify, then 11b* ↘ No (11b)

b. Is there a procedure for completing the record if not signed in (that time/a reasonable time)?

Yes – *State procedure, then 12* ↘ No (12)

12. Determine what list(s) will be used for sampling. Enter name(s) and linkage entries required both here and on the inside front cover of the hospital manual.

Sample from _____

Linkage entries _____

13. HOSPITAL PERSONNEL PRESENT

Name	Title

Notes

Notes

Appendix VI

Medical Abstract—National Hospital Discharge Survey

OMB No. 0920-0212: Approval Expires 01/31/2000

Notice – All information which would permit identification of an individual or an establishment will be held confidential, will be used only by persons engaged in and for the purposes of the survey, and will not be disclosed or released to other persons or used for any other purpose. Public reporting burden of this collection of information is estimated to average 4 minutes per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. An agency may not conduct or sponsor, and a person is not required to respond to a collection of information unless it displays a currently valid OMB control number. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to DHHS Reports Clearance Officer; Paperwork Reduction Project (0920-0212); Room 531-H, Hubert H. Humphrey Building, 200 Independence Avenue, SW; Washington, DC 20201.

FORM **HDS-1**
(5-12-99)

U.S. DEPARTMENT OF COMMERCE
BUREAU OF THE CENSUS
ACTING AS COLLECTING AGENT FOR
DEPARTMENT OF HEALTH AND HUMAN SERVICES
CENTERS FOR DISEASE CONTROL AND PREVENTION
NATIONAL CENTER FOR HEALTH STATISTICS

MEDICAL ABSTRACT - NATIONAL HOSPITAL DISCHARGE SURVEY

A. PATIENT IDENTIFICATION

1. Hospital number	<input type="text"/>	4. Date of admission	Month <input type="text"/> <input type="text"/> - Day <input type="text"/> <input type="text"/> - Year <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>
2. HDS number	<input type="text"/>	5. Date of discharge	Month <input type="text"/> <input type="text"/> - Day <input type="text"/> <input type="text"/> - Year <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>
3. Medical record number	<input type="text"/>	6. Residence ZIP Code	<input type="text"/>

B. PATIENT CHARACTERISTICS

7. Date of birth	Month <input type="text"/> <input type="text"/> - Day <input type="text"/> <input type="text"/> - Year <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>	8. Age (Complete only if date of birth not given)	Units <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/> <ul style="list-style-type: none"> 1 <input type="checkbox"/> Years 2 <input type="checkbox"/> Months 3 <input type="checkbox"/> Days
------------------	--	---	--

9. Sex (Mark (X) one)

1 <input type="checkbox"/> Male	2 <input type="checkbox"/> Female	3 <input type="checkbox"/> Not stated
---------------------------------	-----------------------------------	---------------------------------------

10. Race

1 <input type="checkbox"/> White	5 <input type="checkbox"/> Other (Specify) _____
2 <input type="checkbox"/> Black	
3 <input type="checkbox"/> American Indian/Eskimo/Aleut	
4 <input type="checkbox"/> Asian/Pacific Islander	6 <input type="checkbox"/> Not stated

11. Ethnicity (Mark (X) one)

1 <input type="checkbox"/> Hispanic origin	2 <input type="checkbox"/> Non-Hispanic	3 <input type="checkbox"/> Not stated
--	---	---------------------------------------

12. Marital status (Mark (X) one)

1 <input type="checkbox"/> Married	3 <input type="checkbox"/> Widowed	5 <input type="checkbox"/> Separated
2 <input type="checkbox"/> Single	4 <input type="checkbox"/> Divorced	6 <input type="checkbox"/> Not stated

13. Expected source(s) of payment

	Principal (Mark one only)	Other additional sources (Mark all that apply)
1. Worker's compensation	<input type="checkbox"/>	<input type="checkbox"/>
2. Medicare	<input type="checkbox"/>	<input type="checkbox"/>
3. Medicaid	<input type="checkbox"/>	<input type="checkbox"/>
4. Other government payments	<input type="checkbox"/>	<input type="checkbox"/>
5. Blue Cross/Blue Shield	<input type="checkbox"/>	<input type="checkbox"/>
6. HMO/PPO	<input type="checkbox"/>	<input type="checkbox"/>
7. Other private or commercial insurance	<input type="checkbox"/>	<input type="checkbox"/>
8. Self pay	<input type="checkbox"/>	<input type="checkbox"/>
9. No charge	<input type="checkbox"/>	<input type="checkbox"/>
10. Other (Specify) _____	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/> No source of payment indicated	<input type="checkbox"/>	<input type="checkbox"/>

14. Status/Disposition of patient

(Mark (X) appropriate box(es))	Status	Disposition
	1 <input type="checkbox"/> Alive	a. <input type="checkbox"/> Routine discharge/discharged home b. <input type="checkbox"/> Left against medical advice c. <input type="checkbox"/> Discharged, transferred to another short-term hospital d. <input type="checkbox"/> Discharged, transferred to long-term care institution e. <input type="checkbox"/> Other disposition/not stated
	2 <input type="checkbox"/> Died	
	3 <input type="checkbox"/> Status not stated	

(Over)

Appendix VIII

Statement of Agreement

A. Release

Hospital *Hospital Name*

Authorizes (Automated Source)

to release to the National Center for Health Statistics, CDC selected demographic and medical data on all patients discharged during each calendar year as specified below. The attached National Hospital Discharge Survey medical abstract form contains the specific data items that will be requested and used for the survey.

B. Options

Please specify by checking the appropriate box(s).

1. This institution agrees to participate for the duration of its agreement with its present medical abstracting service (shown above).

2. This institution agrees to participate for the data year(s).

1995 Approve Release

1997 Approve Release

1996 Approve Release

1998 Approve Release

C. Use of Data and Assurances of Confidentiality

The National Center for Health Statistics will use the patient data solely for statistical purposes in providing aggregate data on inpatient utilization. All information obtained which would permit identification of any individual or facility is collected with a guarantee that it will be held in strict confidence, will be used only by NCHS staff and its agents directly engaged in and for the purpose of the survey, and will not be disclosed or released to other persons or used for any other purpose. All information collected which may identify individuals (medical record number and Zip Code) or hospitals (NHDS hospital number) are removed prior to release to the public. The National Center for Health Statistics and its agent physically safeguard the data and are bound by statutory confidentiality restrictions of 42 USC 242m. If your hospital changes its arrangement for processing inpatient data, this agreement will automatically become void.

Signature of Authorized Hospital Official

Date

Title of Authorized Hospital Official

Signature of Authorized NCHS Official

Appendix IX

Abstract Service Agreement

The National Center for Health Statistics (NCHS) agrees to purchase specific hospital discharge abstract data for selected hospitals from *Name of Abstract Service*. The following sections detail the conditions and guidelines for submission of this data which will be incorporated into the National Hospital Discharge Survey (NHDS).

A. GENERAL CONDITIONS

1. The *Name of Abstract Service* will furnish all discharges for each hospital identified by NCHS.
2. The NCHS expects to collect discharge data from the *Name of Abstract Service* semiannually unless otherwise stated. In other words, this agreement will cover a five-year period and, within this period, data will be requested for the following six-month periods.
 - January - June 1996
 - July - December 1996
 - January - June 1997
 - July - December 1997
 - January - June 1998
 - July - December 1998
 - January - June 1999
 - July - December 1999
 - January - June 2000
 - July - December 2000
3. NCHS requests only those items from each discharge record which correspond to the Uniform Hospital Discharge Data Set (UHDDS), plus marital status, if available (see Section E, Computer Tape Layout).
4. The *Name of Abstract Service* agrees to submit its discharge datasets to the NCHS within 120 days after the end of the six-month periods listed above. Any data not forwarded for the first six-month period must be submitted at year end, in the second six-month transmittal.

B. INSTRUCTIONS FOR LABELING AND DELIVERY OF DISCHARGE DATA

1. When sending computer tape(s) to the NCHS, a copy of each tape must be kept for a period of 60 days after the mailing. This will guard against the possibility of added delays in the NCHS processing due to tape loss, defective tapes, or inadvertent damage during tape processing.
2. Each tape sent to the NCHS should be accompanied by a clear description of the tape layout and contents. This description must include a count of the total number of records on each tape and the number of records per hospital.
3. A printout of the first ten (10) records should be included with each tape submitted to enable the NCHS to review recoding and reformatting specifications.
4. Physical standards for all data tapes sent should be clearly stated on the Tape Transmittal form.
 - Encoding structure: EBCDIC
 - Density: 6250 or 38,000 BPI
 - Labels: Standard IBM label/No label
5. Data-specific characteristics must be defined and stated on the Tape Transmittal form:
 - Dataset name(s) and volume serial number(s)
 - Record length(s)
 - Block size(s)
6. The name of the person responsible for preparing the tapes for this project and a phone number where he/she can be reached if questions arise concerning the contents of the tape should be included on the Tape Transmittal form.
7. An invoice should be included with each tape submission. The invoice should specify EIN number, reporting period for data, number of hospitals reporting, total number of records, shipping cost, and tape cost.

8. All data, documentation, invoices, and correspondence should be sent to:

National Center for Health Statistics
Presidential Building, Room 956
6525 Belcrest Road
Hyattsville, MD 20782

C. FORMULA FOR PAYMENT OF PURCHASED ABSTRACT RECORDS

___ per record will be paid for each discharge record.

D. ASSISTANCE TO NCHS IN MONITORING SUBMISSION OF HOSPITAL DATA AND CHANGES IN HOSPITAL

The *Name of Abstract Service* agrees to assist the NCHS in obtaining complete datasets for each hospital and to inform NCHS when any of these hospitals are no longer in business, merged, or have a submission problem.

Appendix XI

Automated File Layout

E. LAYOUT-File Structure

<u>NHDS DATA ITEM</u>	<u>TAPE POSITION</u>	<u>NHDS CODING STRUCTURE</u>
	001	Leave Blank
Hospital ID Number	002 - 005	NHDS Hospital Identifier
Medical Record Number	006 - 016	Enter last 11 digits
Admission Date	017 - 021	Month 01 - 12 MM Day 01 - 31 DD Year 0 - 9 Y (Last Digit of Year)
Discharge Date	022 - 026	Month 01 - 12 MM Day 01 - 31 DD Year 0-9 Y (Last digit of Year)
Date of Birth	027 - 034	Month 01 - 12 MM Day 01 - 31 DD Year 4 digits YYYY
Age	035 - 036	01 - 99 Years 01 - 11 Months 00 - 28 Days
(Leave age blank if date of birth is reported).		
Age Units	037	Years 1 Months 2 Days 3
(Leave age units blank if date of birth is reported).		

E. LAYOUT

<u>NHDS DATA ITEM</u>	<u>TAPE POSITION</u>	<u>NHDS CODING STRUCTURE</u>
Sex	038	Male 1
		Female 2
		Not Reported 3
<hr/>		
Race	039	White 1
		Black 2
		American Indian/ Eskimo/Aleut 3
		Asian/
		Pacific Islander 4
		Other 5
		Not Stated 6
<hr/>		
Marital Status	040	Married 1
		Single 2
		Widowed 3
		Divorced 4
		Separated 5
		Not Stated 6
<hr/>		
Discharge Status	041	Routine Discharge 1
		Left Against Medical Advice 2
		Discharged to Another Short-Term Hospital 3
		Discharged to Long- Term Care Institution 4
		Other Alive Status 5
		Dead 6
		Not Stated 7
<hr/>		
Principal Diagnosis	042 - 046	ICD-9-CM Code

E. LAYOUT

<u>NHDS DATA ITEM</u>	<u>TAPE POSITION</u>	<u>NHDS CODING STRUCTURE</u>
Other Diagnoses	(1) 047 - 051	Up to 6 other 5-digit ICD-9-CM codes
" "	(2) 052 - 056	
" "	(3) 057 - 061	
" "	(4) 062 - 066	
" "	(5) 067 - 071	
" "	(6) 072 - 076	
<hr/>		
Principal Procedure	077 - 080	ICD-9-CM Code
<hr/>		
Other Procedures	(1) 081 - 084	Up to 3 other 4-digit ICD-9-CM codes
" "	(2) 085 - 088	
" "	(3) 089 - 092	
<hr/>		
Zip Code of Patient's Residence	093 - 097	
<hr/>		
Principal Payment Source	098 - 099	Workmen's Comp. 01
<hr/>		Medicare 02
<hr/>		Medicaid 03
Other Payment Source (1)	100 - 101	Title V 04
<hr/>		Other Gov't 05
<hr/>		Blue Cross 06
Other Payment Source (2)	102 - 103	Other Priv./Comm. 07
<hr/>		Self-pay 08
<hr/>		No Charge 09
<hr/>		Other 10
<hr/>		Not stated 11
<hr/>		HMO/PPO 12
<hr/>		
Date of Principal Surgery	104 - 105	Month 01 - 12 MM
	106 - 107	Day 01 - 31 DD
	108	Year 0 - 9 Y
		Last Digit of Year)
<hr/>		

E. LAYOUT

<u>NHDS DATA ITEM</u>	<u>TAPE POSITION</u>	<u>NHDS CODING STRUCTURE</u>
Date of Secondary Surgery	109 - 110 111 - 112 113	Month 01 - 12 MM Day 01 - 31 DD Year 0 - 9 Y (Last Digit of Year)
Ethnicity	114	Hispanic 1 Non-Hispanic 2 Not Stated 3
Calculated length of Stay	115 - 117	Length of Stay
Admission and Discharge Year (4-Digits)	118 - 121 122 - 125	
Blank Fill	125 - 135	Leave Blank

Appendix XII

Information Request Letter



DEPARTMENT OF HEALTH & HUMAN SERVICES

Public Health Service
Centers for Disease Control and Prevention

National Center for Health Statistics
6525 Belcrest Road
Hyattsville, Maryland 20782

March 10, 1998

Dear Medical Records Director:

Thank you for your continued cooperation with the National Hospital Discharge Survey, the Nation's primary resource for providing utilization statistics of short-stay hospitals in the United States. In order to track changes in characteristics of the hospitals in our survey, we need your voluntary response to the information requested below.

Please fill in and return this form to us as soon as possible in the enclosed postage-paid envelope.

Thank you for your assistance.

Sincerely yours,

Robert Pokras
Chief, Hospital Care Statistics Branch

A. Is the information on the address label at the top of this page correct? <input type="checkbox"/> YES <input type="checkbox"/> NO (Please make needed corrections)							
B. Please provide the number of inpatients discharged from this hospital for the 12 months of 1997.							
Month	Disch.	Month	Disch.	Month	Disch.	Month	Disch.
Jan		Apr		Jul		Oct	
Feb		May		Aug		Nov	
Mar		Jun		Sep		Dec	
C. What was the average number of hospital beds , excluding "newborn" bassinets, staffed for inpatient use _____.							
D. Total number of newborns for 1997 _____.							

OMB No. 0920-0212: Approval Expires: 10/31/2000

Public reporting burden for this collection of information is estimated to average 2 minutes per response. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to CDC/ATSDR Reports Clearance Officer; 1600 Clifton Road, MS D-24, Atlanta, GA 30333, ATTN: PRA (0920-0212).

Appendix XIII

Definitions of Terms Related to the Survey

Hospital—Hospital with an average length of stay of less than 30 days for all patients. Hospitals whose specialty was general (medical or surgical) or children’s general were included, regardless of length of stay. Federal hospitals, hospital units of institutions, and hospitals with less than six beds staffed for patients’ use are not included.

Type of ownership of hospital—The type of organization that controls and operates the hospital. Hospitals are grouped as follows:

Not for profit—Hospital operated by a church or another not-for-profit organization.

Government—Hospital operated by State and local government.

Proprietary—Hospital operated by individuals, partnerships, or corporations for profit.

Inpatient—A person who is formally admitted to the inpatient service of a short-stay hospital for observation, care, diagnosis, or treatment. The terms “inpatient” and “patient” are used synonymously.

Newborn infant—A patient admitted by birth to a hospital.

Discharge—The formal release of a patient by a hospital; that is, the termination of a period of hospitalization by death or by disposition to place of residence, nursing home, or another hospital. The terms “discharges” and “patients discharged” are used synonymously.

Average length of stay—The number of days of care accumulated by patients discharged during the year divided by the number of these discharges.

Diagnosis—A disease or injury (or factor that influences health status and contact with health services that is not itself a current illness or injury) listed on the medical record of a patient.

Principal diagnosis—The condition established after study to be chiefly responsible for occasioning the

admission of the patient to the hospital for care.

First-listed diagnosis—The diagnosis specified as the principal diagnosis on the face sheet or discharge summary of the medical record, or if the principal diagnosis is not specified, the diagnosis listed first on the face sheet or discharge summary of the medical record. The number of first-listed diagnoses is equivalent to the number of discharges.

Procedure—A surgical or nonsurgical operation, diagnostic procedure, or special treatment reported on the medical record of a patient.

Rate of procedures—The ratio of the number of procedures during a year to the number of persons in the civilian population on July 1 of that year determines the rate of procedures.

Population—The U.S. resident population excluding members of the Armed Forces.

Age—Patient’s age at the birthday before admission to the hospital.

Geographic region—Hospitals are classified by location in one of the four geographic regions of the United States that correspond to those used by the U.S. Bureau of the Census.

<i>Region</i>	<i>States included</i>
Northeast	Maine, New Hampshire, Vermont, Massachusetts, Rhode Island, Connecticut, New York, New Jersey, and Pennsylvania
Midwest	Michigan, Ohio, Illinois, Indiana, Wisconsin, Minnesota, Iowa, Missouri, North Dakota, South Dakota, Nebraska, and Kansas
South	Delaware, Maryland, District of Columbia, Virginia, West Virginia, North Carolina, South Carolina, Georgia, Florida, Kentucky, Tennessee, Alabama, Mississippi, Arkansas, Louisiana, Oklahoma, and Texas,

West
Montana, Idaho, Wyoming, Colorado, New Mexico, Arizona, Utah, Nevada, Washington, Oregon, California, Hawaii, and Alaska

Vital and Health Statistics series descriptions

- SERIES 1. **Programs and Collection Procedures**—These reports describe the data collection programs of the National Center for Health Statistics. They include descriptions of the methods used to collect and process the data, definitions, and other material necessary for understanding the data.
- SERIES 2. **Data Evaluation and Methods Research**—These reports are studies of new statistical methods and include analytical techniques, objective evaluations of reliability of collected data, and contributions to statistical theory. These studies also include experimental tests of new survey methods and comparisons of U.S. methodology with those of other countries.
- SERIES 3. **Analytical and Epidemiological Studies**—These reports present analytical or interpretive studies based on vital and health statistics. These reports carry the analyses further than the expository types of reports in the other series.
- SERIES 4. **Documents and Committee Reports**—These are final reports of major committees concerned with vital and health statistics and documents such as recommended model vital registration laws and revised birth and death certificates.
- SERIES 5. **International Vital and Health Statistics Reports**—These reports are analytical or descriptive reports that compare U.S. vital and health statistics with those of other countries or present other international data of relevance to the health statistics system of the United States.
- SERIES 6. **Cognition and Survey Measurement**—These reports are from the National Laboratory for Collaborative Research in Cognition and Survey Measurement. They use methods of cognitive science to design, evaluate, and test survey instruments.
- SERIES 10. **Data From the National Health Interview Survey**—These reports contain statistics on illness; unintentional injuries; disability; use of hospital, medical, and other health services; and a wide range of special current health topics covering many aspects of health behaviors, health status, and health care utilization. They are based on data collected in a continuing national household interview survey.
- SERIES 11. **Data From the National Health Examination Survey, the National Health and Nutrition Examination Surveys, and the Hispanic Health and Nutrition Examination Survey**—Data from direct examination, testing, and measurement on representative samples of the civilian noninstitutionalized population provide the basis for (1) medically defined total prevalence of specific diseases or conditions in the United States and the distributions of the population with respect to physical, physiological, and psychological characteristics, and (2) analyses of trends and relationships among various measurements and between survey periods.
- SERIES 12. **Data From the Institutionalized Population Surveys**—Discontinued in 1975. Reports from these surveys are included in Series 13.
- SERIES 13. **Data From the National Health Care Survey**—These reports contain statistics on health resources and the public's use of health care resources including ambulatory, hospital, and long-term care services based on data collected directly from health care providers and provider records.
- SERIES 14. **Data on Health Resources: Manpower and Facilities**—Discontinued in 1990. Reports on the numbers, geographic distribution, and characteristics of health resources are now included in Series 13.
- SERIES 15. **Data From Special Surveys**—These reports contain statistics on health and health-related topics collected in special surveys that are not part of the continuing data systems of the National Center for Health Statistics.
- SERIES 16. **Compilations of Advance Data From Vital and Health Statistics**—Advance Data Reports provide early release of information from the National Center for Health Statistics' health and demographic surveys. They are compiled in the order in which they are published. Some of these releases may be followed by detailed reports in Series 10–13.
- SERIES 20. **Data on Mortality**—These reports contain statistics on mortality that are not included in regular, annual, or monthly reports. Special analyses by cause of death, age, other demographic variables, and geographic and trend analyses are included.
- SERIES 21. **Data on Natality, Marriage, and Divorce**—These reports contain statistics on natality, marriage, and divorce that are not included in regular, annual, or monthly reports. Special analyses by health and demographic variables and geographic and trend analyses are included.
- SERIES 22. **Data From the National Mortality and Natality Surveys**—Discontinued in 1975. Reports from these sample surveys, based on vital records, are now published in Series 20 or 21.
- SERIES 23. **Data From the National Survey of Family Growth**—These reports contain statistics on factors that affect birth rates, including contraception, infertility, cohabitation, marriage, divorce, and remarriage; adoption; use of medical care for family planning and infertility; and related maternal and infant health topics. These statistics are based on national surveys of women of childbearing age.
- SERIES 24. **Compilations of Data on Natality, Mortality, Marriage, Divorce, and Induced Terminations of Pregnancy**—These include advance reports of births, deaths, marriages, and divorces based on final data from the National Vital Statistics System that were published as supplements to the *Monthly Vital Statistics Report* (MVSR). These reports provide highlights and summaries of detailed data subsequently published in *Vital Statistics of the United States*. Other supplements to the MVSR published here provide selected findings based on final data from the National Vital Statistics System and may be followed by detailed reports in Series 20 or 21.

For answers to questions about this report or for a list of reports published in these series, contact:

Data Dissemination Branch
National Center for Health Statistics
Centers for Disease Control and Prevention
6525 Belcrest Road, Room 1064
Hyattsville, MD 20782-2003
(301) 458-4636
E-mail: nchsquery@cdc.gov
Internet: www.cdc.gov/nchs/

**DEPARTMENT OF
HEALTH & HUMAN SERVICES**

Centers for Disease Control and Prevention
National Center for Health Statistics
6525 Belcrest Road
Hyattsville, Maryland 20782-2003

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, \$300

STANDARD MAIL (A)
POSTAGE & FEES PAID
CDC/NCHS
PERMIT NO. G-284