HEALTHCARE COST AND UTILIZATION PROJECT — HCUP
A FEDERAL-STATE-INDUSTRY PARTNERSHIP IN HEALTH DATA
Sponsored by the Agency for Healthcare Research and Quality

INTRODUCTION TO

THE HCUP NATIONWIDE INPATIENT SAMPLE (NIS),

2003

These pages provide only an introduction to the NIS package.

Full documentation is provided on the NIS Documentation CD-ROM.

For documentation updates and changes, be sure to visit: http://www.hcup-us.ahrq.gov.

Table of Contents

Index of Tables

HCUP NATIONWIDE INPATIENT SAMPLE (NIS)
SUMMARY OF DATA USE LIMITATIONS

***** REMINDER *****

All users of the NIS must take the on-line HCUP Data Use Agreement (DUA) training course, and read and sign a Data Use Agreement.[†]

Authorized users of HCUP data agree to the following restrictions:[‡]

- Will not use the data for any purpose other than research or aggregate statistical reporting.

- Will not re-release any data to unauthorized users.

- Will not redistribute HCUP data by posting on any Web site or publicly-accessible online repository.

- Will not identify or attempt to identify any individual, including by the use of vulnerability analysis or penetration testing. Methods that could be used to identify individuals directly or indirectly shall not be disclosed or published.

- Will not publish information that could identify individual establishments (e.g., hospitals) and will not contact establishments.

- Will not use the data concerning individual establishments for commercial or competitive purposes involving those establishments and will not use the data to determine rights, benefits, or privileges of individual establishments.

- Will not use data elements from the proprietary severity adjustment software packages (3M APR-DRGs, HSS APS-DRGs, and Truven Health Analytics Disease Staging) for any commercial purpose or to disassemble, decompile, or otherwise reverse engineer the proprietary software.

- Will acknowledge in reports that data from the "Healthcare Cost and Utilization Project (HCUP)," were used, including names of the specific databases used for analysis.

- Will acknowledge that risk of individual identification of persons is increased when observations (i.e., individual discharge records) in any given cell of tabulated data is less than or equal to 10.

Any violation of the limitations in the Data Use Agreement is punishable under Federal law by a fine of up to $10,000 and up to 5 years in prison. Violations may also be subject to penalties under State statutes.

[†] The on-line Data Use Agreement training session and the Data Use Agreement are available on the HCUP User Support (HCUP-US) Website at http://www.hcup-us.ahrq.gov.
[‡] Specific provisions are detailed in the Data Use Agreement for Nationwide Databases.

HCUP CONTACT INFORMATION

All HCUP data users, including data purchasers and collaborators, must complete the online HCUP Data Use Agreement (DUA) Training Tool, and read and sign the HCUP Data Use Agreement. Proof of training completion and signed Data Use Agreements must be submitted to the HCUP Central Distributor as described below.

The on-line DUA training course is available at: http://www.hcup-us.ahrq.gov/tech_assist/dua.jsp

The HCUP Nationwide Data Use Agreement is available on the AHRQ-sponsored HCUP User Support (HCUP-US) Web site at: http://www.hcup-us.ahrq.gov

HCUP Central Distributor

Data purchasers will be required to provide their DUA training completion code and will execute their DUAs electronically as a part of the online ordering process. The DUAs and training certificates for collaborators and others with access to HCUP data should be submitted directly to the HCUP Central Distributor using the contact information below.

The HCUP Central Distributor can also help with questions concerning HCUP database purchases, your current order, training certificate codes, or invoices, if your questions are not covered in the Purchasing FAQs on the HCUP Central Distributor Web site.

Purchasing FAQs:
https://www.distributor.hcup-us.ahrq.gov/Purchasing-Frequently-Asked-Questions.aspx

Phone: (866) 290-HCUP (4287)
Email: HCUPDistributor@AHRQ.gov
Fax: 866-792-5313 (toll free in the United States)

Mailing address:
HCUP Central Distributor
Social & Scientific Systems, Inc.
8757 Georgia Ave, 12th Floor
Silver Spring, MD 20910

HCUP User Support:

Information about the content of the HCUP databases is available on the HCUP User Support (HCUP-US) Web site (http://www.hcup-us.ahrq.gov). If you have questions about using the HCUP databases, software tools, supplemental files, and other HCUP products, please review the HCUP Frequently Asked Questions or contact HCUP User Support:

HCUP FAQs: http://www.hcup-us.ahrq.gov/tech_assist/faq.jsp

Phone: 866-290-HCUP (4287) (toll free)
Email: hcup@ahrq.gov

WHAT'S NEW IN THE 2003
NATIONWIDE INPATIENT SAMPLE (NIS)?

- Two States have joined the NIS in 2003: Indiana and New Hampshire. Data from Maine were not available for inclusion in the 2003 NIS.

- The data element HOSPSTCO (the hospital's county) is no longer present on the Core file but remains on the hospital file.

- The data elements MDNUM1_R and MDNUM2_R (reidentified physician numbers) replace previous physician number data elements.

- New data elements include the following:

  ‣ ASOURCEUB92 indicates the source of admission and uses the same coding as the source of admission data element on the UB-92 claim form. ASOURCEUB92 has more detailed categories for routine admissions and transfers from other health facilities than the HCUP data element ASOURCE.

  ‣ ECODEn is an array of the external cause of injury codes (i.e., "E codes"). Prior to 2003, E codes are included in the diagnosis array (DXn). Beginning in 2003, any separately reported E codes and any E codes encountered in the diagnosis array are placed in a separate array specific to E codes (ECODEn). NECODES indicates the total number of E codes (valid and invalid) that are coded on the original discharge record. The array E_CCSn contains the Clinical Classifications Software (CCS) category for the corresponding E code in the ECODEn array. The CCS consists of 21 categories of E codes based on a classification system developed by the Centers for Disease Control.

  ‣ PT_UR_CAT4 is an urban-rural measure that characterizes the patient's county of residence. The measure is a consolidated version of the Urban Influence Codes (UIC) developed by the U.S. Department of Agriculture, which expands upon the Metropolitan Statistical Area (MSA) classification as applied to the 2000 Census. In this classification, urban counties are distinguished by the population of their largest metropolitan area. Rural counties are differentiated by the size of the largest metropolitan area they are adjacent to and the size of their largest community.

  ‣ ZIPINC_QRTL is a categorical variable indicating the quartile that contains the median household income of the patient's ZIP Code of residence. For 2003, the median income quartiles are defined as: $1 - $35,999; $36,000 - $44,999; $45,000 - $59,999; and $60,000 or more.

HEALTHCARE COST AND UTILIZATION PROJECT — HCUP
A FEDERAL-STATE-INDUSTRY PARTNERSHIP IN HEALTH DATA
Sponsored by the Agency for Healthcare Research and Quality

**The Agency for Healthcare Research and Quality and
the staff of the Healthcare Cost and Utilization Project (HCUP) thank you for
purchasing the HCUP Nationwide Inpatient Sample (NIS).**

HCUP Nationwide Inpatient Sample (NIS)

ABSTRACT

The Nationwide Inpatient Sample (NIS) is part of the Healthcare Cost and Utilization Project (HCUP), sponsored by the Agency for Healthcare Research and Quality (AHRQ), formerly the Agency for Health Care Policy and Research.

The NIS is a database of hospital inpatient stays. Researchers and policymakers use the NIS to identify, track, and analyze national trends in health care utilization, access, charges, quality, and outcomes.

The NIS is the largest all-payer inpatient care database that is publicly available in the United States, containing data from 5 to 8 million hospital stays from about 1,000 hospitals sampled to approximate a 20-percent stratified sample of U.S. community hospitals. The NIS is available for a 16-year time period, from 1988 to 2003, allowing analysis of trends over time. (Analyses of time trends are recommended from 1993 forward; see *Using the HCUP Nationwide Inpatient Sample to Estimate Trends*, a report available on the HCUP User Support Website and on the NIS Documentation CD-ROM.)

The NIS is the only national hospital database with charge information on all patients, regardless of payer, including persons covered by Medicare, Medicaid, private insurance, and the uninsured. The NIS's large sample size enables analyses of rare conditions, such as congenital anomalies; uncommon treatments, such as organ transplantation; and special patient populations, such as the uninsured.

Inpatient stay records in the NIS include clinical and resource use information typically available from discharge abstracts. Hospital and discharge weights are provided for producing national estimates. The NIS can be linked to hospital-level data from the American Hospital Association's Annual Survey of Hospitals and county-level data from the Bureau of Health Professions' Area Resource File, except in those states that do not allow the release of hospital identifiers.

Beginning in 1998, the NIS differs from previous NIS releases: some data elements were dropped, some were added, for some data elements, the coding was changed, and the sampling and weighting strategy was revised to improve the representativeness of the data. (See the report, *Changes in the NIS Sampling and Weighting Strategy for 1998*, which describes these changes, available on the HCUP User Support Website and on the NIS Documentation CD-ROM.)

Access to the NIS is open to users who sign data use agreements. Uses are limited to research and aggregate statistical reporting.

For more information on the NIS, visit the AHRQ-sponsored HCUP User Support (HCUP-US) Website at http://www.hcup-us.ahrq.gov or see the detailed documentation on the documentation CD that accompanies the data.

INTRODUCTION TO THE HCUP NATIONWIDE INPATIENT SAMPLE (NIS)

OVERVIEW OF NIS DATA

The Nationwide Inpatient Sample (NIS) contains all-payer data on hospital inpatient stays from States participating in the Healthcare Cost and Utilization Project (HCUP). Each year of the NIS provides information on approximately 5 million to 8 million inpatient stays from about 1,000 hospitals. All discharges from sampled hospitals are included in the NIS database.

The NIS contains patient-level clinical and resource use information included in a typical discharge abstract. The NIS can be linked directly to hospital-level data from the American Hospital Association (AHA) Annual Survey of Hospitals and to county-level data from the Health Resources and Services Administration Bureau of Health Professions' Area Resource File (ARF), except in those states that do not allow the release of hospital identifiers.

The NIS is designed to approximate a 20-percent sample of U.S. community hospitals, defined by the AHA to be "all nonfederal, short-term, general, and other specialty hospitals, excluding hospital units of institutions." Included among community hospitals are specialty hospitals such as obstetrics-gynecology, ear-nose-throat, short-term rehabilitation, orthopedic, and pediatric institutions. Also included are public hospitals and academic medical centers. Excluded are short-term rehabilitation hospitals (beginning with 1998 data), long-term hospitals, psychiatric hospitals, and alcoholism/chemical dependency treatment facilities.

This universe of U.S. community hospitals is divided into strata using five hospital characteristics: ownership/control, bed size, teaching status, urban/rural location, and U.S. region.

The NIS is a stratified probability sample of hospitals in the frame, with sampling probabilities proportional to the number of U.S. community hospitals in each stratum. The frame is limited by the availability of inpatient data from the data sources.

In order to improve the representativeness of the NIS, the sampling and weighting strategy was modified beginning with the 1998 data. The full description of this process can be found in the special report on *Changes in NIS Sampling and Weighting Strategy for 1998.* This report is available on the 2003 NIS Documentation CD-ROM and on the AHRQ-sponsored HCUP User Support (HCUP-US) Website at http://www.hcup-us.ahrq.gov. To facilitate the production of national estimates, both hospital and discharge weights are provided, along with information necessary to calculate the variance of estimates. Detailed information on the design of the NIS is available in the year-specific special reports on *Design of the Nationwide Inpatient Sample* found on the HCUP-US website and the NIS Documentation CD-ROM.

NIS data sets are currently available for multiple years, as shown in Table 1. Each release of the NIS includes:

- X   Data in fixed-width ASCII format on CD-ROM.
- X   Patient-level hospital discharge abstract data for 100 percent of discharges from a sample of hospitals in participating States.
- X   5 million to 8 million inpatient records per year.
- X   800-1,000 hospitals per year.
- X   Two 10% subsamples of discharges from all NIS hospitals.
- X   Discharge-level weights to calculate national estimates for discharges.
- X   Hospital Weights File to produce national estimates for hospitals and to link the NIS to data from the American Hospital Association Annual Survey of Hospitals.
- X   NIS Documentation and tools, also on CD-ROM – including file specifications, programming source code for loading ASCII data into SAS and SPSS, and value labels.

Table 1. Summary of NIS Releases

Data from	Media/format options	Structure of Releases
▪ 1988-1992 ▪ 8 States in 1988 ▪ 11 States in 1989-1992	On CD-ROM, in ASCII format	5 years of data in a 6-CD set, compressed files Two 10% subsamples of discharges for each year
▪ 1993 ▪ 17 states ▪ 1994 ▪ 17 states ▪ 1995 ▪ 19 states ▪ 1996 ▪ 19 states ▪ 1997 ▪ 22 states ▪ 1998 ▪ 22 states ▪ 1999 ▪ 24 states ▪ 2000 ▪ 28 states ▪ 2001 ▪ 33 states ▪ 2002 ▪ 35 states ▪ 2003 ▪ 37 states	On CD-ROM, in ASCII format	1 year of data in a 2-CD set, compressed files Two 10% subsamples of discharges for each year Beginning in 2002, a companion file with four different sets of severity measures

NIS Data Sources, Hospitals, and Inpatient Stays

Table 2 summarizes the data sources, number of hospitals, and number of unweighted and weighted inpatient stays in NIS data.

Table 2. Summary of NIS Data Sources, Hospitals and Inpatient Stays, 1988-2003

Year	Data sources	Number of hospitals	Number of discharges in the NIS, unweighted	Number of discharges in the NIS, weighted for national estimates
1988	CA CO FL IL IA MA NJ WA	759	5,265,756	35,171,448
1989	AZ CA CO FL IL IA MA NJ PA WA WI *(Added AZ, PA, WI)*	882	6,110,064	35,104,645
1990	AZ CA CO FL IL IA MA NJ PA WA WI *(No change)*	871	6,268,515	35,215,397
1991	AZ CA CO FL IL IA MA NJ PA WA WI *(No change)*	859	6,156,188	35,036,492
1992	AZ CA CO FL IL IA MA NJ PA WA WI *(No change)*	856	6,195,744	35,011,385
1993	AZ CA CO CT FL IL IA KS MD MA NJ NY OR PA SC WA WI *(Added CT, KS, MD, NY, OR, SC)*	913	6,538,976	34,714,530
1994	AZ CA CO CT FL IL IA KS MD MA NJ NY OR PA SC WA WI *(No change)*	904	6,385,011	34,622,203
1995	AZ CA CO CT FL IL IA KS MD MA MO NJ NY OR PA SC TN WA WI *(Added MO, TN)*	938	6,714,935	34,791,998
1996	AZ CA CO CT FL IL IA KS MD MA MO NJ NY OR PA SC TN WA WI *(No change)*	906	6,542,069	34,874,386
1997	AZ CA CO CT FL GA HI IL IA KS MD MA MO NJ NY OR PA SC TN UT WA WI *(Added GA, HI, UT)*	1,012	7,148,420	35,408,207
1998	AZ CA CO CT FL GA HI IL IA KS MD MA MO NJ NY OR PA SC TN UT WA WI *(No change)*	984	6,827,350	34,874,001
1999	AZ CA CO CT FL GA HI IL IA KS MD MA ME MO NJ NY OR PA SC TN UT VA WA WI *(Added ME, VA)*	984	7,198,929	35,467,673

Year	Data sources	Number of hospitals	Number of discharges in the NIS, unweighted	Number of discharges in the NIS, weighted for national estimates
2000	AZ CA CO CT FL GA HI IL IA KS KY MD MA ME MO NC NJ NY OR PA SC TN TX UT VA WA WI WV *(Added KY, NC, TX, WV)*	994	7,450,992	36,417,565
2001	AZ CA CO CT FL GA HI IL IA KS KY MD MA ME MI MN MO NC NE NJ NY OR PA RI SC TN TX UT VA VT WA WI WV *(Added MI, MN, NE, RI, VT)*	986	7,452,727	37,187,641
2002	CA CO CT FL GA HI IL IA KS KY MD MA ME MI MN MO NC NE NJ NY NV OH OR PA RI SC SD TN TX UT VA VT WA WI WV *(Added NV, OH, SD; AZ data were not available)*	995	7,853,982	37,804,021
2003	AZ CA CO CT FL GA HI IL IN IA KS KY MD MA MI MN MO NC NE NH NJ NY NV OH OR PA RI SC SD TN TX UT VA VT WA WI WV *(Added AZ, IN, NH; ME data were not available)*	994	7,977,728	38,220,659

State-Specific Restrictions

Some data sources that contributed data to the NIS imposed restrictions on the release of certain data elements or on the number and types of hospitals that could be included in the database. Because of confidentiality laws, some data sources were prohibited from providing HCUP with discharge records that indicated specific medical conditions, such as HIV/AIDS or behavioral health. Detailed information on these state-specific restrictions is available in the report on *Sources of NIS Data and State-specific Restrictions* on the NIS Documentation CD-ROM.

Contents of CD-ROM Set

There are two types of files included in the NIS: 1) data files and 2) documentation and tools files.

1)   Data Files - four types of fixed-width ASCII formatted data files are included in the NIS, beginning with the 2002 data:

Inpatient Core File: This inpatient discharge-level file contains data for 100% of the discharges from a sample of hospitals in participating states. The unit of observation is an *inpatient stay record*. See Table 3 for a list of data elements in the Inpatient Core File. This file is available in all years.

Subsample Inpatient Core Files: Each of these discharge-level files contains all data elements from the Core File for a 10% subsample of discharges from the NIS; these can be

combined to create a 20% NIS subsample. These files can be useful for testing programs or validating models. The unit of observation is an *inpatient stay record*. See Table 3 for a list of data elements in the Inpatient Core Files. These files are available in all years.

Hospital Weights File: This hospital-level file contains one observation for each hospital included in the NIS and contains weights and variance estimation data elements, as well as linkage data elements. The unit of observation is the *hospital*. The HCUP hospital identifier (HOSPID) provides the linkage between the NIS Inpatient Core files and the Hospital Weights file.   See Table 4 for a list of data elements in the Hospital Weights File. This file is available in all years.

Disease Severity Measures Files: These discharge-level files contain information from four different sets of disease severity measures. Information from these severity files is to be used in conjunction with the inpatient Core files. There is one severity measure file for each of the Core and subsample Core files. The unit of observation is an *inpatient stay record*. The HCUP unique record identifier (KEY) provides the linkage between the Core files and the Disease Severity Measures files.   See Table 5 for a list of data elements in the Severity Measures Files. These files are available beginning with the 2002 NIS.

2) Documentation and Tools Files

Documentation: Complete file documentation, variable notes, and summary statistics are provided in a series of Portable Document Format (*.pdf) files. These files are detailed in Table 7.

SAS source code: Code is included for the format library for the variables and for loading ASCII data into SAS format.

SPSS source code: Code is included for the variable library and for loading ASCII data into SPSS format.

Labels: Labels for the Clinical Classifications Software (CCS), formerly called the Clinical Classifications for Health Policy Research (CCHPR), and for the Diagnosis-Related Groups (multiple versions).

File Specifications: Record layouts for all data files.

NIS Data Elements

All releases of the NIS contain two types of data: inpatient stay records and hospital information with weights. Table 3 and Table 4 (on the following pages) identify the data elements that can be found in the inpatient stay and hospital weights files, respectively. Beginning with the 2002 NIS, an additional file including four different sets of disease severity measures is available. Table 5 identifies the data elements in the disease severity measures files.

Not all data elements in the NIS are uniformly coded or available across all States. This is not complete documentation for the data; please refer to the NIS Documentation CD-ROM for full documentation on all data elements, for summary statistics, and for the record layout.

Table 3. Data Elements in the NIS Inpatient Core Files

Data elements that are *italicized* are not included in the 2003 NIS Inpatient Core Files, but are only available in previous years' files.

Type of Data Element	HCUP Variable Name	Years Available	Coding Notes	Unavailable in 2003 for:
Admission day of week or weekend	AWEEKEND	1998-2003	Admission on weekend: (0) admission on Monday-Friday, (1) admission on Saturday-Sunday	
	ADAYWK	*1988-1997*	*Admission day of week: (1) Sunday, (2) Monday, (3) Tuesday, (4) Wednesday, etc.*	
Admission month	AMONTH	1988-2003	Admission month coded from (1) January to (12) December	FL
Admission source	ASOURCE	1988-2003	Admission source, uniform coding: (1) ER, (2) another hospital, (3) another facility including long-term care, (4) court/law enforcement, (5) routine/birth/other	KS
	ASOURCE_X	1998-2003	Admission source, as received from data source using State-specific coding	KS
	ASOURCEUB92	2003	Admission source (UB-92 standard coding). For newborn admissions (ATYPE = 4): (1) normal delivery, (2) premature delivery, (3) sick baby, (4) extramural birth; For non-newborn admissions (ATYPE NE 4): (1) physician referral, (2) clinic referral, (3) HMO referral, (4) transfer from a hospital, (5) transfer from a skilled nursing facility, (6) transfer from a another health care facility, (7) emergency room, (8) court/law enforcement, (A) transfer from a critical access hospital	
Admission type	ATYPE	1988-2003	Admission type, uniform coding: (1) emergency, (2) urgent, (3) elective, (4) newborn, (5) trauma center beginning in 2003 data, (6) other	CA
	ELECTIVE	2002-2003	Indicates elective admission: (1) elective, (0) non-elective admission	
Age at admission	AGE	1988-2003	Age in years coded 0-124 years	
	AGEDAY	1988-2003	Age in days coded 0-365 only when the age in years is less than 1	FL, MA, NH, SC, TX
Clinical Classifications Software (CCS) category	DXCCS1 - DXCCS15	1998-2003	CCS category for all diagnoses for NIS beginning in 1998	
	DCCHPR1	*1988-1997*	*CCS category for principal diagnosis for NIS prior to 1998. CCS was formerly called the Clinical Classifications for Health Policy Research (CCHPR)*	
	PRCCS1 - PRCCS15	1998-2003	CCS category for all procedures for NIS beginning in 1998	

Type of Data Element	HCUP Variable Name	Years Available	Coding Notes	Unavailable in 2003 for:
Clinical Classifications Software (CCS) category (continued)	*PCCHPR1*	*1988-1997*	*CCS category for principal procedure for NIS prior to 1998. CCS was formerly called the Clinical Classifications for Health Policy Research (CCHPR)*	
Data source information	*DSNUM*	*1988-1997*	*Data source number*	
	DSTYPE	*1988-1997*	*Data source type: (1) State data organization, (2) Hospital association, (3) Consortia*	
Diagnosis information	DX1 - DX15	1988-2003	Diagnoses, principal and secondary (ICD-9-CM). Beginning in 2003, the diagnosis array does not include any of external cause of injury codes. These codes have been stored in a separate array ECODEn.	
	NDX	1988-2003	Number of diagnoses coded on the original record	
	DSNDX	*1988-1997*	*Number of diagnosis fields provided by the data source*	
	DXSYS	*1988-1997*	*Diagnosis system (ICD-9-CM)*	
	DXV1 - DXV15	*1988-1997*	*Diagnosis validity flags*	
Diagnosis Related Group (DRG)	DRG	1988-2003	DRG in use on discharge date	
	DRGVER	1988-2003	Grouper version in use on discharge date	
	DRG10	*1988-1999*	*DRG Version 10 (effective October 1992 - September 1993)*	
	DRG18	1998-2003	DRG Version 18 (effective October 2000 - September 2001)	
Discharge quarter	DQTR	1988-2003	Coded: (1) Jan - Mar, (2) Apr - Jun, (3) Jul - Sep, (4) Oct - Dec	
Discharge weights (Weights for 1988-1993 are on Hospital Weights file)	DISCWT	1998-2003	Discharge weight on Core file and Hospital Weights file for NIS beginning in 1998. In all data years except 2000, this weight is used to create national estimates for all analyses. In 2000 only, this weight is used to create national estimates for all analyses excluding those that involve total charges.	
	DISCWT_U	*1993-1997*	*Discharge weight on Core file and Hospital Weights file for NIS prior to 1998*	
	DISCWTcharge	*2000*	*Discharge weight for national estimates of total charges. In 2000 only, this weight is used to create national estimates for analyses that involve total charges*	

Type of Data Element	HCUP Variable Name	Years Available	Coding Notes	Unavailable in 2003 for:
Discharge weights (continued)	DISCWT10	1998-2003	Discharge weight on 10% subsample Core file for NIS beginning in 1998. In all data years except 2000, this weight is used to create national estimates for all analyses. In 2000 only, this weight is used to create national estimates for all analyses excluding those that involve total charges	
	D10CWT_U	*1993-1997*	*Discharge weight on 10% subsample Core file for NIS prior to 1998*	
	DISCWTcharge10	*2000*	*Discharge weight for national estimates of total charges on 10% subsample file. In 2000 only, this weight is used to create national estimates for analyses that involve total charges*	
Discharge year	YEAR	1988-2003		
Disposition of patient (discharge status)	*DISP*	*1988-1997*	*Disposition of patient, uniform coding used prior to 1998: (1) routine, (2) short-term hospital, (3) skilled nursing facility, (4) intermediate care facility, (5) another type of facility, (6) home health care, (7) against medical advice, (20) died*	
	DIED	1988-2003	Indicates in-hospital death: (0) did not die during hospitalization, (1) died during hospitalization	
	DISPUB92	1998-2003	Disposition of patient, UB-92 coding: (1) routine, (2) short term hospital, (3) skilled nursing facility, (4) intermediate care, (5) another type of facility, (6) home health care, (7) against medical advice, (8) home IV provider, (20) died in hospital, (40) died at home, (41) died in a medical facility, (42) died, place unknown, (43) Federal health facility beginning in 2003 data, (50) Hospice, home, (51) Hospice, medical facility, (61) hospital-based Medicare approved swing bed beginning in 2000 data, (62) another rehabilitation facility beginning in 2001 data, (63) long term care hospital beginning in 2001 data, (64) certified nursing facility beginning in 2002 data, (71) another institution for outpatient services beginning in 2000 data, (72) this institution for outpatient services beginning in 2000 data, (99) discharged alive, destination unknown beginning in 2001 data	CA, MD

Type of Data Element	HCUP Variable Name	Years Available	Coding Notes	Unavailable in 2003 for:
Disposition of patient (discharge status) (continued)	DISPUNIFORM	1998-2003	Disposition of patient, uniform coding used beginning in 1998: (1) routine, (2) transfer to short term hospital, (5) other transfers, including skilled nursing facility, intermediate care, and another type of facility, (6) home health care, (7) against medical advice, (20) died in hospital, (99) discharged alive, destination unknown	
External causes of injury and poisoning	ECODE1 - ECODE4	2003	External cause of injury and poisoning code, primary and secondary (ICD-9-CM). Beginning in 2003, external cause of injury codes are stored in a separate array ECODEn from the diagnosis codes in the array DXn. Prior to 2003, these codes are contained in the diagnosis array (DXn).	
	E_CCS1 - E_CCS4	2003	CCS category for the external cause of injury and poisoning codes	
	NECODE	2003	Number of external cause of injury codes on the original record. A maximum of 4 codes are retained on the NIS.	
Gender of patient	FEMALE	1998-2003	Indicates gender for NIS beginning in 1998: (0) male, (1) female	
	SEX	*1988-1997*	*Indicates gender for NIS prior to 1998: (1) male, (2) female*	
Hospital information	DSHOSPID	1988-2003	Hospital number as received from the data source	GA, HI, IN, KS, MI, NE, OH, SC, SD, TN, TX
	HOSPID	1988-2003	HCUP hospital number (links to Hospital Weights file)	
	HOSPST	1988-2003	State postal code for the hospital (e.g., AZ for Arizona)	
	HOSPSTCO	*1988-2002*	*Modified Federal Information Processing Standards (FIPS) State/county code for the hospital links to Area Resource File (available from the Bureau of Health Professions, Health Resources and Services Administration) Beginning in 2003, this data element is available only on the hospital file.*	
	NIS_STRATUM	1998-2003	Stratum used to sample hospitals, based on geographic region, control, location/teaching status, and bed size. Stratum information is also in the Hospital Weights file.	

Type of Data Element	HCUP Variable Name	Years Available	Coding Notes	Unavailable in 2003 for:
Length of Stay	LOS	1988-2003	Length of stay, edited	
	LOS_X	1988-2003	Length of stay, as received from data source	
Location of the patient	PL_UR_CAT4	2003	Urban–rural designation for patient's county of residence: (1) large metropolitan, (2) small metropolitan, (3) micropolitan, (4) non-core	
Major Diagnosis Category (MDC)	MDC	1988-2003	MDC in use on discharge date	
	MDC10	*1988-1999*	*MDC Version 10 (effective October 1992 - September 1993)*	
	MDC18	1998-2003	MDC Version 18 (effective October 2000 - September 2001)	
Median household income for patient's ZIP Code	ZIPINC_QRTL	2003	Median household income quartiles for patient's ZIP Code. For 2003, the median income quartiles are defined as: $1 - $35,999; $36,000 - $44,999; $45,000 - $59,999; and $60,000 or more.	
	ZIPINC	*1998-2002*	*Median household income category in files beginning in 1998: (1) $1-$24,999, (2) $25,000-$34,999, (3) $35,000-$44,999, (4) $45,000 and above*	
	ZIPINC4	*1988-1997*	*Median household income category in files prior to 1998: (1) $1-$25,000, (2) $25,001-$30,000, (3) $30,001-$35,000, (4) $35,001 and above*	
	ZIPINC8	*1988-1997*	*Median household income category in files prior to 1998: (1) $1-$15,000, (2) $15,001-$20,000, (3) $20,001-$25,000, (4) $25,001-$30,000, (5) $30,001-$35,000, (6) $35,001-$40,000, (7) $40,001-$45,000, (8) $45,001 or more*	
Neonatal/ maternal flag	NEOMAT	1988-2003	Assigned from diagnoses and procedure codes: (0) not maternal or neonatal, (1) maternal diagnosis or procedure, (2) neonatal diagnosis, (3) maternal and neonatal on same record	
Payer information	PAY1	1988-2003	Expected primary payer, uniform: (1) Medicare, (2) Medicaid, (3) private including HMO, (4) self-pay, (5) no charge, (6) other	
	PAY1_N	*1988-1997*	*Expected primary payer, nonuniform: (1) Medicare, (2) Medicaid, (3) Blue Cross, Blue Cross PPO, (4) commercial, PPO, (5) HMO, PHP, etc., (6) self-pay, (7) no charge, (8) Title V, (9) Worker's Compensation, (10) CHAMPUS, CHAMPVA, (11) other government, (12) other*	

Type of Data Element	HCUP Variable Name	Years Available	Coding Notes	Unavailable in 2003 for:
Payer information (continued)	PAY1_X	1998-2003	Expected primary payer, as received from the data source	
	PAY2	1988-2003	Expected secondary payer, uniform: (1) Medicare, (2) Medicaid, (3) private including HMO, (4) self-pay, (5) no charge, (6) other	CA, CO, FL, GA, HI, IA, OH, RI, SD, VA
	PAY2_N	*1988-1997*	*Expected secondary payer, nonuniform: (1) Medicare, (2) Medicaid, (3) Blue Cross, Blue Cross PPO, (4) commercial, PPO, (5) HMO, PHP, etc., (6) self-pay, (7) no charge, (8) Title V, (9) Worker's Compensation, (10) CHAMPUS, CHAMPVA, (11) other government, (12) other*	
	PAY2_X	1998-2003	Expected secondary payer, as received from the data source	
Physician identifiers, synthetic	*MDID_S*	*1988-2000*	*Synthetic attending physician number in files prior to 2001*	
	MDNUM1_R	2003	Re-identified attending physician number in files starting in 2003	CA, CT, GA, HI, IN, IL, KS, MA, NC, OH, UT, VT, WI
	MDNUM1_S	*2001-2002*	*Synthetic attending physician number in files beginning in 2001 and discontinued in 2003*	
	SURGID_S	*1988-2000*	*Synthetic secondary physician number in files prior to 2001*	
	MDNUM2_R	2003	Re-identified secondary physician number in files starting in 2003	CA, CT, HI, IL, IN, KS, OH, WI
	MDNUM2_S	*2001-2002*	*Synthetic secondary physician number in files beginning in 2001 and discontinued in 2003*	
Procedure information	PR1 - PR15	1988-2003	Procedures, principal and secondary (ICD-9-CM)	
	NPR	1988-2003	Number of procedures coded on the original record	
	DSNPR	*1988-1997*	*Number of procedure fields in this data source*	
	PRSYS	*1988-1997*	*Procedure system (ICD-9-CM)*	
	PRV1 -PRV15	*1988-1997*	*Procedure validity flag*	

Type of Data Element	HCUP Variable Name	Years Available	Coding Notes	Unavailable in 2003 for:
Procedure information (continued)	PRDAY1	1988-2003	Number of days from admission to principal procedure.	IL, KS, OH, UT, WA, WV
	PRDAY2 - PRDAY15	1998-2003	Number of days from admission to secondary procedures.	IL, KS, OH, UT, WA, WV
Race of Patient	RACE	1988-2003	Race, uniform coding: (1) white, (2) black, (3) Hispanic, (4) Asian or Pacific Islander, (5) Native American, (6) other	GA, IL, KY, ME, MN, NE, NV, OH, OR, WA, WV
Record identifier, synthetic	KEY	1998-2003	Unique record number for file beginning in 1998	
	SEQ	*1988-1997*	*Unique record number for NIS prior to 1998*	
	SEQ_SID	*1988-1997*	*Unique record number for NIS prior to 1998*	
	PROCESS	*1988-1997*	*Processing number for NIS prior to 1998*	
Total Charges	TOTCHG	1988-2003	Total charges, edited	
	TOTCHG_X	1988-2003	Total charges, as received from data source	

Table 4. Data Elements in the NIS Hospital Weights File

Data elements that are *italicized* are not included in the 2003 NIS Hospital Weights File, but are only available in previous years' files.

Type of Data Element	HCUP Variable Name	Years Available	Coding Notes	Unavailable in 2003 for:
Discharge counts	*N_DISC_U*	*1988-1997*	*Number of AHA universe discharges in the stratum*	
	S_DISC_U	1988-2003	Number of sampled discharges in the sampling stratum (NIS_STRATUM or STRATUM)	
	S_DISC_S	*1988-1997*	*Number of sampled discharges in the stratum STRAT_ST*	
	N_DISC_F	*1988-1997*	*Number of frame discharges in the stratum*	
	N_DISC_S	*1988-1997*	*Number of State's discharges in the stratum*	
	TOTAL_DISC	1998-2003	Total number of discharges from this hospital in the NIS	
	TOTDSCHG	*1988-1997*	*Total number of discharges from this hospital in the NIS*	
Discharge weights	DISCWT	1998-2003	Discharge weight used in the NIS beginning in 1998. In all data years except 2000, this weight is used to create national estimates for all analyses. In 2000 only, this weight is used to create national estimates for all analyses excluding those that involve total charges.	
	DISCWT_U	*1988-1997*	*Discharge weights used in the NIS prior to 1998.*	
	DISCWT_F	*1988-1997*	*Discharge weights to the sample frame are available only in 1988-1997*	
	DISCWT_S	*1988-1997*	*Discharge weights to the State are available only in 1988-1997*	
	DISCWTcharge	*2000*	*Discharge weight for national estimates of total charges for 2000 only.*	
Discharge Year	YEAR	1988-2003	Discharge year	
Hospital counts	*N_HOSP_F*	*1988-1997*	*Number of frame hospitals in the stratum*	
	N_HOSP_S	*1988-1997*	*Number of State's hospitals in the stratum*	
	N_HOSP_U	1988-2003	Number of AHA universe hospitals in the stratum	
	S_HOSP_S	*1988-1997*	*Number of sampled hospitals in STRAT_ST*	
	S_HOSP_U	1988-2003	Number of sampled hospitals in the stratum (NIS_STRATUM or STRATUM)	
Hospital identifiers	HOSPID	1988-2003	HCUP hospital number (links to inpatient Core files)	
	AHAID	1988-2003	AHA hospital identifier that matches AHA Annual Survey of Hospitals (not available for all states)	GA, HI, IN, KS, MI, NE, OH, SC, SD, TN, TX

Type of Data Element	HCUP Variable Name	Years Available	Coding Notes	Unavailable in 2003 for:
Hospital identifiers (continued)	IDNUMBER	1988-2003	AHA hospital identifier without the leading 6 (not available for all states)	GA, HI, IN, KS, MI, NE, OH, SC, SD, TN, TX
	HOSPNAME	1993-2003	Hospital name from AHA Annual Survey of Hospitals (not available for all states)	GA, HI, IN, KS, MI, NE, OH, SC, SD, TN, TX
Hospital location	HOSPADDR	1993-2003	Hospital address from AHA Annual Survey of Hospitals (not available for all states)	GA, HI, IN, KS, MI, NE, OH, SC, SD, TN, TX
	HOSPCITY	1993-2003	Hospital city from AHA Annual Survey of Hospitals (not available for all states)	GA, HI, IN, KS, MI, NE, OH, SC, SD, TN, TX
	HOSPST	1988-2003	Hospital state postal code for hospital (e.g., AZ for Arizona)	
	HOSPSTCO	2002-2003	Modified Federal Information Processing Standards (FIPS) State/county code for the hospital links to Area Resource File (available from the Bureau of Health Professions, Health Resources and Services Administration)	GA, HI, IN, KS, MI, NE, OH, SC, SD, TN, TX
	HOSPZIP	1993-2003	Hospital ZIP Code from AHA Annual Survey of Hospitals (not available for all states)	GA, HI, IN, KS, MI, NE, OH, SC, SD, TN, TX
Hospital characteristics	HOSP_BEDSIZE	1998-2003	Bed size of hospital: (1) small, (2) medium, (3) large	
	H_BEDSZ	*1993-1997*	*Bed size of hospital: (1) small, (2) medium, (3) large*	
	ST_BEDSZ	*1988-1992*	*Bed size of hospital: (1) small, (2) medium, (3) large*	
	HOSP_CONTROL	1998-2003	Control/ownership of hospital: (0) government or private, collapsed category, (1) government, nonfederal, public, (2) private, non-profit, voluntary, (3) private, invest-own, (4) private, collapsed category	
	H_CONTRL	*1993-1997*	*Control/ownership of hospital: (1) government, nonfederal (2) private, non-profit (3) private, invest-own*	
	ST_OWNER	*1988-1992*	*Control/ownership of hospital: (1) public (2) private, non-profit (3) private for profit*	
	HOSP_ LOCATION	1998-2003	Location: (0) rural, (1) urban	
	H_LOC	*1993-1997*	*Location: (0) rural, (1) urban*	
	HOSP_ LOCTEACH	1998-2003	Location/teaching status of hospital: (1) rural, (2) urban non-teaching, (3) urban teaching	

Type of Data Element	HCUP Variable Name	Years Available	Coding Notes	Unavailable in 2003 for:
Hospital characteristics (continued)	*H_LOCTCH*	*1993-1997*	*Location/teaching status of hospital: (1) rural, (2) urban non-teaching, (3) urban teaching*	
	LOCTEACH	*1988-1992*	*Location/teaching status of hospital: (1) rural, (2) urban non-teaching, (3) urban teaching*	
	HOSP_REGION	1998-2003	Region of hospital: (1) Northeast, (2) Midwest, (3) South, (4) West	
	H_REGION	*1993-1997*	*Region of hospital: (1) Northeast, (2) Midwest, (3) South, (4) West*	
	ST_REG	*1988-1992*	*Region of hospital: (1) Northeast, (2) Midwest, (3) South, (4) West*	
	HOSP_TEACH	1998-2003	Teaching status of hospital: (0) non-teaching, (1) teaching	
	H_TCH	*1993-1997*	*Teaching status of hospital: (0) non-teaching, (1) teaching*	
	NIS_STRATUM	1998-2003	Stratum used to sample hospitals beginning in 1998; includes geographic region, control, location/teaching status, and bed size	
	STRATUM	*1988-1997*	*Stratum used to sample hospitals prior to 1998; includes geographic region, control, location/teaching status, and bed size*	
	STRAT_ST	*1988-1997*	*Stratum for State-specific weights*	
Hospital weights	HOSPWT	1998-2003	Weight to hospitals in AHA universe (i.e., total U.S.) beginning in 1998	
	HOSPWT_U	*1988-1997*	*Weight to hospitals in AHA universe (i.e., total U.S.) prior to 1998*	
	HOSPWT_F	*1988-1997*	*Weight to hospitals in the sample frame.*	
	HOSPWT_S	*1988-1997*	*Weight to hospitals in the State*	

Table 5. Data Elements in the NIS Disease Severity Measures Files

All data elements listed below are available in the 2003 NIS Disease Severity Measures Files.

Type of Data Element	HCUP Variable Name	Years Available	Coding Notes	Unavailable in 2003 for:
AHRQ Comorbidity Software (AHRQ)	CM_AIDS	2002-2003	AHRQ comorbidity measure: Acquired immune deficiency syndrome	PA
	CM_ALCOHOL	2002-2003	AHRQ comorbidity measure: Alcohol abuse	PA
	CM_ANEMDEF	2002-2003	AHRQ comorbidity measure: Deficiency anemias	PA
	CM_ARTH	2002-2003	AHRQ comorbidity measure: Rheumatoid arthritis/collagen vascular diseases	PA
	CM_BLDLOSS	2002-2003	AHRQ comorbidity measure: Chronic blood loss anemia	PA
	CM_CHF	2002-2003	AHRQ comorbidity measure: Congestive heart failure	PA
	CM_CHRNLUNG	2002-2003	AHRQ comorbidity measure: Chronic pulmonary disease	PA
	CM_COAG	2002-2003	AHRQ comorbidity measure: Coagulopathy	PA
	CM_DEPRESS	2002-2003	AHRQ comorbidity measure: Depression	PA
	CM_DM	2002-2003	AHRQ comorbidity measure: Diabetes, uncomplicated	PA
	CM_DMCX	2002-2003	AHRQ comorbidity measure: Diabetes with chronic complications	PA
	CM_DRUG	2002-2003	AHRQ comorbidity measure: Drug abuse	PA
	CM_HTN_C	2002-2003	AHRQ comorbidity measure: Hypertension, uncomplicated and complicated	PA
	CM_HYPOTHY	2002-2003	AHRQ comorbidity measure: Hypothyroidism	PA
	CM_LIVER	2002-2003	AHRQ comorbidity measure: Liver disease	PA
	CM_LYMPH	2002-2003	AHRQ comorbidity measure: Lymphoma	PA
	CM_LYTES	2002-2003	AHRQ comorbidity measure: Fluid and electrolyte disorders	PA
	CM_METS	2002-2003	AHRQ comorbidity measure: Metastatic cancer	PA
	CM_NEURO	2002-2003	AHRQ comorbidity measure: Other neurological disorders	PA
	CM_OBESE	2002-2003	AHRQ comorbidity measure: Obesity	PA
	CM_PARA	2002-2003	AHRQ comorbidity measure: Paralysis	PA
	CM_PERIVASC	2002-2003	AHRQ comorbidity measure: Peripheral vascular disorders	PA
	CM_PSYCH	2002-2003	AHRQ comorbidity measure: Psychoses	PA
	CM_PULMCIRC	2002-2003	AHRQ comorbidity measure: Pulmonary circulation disorders	PA
	CM_RENLFAIL	2002-2003	AHRQ comorbidity measure: Renal failure	PA
	CM_TUMOR	2002-2003	AHRQ comorbidity measure: Solid tumor without metastasis	PA
	CM_ULCER	2002-2003	AHRQ comorbidity measure: Peptic ulcer disease excluding bleeding	PA

Type of Data Element	HCUP Variable Name	Years Available	Coding Notes	Unavailable in 2003 for:
AHRQ Comorbidity Software (continued)	CM_VALVE	2002-2003	AHRQ comorbidity measure: Valvular disease	PA
	CM_WGHTLOSS	2002-2003	AHRQ comorbidity measure: Weight loss	PA
All Patient Refined DRG (3M)	APRDRG	2002-2003	All Patient Refined DRG	PA
	APRDRG_Risk_Mortality	2002-2003	All Patient Refined DRG: Risk of Mortality Subclass	PA
	APRDRG_Severity	2002-2003	All Patient Refined DRG: Severity of Illness Subclass	PA
All-Payer Severity-adjusted DRG (HSS, Inc.)	APSDRG	2002-2003	All-Payer Severity-adjusted DRG	PA
	APSDRG_Mortality_Weight	2002-2003	All-Payer Severity-adjusted DRG: Mortality Weight	PA
	APSDRG_LOS_Weight	2002-2003	All-Payer Severity-adjusted DRG: Length of Stay Weight	PA
	APSDRG_Charge_Weight	2002-2003	All-Payer Severity-adjusted DRG: Charge Weight	PA
Disease Staging (Medstat)	DS_DX_Category1	2002-2003	Disease Staging: Principal Disease Category	PA
	DS_Stage1	2002-2003	Disease Staging: Stage of Principal Disease Category	PA
	DS_LOS_Level	2002-2003	Disease Staging: Length of Stay Level	PA
	DS_LOS_Scale	2002-2003	Disease Staging: Length of Stay Scale	PA
	DS_Mrt_Level	2002-2003	Disease Staging: Mortality Level	PA
	DS_Mrt_Scale	2002-2003	Disease Staging: Mortality Scale	PA
	DS_RD_Level	2002-2003	Disease Staging: Resource Demand Level	PA
	DS_RD_Scale	2002-2003	Disease Staging: Resource Demand Scale	PA
Linkage Variables	HOSPID	2002-2003	HCUP hospital identification number	PA
	KEY	2002-2003	HCUP record identifier	PA

SAMPLING OF HOSPITALS INCLUDED IN THE NIS

The hospital universe is defined by all hospitals that were open during any part of each calendar year and were designated as community hospitals in the AHA Annual Survey of Hospitals.

For more information on how hospitals in the data were mapped to hospitals as defined by the AHA, refer to the special report: *HCUP Hospital Identifiers*. For a list of all data sources, refer to: *Sources of NIS Data and State-Specific Restrictions*. For more detailed descriptions of the sampling design, refer to the year-specific special reports *Design of the HCUP Nationwide Inpatient Sample*. All reports can be found on the NIS Documentation CD-ROM.

Stratification Variables

To help ensure generalizability, five hospital sampling strata were defined based on hospital characteristics contained in the AHA Annual Survey of Hospitals. The stratification variables are:

1) *Geographic Region – Northeast, Midwest, West, or South*. This is based on the U.S. Census regions.
2) *Location – urban or rural*. A metropolitan statistical area is considered urban.
3) *Teaching Status – teaching or non-teaching*. A hospital is considered to be a teaching hospital if it has an AMA-approved residency program, is a member of the Council of Teaching Hospitals (COTH) or has a ratio of full-time equivalent interns and residents to beds of .25 or higher.
4) *Control – government nonfederal (public), private not-for-profit (voluntary) or private investor-owned (proprietary).* When there were enough hospitals of each type to allow it (southern rural, southern urban non-teaching, and western urban non-teaching), hospitals were stratified as public, voluntary, and proprietary. For smaller strata (north central rural and western rural hospitals) a collapsed stratification of public versus private was used, with the voluntary and proprietary hospitals combined to form a single 'private' category. For all other combinations of region, location and teaching status, no stratification based on control was advisable given the number of hospitals in these cells.
5) *Bed size – small, medium, or large.* Bed size categories are based on hospital beds, and are specific to the hospital's location and teaching status, as shown in Table 6 (on the following page). Bed size cutpoints were chosen so that approximately one-third of the hospitals in a given region and location/teaching combination would be in each bed size category (small, medium, or large). Rural hospitals were not split according to teaching status, because rural teaching hospitals were rare.

Table 6. Bed Size Categories, by Region

Location and Teaching Status	Hospital Bed size		
	Small	Medium	Large
NORTHEAST			
Rural	1-49	50-99	100+
Urban, non-teaching	1-124	125-199	200+
Urban, teaching	1-249	250-424	425+
MIDWEST			
Rural	1-29	30-49	50+
Urban, non-teaching	1-74	75-174	175+
Urban, teaching	1-249	250-374	375+
SOUTH			
Rural	1-39	40-74	75+
Urban, non-teaching	1-99	100-199	200+
Urban, teaching	1-249	250-449	450+
WEST			
Rural	1-24	25-44	45+
Urban, non-teaching	1-99	100-174	175+
Urban, teaching	1-199	200-324	325+

To further ensure geographic representativeness, implicit stratification variables included state and three-digit zip code (the first three digits of the hospital's five digit zip code). The hospitals were sorted according to these variables prior to systematic random sampling. For more detailed descriptions of the stratification and sample design, refer to the year-specific special reports *Design of the HCUP Nationwide Inpatient Sample* which can be found on the HCUP-US Website and on the NIS Documentation CD-ROM.

GETTING STARTED

NIS information is provided on two CD-ROMs. The NIS inpatient core and hospital weights data files are on CD-ROM #1, and the NIS documentation and tools and the disease severity measures files are on CD-ROM #2.

NIS Data Files

In order to load and analyze the NIS data onto your PC, you will need 10 gigabytes of space available. Because of the size of the files, the data are distributed as self-extracting PKZIP compressed files. To decompress the data, you should follow these steps:

1. Create a directory for the NIS on your hard drive.
2. Copy the self-extracting data files from the NIS Data Files CD-ROM into the new directory.
3. Unzip each file by running the corresponding *.exe file.
    - Type the file name within DOS or click on the name within Windows Explorer.
    - Edit the name of the "Unzip To Folder" in the WinZip Self-Extractor dialog to select the desired destination directory for the extracted file.
    - Click on the "Unzip" button.

The ASCII data files will then be uncompressed into this directory. After the files are uncompressed, the *.exe files can be deleted.

NIS Documentation

NIS documentation files on the Documentation CD-ROM provide important resources for the user. Refer to these resources to understand the structure and content of the NIS and to aid in using the NIS. Many of the documentation files are provided in portable document format (*.pdf) files. Files with the *.pdf extension can be viewed, searched, and printed using the Adobe Acrobat Reader®.

You must have the Adobe Acrobat Reader software on your computer to access the NIS documentation. If you do not have Adobe Acrobat Reader software on your computer, see the DOCUMENTATION.README.TXT file on NIS Documentation CD-ROM for instructions on installing or obtaining the software.

The Acrobat Reader provided on the NIS Documentation CD-ROM is for IBM-compatible microcomputers running Microsoft Windows 98, second edition, or higher. More information and Acrobat Reader software for other platforms (DOS, Windows 3.1, Macintosh, Sun Systems, etc.) may be obtained free of charge from the Adobe Home Page at http://www.adobe.com/. For further assistance in installing and running the Adobe Acrobat Reader on your computer platform, please consult your local support personnel.

Table 7 describes the documentation, tools, and the severity adjustment data files that can be found on the NIS Documentation CD-ROM and illustrates the structure of the directories and subdirectories on the CD. All NIS documentation is also available on the AHRQ-sponsored HCUP User Support Website at http://www.hcup-us.ahrq.gov.

Table 7. NIS Documentation CD-ROM, including Severity Adjustment Files

Directory	Description
Root	Includes DOCUMENTATION.README.TXT file with introductory information on accessing the NIS documentation
/Adobe Acrobat Reader	Adobe Acrobat Reader files for IBM compatible for Microsoft Windows 95 or higher. (One text, one HTML, and one application file)
/Description of Data Elements	Includes information on all NIS variables such as uniform coding and state-specific information: Core (for data elements starting with letters A-L and letters M-Z), Disease Severity Measures, and Hospital Weights. (PDF files)
/File Specifications	Includes data set name, number of records, record length, and record layout. One file per data file: Core, Core Subsample #1, Core Subsample #2, Disease Severity Measures, Subsample #1 Disease Severity Measures, Subsample #2 Disease Severity Measures, and Hospital Weights. (Text files)
/General Information	Includes: • Introduction to the NIS, 2003 (PDF file) – *this document* • Availability of NIS data elements across all years (PDF file) • Sources of NIS Data and State-Specific Restrictions (PDF file)
/HCUP Tools_Labels	Includes: • Label file for the Clinical Classifications Software (CCS), a categorization scheme that groups ICD-9-CM diagnosis and procedure codes into mutually exclusive categories. (Text file) • Label file for Diagnosis Related Groups (DRGs), multiple versions provided (Text file) • SAS code to create format library of variable labels (Text file)
/SAS Load Programs	SAS programming code to convert ASCII data files into SAS. One file per data file: Core, Core Subsample #1, Core Subsample #2, Disease Severity Measures, Subsample #1 Disease Severity Measures, Subsample #2 Disease Severity Measures, and Hospital Weights. (Text files)
/Severity Adjustment Files	Includes self-extracting compressed data files with the four sets of severity measures: Disease Severity Measures, Subsample #1 Disease Severity Measures, Subsample #2 Disease Severity Measures.
/Severity Adjustment Reports	Includes information on each of the severity measures: Overview of Severity Systems, documentation for APR-DRGS, documentation of APS-DRGS, and documentation for Disease Staging.
/Special Reports	Includes: • Design of the Nationwide Inpatient Sample, 2003 (PDF file) • Changes in NIS Sampling and Weighting Strategy for 1998 (PDF file) • NIS Comparison Report (PDF file) • Calculating Nationwide Inpatient Sample Variances (PDF file) • Using the HCUP Nationwide Inpatient Sample to Estimate Trends • HCUP Coding Practices (PDF file) • HCUP Quality Control Procedures (PDF file) • HCUP Hospital Identifiers (PDF file)
/SPSS Load Programs	SPSS programming code to convert ASCII data files into SPSS. One file per data file: Core, Core Subsample #1, Core Subsample #2, Disease Severity Measures, Subsample #1 Disease Severity Measures, Subsample #2 Disease Severity Measures, and Hospital Weights. (Text files)
/Summary Statistics	Includes summary statistics (means and frequencies) on NIS data. At least one file per data file: Core (weighted and unweighted statistics and data quality report), Core Subsample #1, Core Subsample #2, Disease Severity Measures, Subsample #1 Disease Severity Measures, Subsample #2 Disease Severity Measures, and Hospital Weights. (PDF files)

HOW TO USE THE NIS

This section provides a brief synopsis of special considerations when using the NIS. For more details see detailed documentation under Special Reports, Description of Data Elements, and Summary Statistics.

- If anyone other than the original purchaser uses the NIS data, be sure to have them read and sign a data use agreement. A copy of the signed data use agreements must be sent to AHRQ. See page 2 for the mailing address. A copy of the NIS Data Use Agreement is provided in the binder you receive when you purchase the NIS and is available on the HCUP User Support Website: http://www.hcup-us.ahrq.gov.

- The NIS contains <u>discharge</u>-level records, not <u>patient</u>-level records. This means that individual patients who are hospitalized multiple times in one year may be present in the NIS multiple times. There is no uniform patient identifier available that allows a patient-level analysis with the NIS. This will be especially important to remember for certain conditions for which patients may be hospitalized multiple times in a single year.

Creating National Estimates

- To produce national estimates, use one of the following discharge weights to weight discharges in the NIS Core files to the discharges from all U.S. community, non-rehabilitation hospitals. The name of the discharge weight data element depends on the year of data, the type of analysis, and whether you are using the Core file or the 10% subsample Core file.

NIS Year	Name of Discharge Weight on the **Core File** to Use for Creating Nationwide Estimates	Name of Discharge Weight on the **10% Subsample Core File** to Use for Creating Nationwide Estimates
2001 - 2003	• DISCWT for all analyses	• DISCWT10 for all analyses
2000	• DISCWT to create nationwide estimates for all analyses <u>except</u> those that involve total charges.	• DISCWT10 to create nationwide estimates for all analyses <u>except</u> those that involve total charges.
	• DISCWTCHARGE to create nationwide estimates of total charges.	• DISCWTCHARGE10 to create nationwide estimates of total charges.
1998-1999	• DISCWT for all analyses	• DISCWT10 for all analyses
1988-1997	• DISCWT_U for all analyses	• D10CWT_U for all analyses

- Because the NIS is a stratified sample, proper statistical techniques must be used to calculate standard errors and confidence intervals. For detailed instructions, refer the special report *Calculating Nationwide Inpatient Sample Variances* on the HCUP-US website and the NIS Documentation CD-ROM.

- The NIS Comparison Report assesses the accuracy of NIS estimates. The most recent report is available on the NIS Documentation CD-ROM and provides a comparison of a previous year's NIS with other data sources. The updated report for the current NIS will be posted on the HCUP User Support Website (www.hcup-us.ahrq.gov) as soon as it is completed.

- When creating national estimates, it is a good idea to check your estimates against other data sources, if available. For example, the National Hospital Discharge Survey (http://www.cdc.gov/nchs/products/pubs/pubd/series/sr13/ser13.htm) can provide benchmarks against which to check your national estimates for hospitalizations with more than 5000 cases.

- When using the 10% subsample files, use the specially created subsample weight, DISCWT10 (the discharge weight multiplied by 10).   When using the hospital weights with the subsample files, there is no need to multiply the hospital weights because all hospitals will be represented in the subsample files.

Why the NIS Should Not Be Used to Make State-Level Estimates

AHRQ strongly advises researchers against using the NIS to estimate State-specific statistics. Prior to 2012, State is available as a NIS data element. However, these NIS samples were not designed to yield a representative sample of hospitals at the State level. AHRQ recommends that researchers employ the SID for State-level estimates.

Each NIS sample is drawn from the sampling frame consisting of discharge data submitted by HCUP Partners-statewide data organizations that agree to participate in the NIS. Data from non-Partner States are missing completely from the sampling frame, and data from Partner States are sometimes incomplete because of different State reporting requirements, different State restrictions, or other data omissions. The NIS is designed to represent hospitals and discharges nationally, including those outside the sampling frame.

To accomplish this, within each hospital sampling stratum the NIS draws a number of hospitals from the sampling frame required to net a total of 20 percent of hospitals nationally. The sampling strata are defined by census region (4 regions), hospital ownership (3 categories), urban-rural location, teaching status, and bed size (3 categories). As a result, the proportion of NIS hospitals in a stratum that are from a given State is unlikely to equal the State's actual proportion of hospitals in that stratum. Consequently, the sample of NIS hospitals is unlikely to be representative of hospitals in the State, and the NIS sample weights will not be appropriate at the State level.

The level of this "misrepresentation" varies across the States in any given year of the NIS, which further confounds State-to-State comparisons on the basis of State-specific estimates from the NIS. Moreover, for a given State the level of misrepresentation changes from year to year as States (and hospitals) enter and exit the sampling frame over time. This further confounds State-specific trends on the basis of State-specific estimates from the NIS.

Finally, because the NIS was not designed to be representative at the State level, design-based estimates of standard errors are not possible, which severely hampers State-level inferences. Moreover, the NIS is composed of all discharges from a sample of hospitals (a cluster sample). The hospital-to-hospital variation and the small number of hospitals available in the NIS for many States make Statelevel estimates very imprecise at best and biased at worst.

Studying Trends

- When studying trends over time using the NIS, be aware that the sampling frame for the NIS changes almost annually, i.e., more states have been added over time. Estimates from earlier years of the NIS may be subject to more sampling bias than later years of the NIS. In order to facilitate analysis of

trends using multiple years of NIS data, an alternate set of NIS discharge and hospital weights for the 1988-1997 HCUP NIS were developed. These alternative weights were calculated in the same way as the weights for the 1998 and later years of the NIS. The *NIS Trends Report* includes details regarding the alternate weights and other recommendations for trends analysis. Both the NIS Trends Report and the alternative weights are available on the HCUP-US web site under Method Series (http://www.hcup-us.ahrq.gov/reports/methods.jsp). The NIS Trends Report is also available on the NIS Documentation CD-ROM.

- Short-term rehabilitation hospitals are included in the 1988-1997 NIS, but are excluded from the NIS beginning in 1998. Patients treated in short-term rehabilitation hospitals tend to have lower mortality rates and longer lengths of stay than patients in other community hospitals. The elimination of rehabilitation hospitals may affect trends but the effect is likely small since only about 3 percent of community hospitals are short-term rehabilitation hospitals and not all state data sources included short term rehabilitation hospitals

Choosing Data Elements for Analysis

- For all data elements you plan to use in your analysis, first perform descriptive statistics and examine the range of values, including number of missing cases. Summary statistics for the entire NIS are provided in the "Summary Statistics" directory on the NIS Documentation CD-ROM. When you detect anomalies (such as large numbers of missing cases), perform descriptive statistics by state for that variable to detect if there are state-specific differences. Sometimes performing descriptive statistics by hospital can be helpful to detect hospital-specific data anomalies.

- Not all data elements in the NIS are provided by each state data source. These data elements are provided on the NIS because they can be valuable for research purposes but they should be used cautiously. For example, RACE is missing for a number of states thus national estimates using RACE should be interpreted and reported with caveats. Check the documentation and run frequencies by state to identify if a data element is not available in one or more states.

- There are differences across the state data sources in the collection of information that could not be accounted for during HCUP processing to make the data uniform. Be sure to read state-specific notes for each data element that you use in your analysis – this information can be found under "Description of Data Elements" on the NIS Documentation CD-ROM.

- Data elements with "_X" suffixes contain state-specific coding, i.e., these data elements are provided by the data sources and have not been altered in any way. For some data elements (e.g., LOS_X and TOTCHG_X) this means that no edit checks have been applied. For other data elements (e.g., PAY1_X), the coding is specific to each state and may not be comparable to any other state.

ICD-9-CM Diagnosis and Procedure Codes

- ICD-9-CM diagnosis and procedure codes provide valuable insights into the reasons for hospitalization and what procedures patients receive, but these codes need to be carefully used and interpreted. ICD-9-CM codes change every October as new codes are introduced and some codes are retired.

- Although the NIS contains up to 15 diagnoses and 15 procedures, the number of diagnoses and procedures varies by state. Some states provide as many as 30 diagnoses and 21 procedures, while other states provide as few as 10 diagnoses and 6 procedures. Because very few cases have more than 15 diagnoses or procedures, the diagnosis and procedure vectors were truncated to save space in the NIS data files. Two variables are provided which tell you exactly how many diagnoses and

procedures were on the original records (NDX and NPR).

- The collection and reporting of external cause of injury (E-codes) varies greatly across states. Some States have laws or mandates for the collection of E-codes; others do not. Some States do not require hospitals to report E-codes in the range 870-879 - "misadventures to patients during surgical and medical care" - which means that these occurrences will be underreported. Be sure to read the state-specific notes on diagnoses for more details - this can be found under "Description of Data Elements" on the NIS Documentation CD-ROM.

OTHER HCUP PRODUCTS

Information on HCUP products and services is available on the World Wide Web on the AHRQ Website http://www.ahrq.gov/data/hcup/. HCUP User Support is available at http://www.hcup-us.ahrq.gov.

DATABASES

Nationwide Inpatient Sample (NIS) is a nationwide database of hospital inpatient stays. The NIS is the largest all-payer inpatient care database that is publicly available in the United States, containing data from 5 to 8 million hospital stays from about 1,000 hospitals sampled to approximate a 20-percent stratified sample of U.S. community hospitals. The NIS is available for a 16-year time period, from 1988 to 2003, allowing analysis of trends over time. For more information, visit the HCUP User Support Website at http://www.hcup-us.ahrq.gov or contact the HCUP Central Distributor (detailed below).

State Inpatient Databases (SID) are hospital inpatient databases from Data Organizations participating in HCUP. The SID contain the universe of the inpatient discharge abstracts in the participating HCUP States, translated into a uniform format to facilitate multi-State comparisons and analyses. Together, the SID encompass about 90 percent of all U.S. community hospital discharges in 2003. For more information, visit the HCUP User Support Website at http://www.hcup-us.ahrq.gov or contact the HCUP Central Distributor (see below).

State Ambulatory Surgery Databases (SASD) are outpatient databases from Data Organizations in participating HCUP States, which capture surgeries performed on the same day in which patients are admitted and released. The SASD contain the ambulatory surgery encounter abstracts in participating States, translated into a uniform format to facilitate multi-State comparisons and analyses. All of the databases include abstracts from hospital-affiliated ambulatory surgery sites. Some contain the universe of ambulatory surgery encounter abstracts for that State, including records from both hospital-affiliated and freestanding surgery centers. Composition and completeness of data files may vary from State to State. For more information, visit the HCUP User Support Website at http://www.hcup-us.ahrq.gov or contact the HCUP Central Distributor (see below).

Kids' Inpatient Database (KID) is a unique database of hospital inpatient stays for children 18 years of age and younger. The 1997 and 2000 KID was specifically designed to permit researchers to study a broad range of conditions and procedures related to child health issues. The KID is created and released every three years. For more information, visit the HCUP User Support Website at http://www.hcup-us.ahrq.gov or contact the HCUP Central Distributor (see below).

HCUP CENTRAL DISTRIBUTOR

HCUP databases are available for purchase through the AHRQ-sponsored HCUP Central Distributor. All years of the NIS and KID are released through the HCUP Central Distributor. In addition, many of the HCUP State Partners allow the public release of the HCUP State Inpatient Databases (SID) and State Ambulatory Surgery Databases (SASD) through the HCUP Central Distributor. Application Kits for purchasing the HCUP databases are available online at http://www.hcup-us.ahrq.gov or contact the HCUP Central Distributor directly. Information on how to obtain uniformly-formatted HCUP files from States not participating in the HCUP Central Distributor is also available from the HCUP Central Distributor:

HCUP Central Distributor
Phone: (866) 556-4287 (toll-free)
FAX: (301) 628-3201
E-mail: hcup@s-3.com

HCUP USER SUPPORT

HCUP User Support (HCUP-US) provides technical assistance to all HCUP users and is designed to facilitate the use of HCUP data, software tools, and products. The goals of this service are to increase awareness of the strengths and uses of HCUP data and to enhance the skills of individuals using the data for research, education, and policy analysis. A user-friendly Website for HCUP-US is located at http://www.hcup-us.ahrq.gov. This site includes links to information on how to purchase and understand the HCUP databases, as well as links to HCUP User Support Services and Frequently Asked Questions. For further information, consultants are available via both telephone and E-mail to help in planning analytic research and to offer advice about appropriate uses of HCUP data.

HCUPnet

HCUPnet is a Web-based query tool for identifying, tracking, analyzing, and comparing statistics on hospitals at the national, regional, and state level. With HCUPnet you have easy access to national statistics and trends and selected state statistics about hospital stays. HCUPnet guides you step-by-step to obtain the statistics you need. HCUPnet generates statistics using the Nationwide Inpatient Sample (NIS), the Kids' Inpatient Database (KID), and the State Inpatient Databases (SID) for those states that have agreed to participate. In addition, HCUPnet provides Quick Statistics – ready-to-use tables on commonly requested information – and national statistics based on the AHRQ Quality Indicators. HCUPnet can be found at: http://www.ahrq.gov/hcupnet.

TOOLS

AHRQ Quality Indicators (QIs) are clinical performance measures for use with readily available inpatient data. Methods and software for the AHRQ Quality Indicators can be downloaded from http://www.qualityindicators.ahrq.gov.

Clinical Classifications Software (CCS), formerly known as the Clinical Classifications for Health Policy Research (CCHPRs), are classification systems that group ICD-9-CM diagnoses and procedures into a limited number of clinically meaningful categories. Methods and software can be downloaded from the HCUP User Support Website, Tools and Software page, at http://www.hcup-us.ahrq.gov/tools_software.jsp.

Comorbidity Software assigns variables that identify comorbidities in hospital discharge records using ICD-9-CM diagnosis codes (International Classification of Diseases, Ninth Revision, Clinical Modification). Methods and software can be downloaded from the HCUP User Support Website, Tools and Software page, at http://www.hcup-us.ahrq.gov/tools_software.jsp.

Procedure Classes identify whether a procedure is (a) diagnostic or therapeutic, and (b) minor or major in terms of invasiveness and/or resource use. Software can be found at http://www.hcup-us.ahrq.gov/tools_software.jsp.

PUBLICATIONS

HCUP Research Notes and Fact Books report aggregate statistics and detailed analyses using HCUP data. To request copies, contact the AHRQ Publications Clearinghouse at (800) 358-9295 or send a postcard to: AHRQ Publications Clearinghouse, P.O. Box 8547, Silver Spring, MD 20907 or visit http://www.hcup-us.ahrq.gov/reports.jsp.