

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

OVERVIEW OF
THE HCUP STATE INPATIENT DATABASES (SID)

These pages provide only an introductory overview of the SID.
Full documentation is provided on the SID Documentation CD-ROM.

Issued June 2004

Agency for Healthcare Research and Quality
Healthcare Cost and Utilization Project (HCUP)
540 Gaither Road, 5th Floor
Rockville, Maryland 20850

Phone: (301) 427-1410

Fax: (301) 427-1430

E-mail: hcup@ahrq.gov

Website: <http://www.hcup-us.ahrq.gov>

SID Data and Documentation Distributed by:

HCUP Central Distributor
Social & Scientific Systems, Inc.

Phone: (866) 556-4287 (toll-free)

Fax: (301) 628-3201

Internet: hcup@s-3.com

Table of Contents

SUMMARY OF DATA USE LIMITATIONS	1
HCUP CONTACT INFORMATION	2
ABSTRACT	3
INTRODUCTION TO THE HCUP STATE INPATIENT DATABASES (SID)	4
OVERVIEW OF THE SID	4
How the HCUP SID Differ from State Data Files.....	6
What Types of Hospitals Are Included in the SID?	6
How to Identify Hospitals in the SID	6
What is the File Structure of the SID in the 1998-2002 Files?	7
What is the File Structure of the SID in the 1995-1997 Files?	8
GETTING STARTED	9
SID Data Files.....	9
SID Documentation.....	9
OTHER HCUP PRODUCTS	11
Databases.....	11
HCUP Central Distributor	11
HCUP User Support	12
HCUPnet.....	12
Tools	12
Publications	12
DATA USE AGREEMENT FOR THE STATE INPATIENT DATABASES	13

HCUP STATE INPATIENT DATABASES (SID) SUMMARY OF DATA USE LIMITATIONS

***** REMINDER *****

All users of the SID must sign a data use agreement and send a copy to AHRQ. †

Authorized users of HCUP data agree to the following limitations: ‡

- 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 identify or attempt to identify any individual.
- Will not link HCUP data to data from another source that identifies individuals.
- Will not report information that could identify individual establishments (e.g., hospitals).
- Will not use the data concerning individual establishments for commercial or competitive purposes involving those establishments.
- Will not use the data to determine rights, benefits, or privileges of individual establishments.
- Will not identify or attempt to identify any establishment when its identity has been concealed on the database.
- Will not contact establishments included in the data.
- Will not attribute to data contributors any conclusions drawn from the data.
- Must acknowledge the "Healthcare Cost and Utilization Project, (HCUP)", as described in the Data Use Agreement, in reports.

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.

† A copy of the Data Use Agreement is included on the SID Documentation CD-ROM and is included at the end of this document. See next page for AHRQ's address.

‡ Specific provisions are detailed in the Data Use Agreement for HCUP State Inpatient Databases.

HCUP CONTACT INFORMATION

Please submit signed data use agreements to HCUP at:

Agency for Healthcare Research and Quality
Healthcare Cost and Utilization Project (HCUP)
540 Gaither Road, 5th Floor
Rockville, MD 20850

Phone: (301) 427-1410

Fax: (301) 427-1430

For technical assistance,

Visit the AHRQ-sponsored HCUP User Support Website at

<http://www.hcup-us.ahrq.gov>

Or send an e-mail to HCUP User Support at

hcup@ahrq.gov

Or contact the HCUP Central Distributor at

HCUP Central Distributor
Social & Scientific Systems, Inc.
Phone: (866) 556-4287 (toll-free between the hours of 9 a.m.
and 5 p.m. (ET). If the HCUP Central Distributor is not
immediately available, please leave a message on the voice
mail, and your call will be returned within one business day.)

Fax: (301) 628-3201

E-mail: hcup@s-3.com

We would like to receive your feedback on the HCUP SID data products.

Our Internet address for user feedback is

hcup@ahrq.gov.

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 State Inpatient Databases (SID).**

HCUP State Inpatient Databases (SID)

ABSTRACT

The State Inpatient Databases (SID) are 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 HCUP State Inpatient Databases (SID) are a powerful set of hospital databases from data organizations in participating States.

- The SID contain the universe of the inpatient discharge abstracts in participating States, translated into a uniform format to facilitate multi-State comparisons and analyses.
- Together, the SID encompass about 85 percent of all U.S. hospital discharges. Some States include discharges from specialty hospitals, such as acute psychiatric hospitals.
- The SID contain a core set of clinical and nonclinical information on all patients, regardless of payer, including persons covered by Medicare, Medicaid, private insurance, and the uninsured.
- In addition to the core set of uniform data elements common to all SID, some include other elements, such as the patient's race.

Researchers and policymakers use SID to investigate questions unique to one State; to compare data from two or more States; to conduct market area variation analyses; and to identify State-specific trends in inpatient care utilization, access, charges, and outcomes.

Twenty of the Data Organizations participating in the HCUP have agreed to release their SID files through the HCUP Central Distributor under the auspices of the AHRQ. The individual state databases are in the same HCUP uniform format and represent 100% of records processed by AHRQ. However, the participating Data Organizations control the release of specific data elements. AHRQ is currently assisting the Data Organizations in the release of the 1995-2002 SID.

The SID 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.

Access to the SID is available through the HCUP Central Distributor. Uses are limited to research and aggregate statistical reporting.

INTRODUCTION TO THE HCUP STATE INPATIENT DATABASES (SID)

OVERVIEW OF THE SID

The Healthcare Cost and Utilization Project (HCUP) State Inpatient Databases (SID) consist of individual data files from Data Organizations in 36 participating States. In general, the SID contain the universe of that state's hospital inpatient discharge records. They are composed of annual, state-specific files that share a common structure and common data elements. Most data elements are coded in a uniform format across all states. In addition to the core set of uniform data elements, the SID include state-specific data elements or data elements available only for a limited number of states. The uniform format of the SID helps facilitate cross-state comparisons. In addition, the SID are well suited for research that requires complete enumeration of hospitals and discharges within market areas or states.

Twenty of the Data Organizations participating in the HCUP have agreed to release their state-specific files through a Central Distributor under the auspices of the AHRQ. The individual state databases are in the same HCUP uniform format and represent 100% of records processed by AHRQ. However, the participating Data Organizations control the release of specific data elements.

SID data sets are currently available for multiple States and years, as shown in Table 1 on the next page. Each release of the SID includes:

- C Data in ASCII format on CD-ROM.
- C Patient-level hospital discharge abstract data for 100 percent of discharges from hospitals in participating States.
- C AHA Linkage File to link the SID to data from the American Hospital Association Annual Survey of Hospitals. This is only available for those states that allow the release of hospital identifiers.
- C SID 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. Number of Discharges in the SID by Year

State	1995	1996	1997	1998	1999	2000	2001	2002
AZ	476,833	493,005	516,064	538,585	560,237	603,911	634,019	Not Available
CA	3,629,322	3,632,167	3,685,706	3,725,422	3,775,711	3,816,887	3,864,090	Not Available
CO	387,772	392,379	402,615	414,367	427,941	441,943	455,363	Not Available
FL	1,899,586	1,931,135	2,004,247	2,095,009	2,151,589	2,232,533	2,343,137	2,386,628
IA	358,667	357,099	355,502	353,365	353,393	361,493	368,092	364,949
KY	Not Available	Not Available	Not Available	Not Available	Not Available	537,006	600,259	622,066
MA	775,216	759,869	763,363	777,922	777,429	788,501	808,583	817,207
MD	625,037	620,796	610,343	614,788	626,955	644,386	669,850	687,199
ME	Not Available	Not Available	Not Available	Not Available	158,294	161,355	161,988	161,898
MI	Not Available	Not Available	Not Available	Not Available	1,214,123	1,238,059	1,270,817	1,274,239
NC	Not Available	Not Available	Not Available	Not Available	Not Available	1,013,444	1,026,899	1,043,439
NE	Not Available	Not Available	Not Available	Not Available	Not Available	Not Available	188,015	198,047
NJ	1,079,678	1,070,992	1,070,524	1,082,746	1,091,947	1,119,445	1,136,089	1,153,654
NY	2,530,206	2,469,668	2,432,632	2,448,148	2,429,753	2,453,206	2,467,228	2,470,404
OR	350,263	369,814	376,069	381,134	386,485	350,747	361,095	Not Available
SC	440,315	446,156	461,308	483,202	498,955	517,018	540,035	Not Available
UT	Not Available	Not Available	225,487	232,476	233,328	241,450	247,056	253,659
WA	523,711	523,672	536,739	540,421	549,696	573,978	583,224	583,758
WI	622,933	614,102	610,962	611,307	621,235	635,166	648,015	Not Available
WV	Not Available	Not Available	Not Available	Not Available	Not Available	279,621	274,289	Not Available

How the HCUP SID Differ from State Data Files

The SID available through the HCUP Central Distributor differ from the data files available from the Data Organizations in the following ways:

- data elements available on the files and
- coding of data elements.

Because the Data Organizations dictate the data elements that may be released through the HCUP Central Distributor, the data elements on the SID are a subset of the data collected by the corresponding Data Organizations. HCUP uniform coding is used on most data elements on the SID. A few state-specific data elements retain the original values provided by the respective Data Organizations.

What Types of Hospitals Are Included in the SID?

What types of hospitals are included in the SID depends on the information provided by the Data Organizations and how the files were handled during HCUP processing. Most state government data organizations provide information on all acute care hospitals in the respective state. Private data organizations are often restricted to member hospitals and may not provide information on all hospitals in their state.

For the 1995-2002 SID, all hospitals reported by the Data Organizations were retained in the SID files and may include discharges from facilities such as psychiatric facilities, alcohol and drug dependency facilities, ambulatory surgery facilities, and State, Federal, and Veterans hospitals, if reported by the data source.

Some community hospitals as defined by the American Hospital Association (AHA) Annual Survey of Hospitals may not be included in the SID because their data were not provided by the data source. To identify community hospitals, the SID must be linked to the AHA Annual Survey of Hospitals by the AHA hospital identifier.

How to Identify Hospitals in the SID

Up to three hospital identifiers are on the SID:

- Some Data Organizations allow the original hospital identifier (DSHOSPID) to be included on the SID. If available on the SID, this identifier is coded for all hospitals and may distinguish different units within a hospital.
- Some Data Organizations allow the AHA hospital identifier (AHAID) to be included on the SID. This variable enables the SID to be linked to the AHA Annual Survey of Hospitals that contains information on hospital characteristics. The AHA hospital identifier is coded for most hospitals. The AHA hospital identifier is missing if the hospital is not registered with the AHA or the source-provided information could not be matched to the AHA.
- Some Data Organizations allow the HCUP-specific hospital identifier (HOSPID) to be included on the SID. HOSPID is coded for all hospitals with a nonmissing AHA hospital identifier.

What is the File Structure of the SID in the 1998-2002 Files?

Based on the availability of data elements across states, data elements included in the SID are structured as follows:

- Core file,
- Charges file, and
- AHA Linkage file.

The **Core file** contains:

- core data elements that form the nucleus of the SID, and
- State-specific data elements intended for limited use.

Core data elements meet at least one of the following criteria:

- are available from all or nearly all data sources,
- lend themselves to uniform coding across sources, or
- are needed for day-to-day applications (e.g., length of stay, patient age).

State-specific data elements meet at least one of the following criteria:

- are available from a limited number of sources,
- do not lend themselves to uniform coding across sources, or
- are not needed for day-to-day applications.

The **Charges file** contains detailed charge information (e.g., charges, rates, units, and revenue codes).

The **AHA Linkage file** contains AHA linkage data elements that allow the SID to be used in conjunction with the AHA Annual Survey of Hospitals data files. These files contain information about hospital characteristics and are available for purchase through the AHA. Since the Data Organizations in participating states determine whether the AHA linkage data elements may be released through the HCUP Central Distributor with the SID, not all SID include AHA linkage data elements.

The Core and Charges files are discharge-level files with one observation per abstract. The same record is represented in each file, but contains different data elements. To combine data elements across discharge-level files, merge the files by the unique record identifier (KEY). There will be a one-to-one correspondence of records.

The AHA Linkage file is a hospital-level file with one observation per hospital or facility. To combine discharge-level files with the hospital-level file (AHA Linkage file), merge the files by the hospital identifier provided by the data source (DSHOSPID), but be careful of the different levels of aggregation. For example, the Core file may contain 5,000 discharges for DSHOSPID "A", but the Hospital file contains only 1 record for DSHOSPID "A".

What is the File Structure of the SID in the 1995-1997 Files?

Based on the availability of data elements across states, data elements included in the SID are structured as follows:

- Core file,
- State-specific file, and
- AHA Linkage file.

The **Core file** contains core data elements that form the nucleus of the SID. Core data elements meet at least one of the following criteria:

- are available from all or nearly all data sources,
- lend themselves to uniform coding across sources, or
- are needed for day-to-day applications (e.g., length of stay, patient age).

The **State-specific file** contains state-specific data elements intended for limited use. State-specific data elements meet at least one of the following criteria:

- are available from a limited number of sources,
- do not lend themselves to uniform coding across sources, or
- are not needed for day-to-day applications.

The **AHA Linkage file** contains AHA linkage data elements that allow the SID to be used in conjunction with the American Hospital Association (AHA) Annual Survey of Hospitals data files. These files contain information about hospital characteristics and are available for purchase through the AHA. Since the Data Organizations in participating states determine whether the AHA linkage data elements may be released through the HCUP Central Distributor with the SID, not all SID include AHA linkage data elements.

The Core and State-specific files are discharge-level files with one observation per abstract. The same record is represented in each file, but contains different data elements. To combine data elements across discharge-level files merge the files by the unique record identifier (SEQ_SID). There will be a one-to-one correspondence of records.

The AHA Linkage file is a hospital-level file with one observation per hospital or facility. To combine discharge-level files with the AHA Linkage file, merge the files by the hospital identifier provided by the data source (DSHOSPID), but be careful of the different levels of aggregation. For example, the Core may contain 5,000 discharges for DSHOSPID "A", but the AHA Linkage file contains only 1 record for DSHOSPID "A".

GETTING STARTED

SID information is provided on CD-ROMs. The number of CD-ROMs depends on the state and year of data. The SID documentation and tools for all states and all years from 1995-2002 are on one CD-ROM. The SID data are on one or more CD-ROMs.

SID Data Files

In order to load SID data onto your PC, you will need between 1 and 4 gigabytes of space available, depending on which SID database you are using. 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 state-specific SID on your hard drive.
2. Copy the self-extracting data files from the SID Data Files CD-ROM(s) 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.

SID Documentation

SID documentation files on the SID Documentation CD-ROM provide important resources for the user. Refer to these resources to understand the structure and content of the SID and to aid in using the SID. 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 SID documentation. If you do not have Adobe Acrobat Reader software on your computer, it may be obtained free of charge from the Adobe Home Page at <http://www.adobe.com/>.

Table 2 describes the documentation and tools files that can be found on the SID Documentation CD-ROM and illustrates the structure of the directories and subdirectories on the CD.

Table 2. SID Documentation CD-ROM

Directory	Description
Root	Includes: <ul style="list-style-type: none"> • SID_DOCUMENTATION_README.TXT file with introductory information on accessing the SID documentation
/General Information	Includes: <ul style="list-style-type: none"> • Overview of the SID (PDF file) • SID File Composition (PDF file) • Data Use Agreement for the State Inpatient Databases (PDF file)
/Special Reports	Includes: <ul style="list-style-type: none"> • HCUP Coding Practices (PDF file) • HCUP Quality Control Procedures (PDF file) • HCUP Hospital Identifiers (PDF file)
/File Specifications	Includes data set name, number of records, record length, and record layout. One file per data file per year: Core, Charges (if applicable in 1998-2002), State-specific (if applicable in 1995-1997), and AHA Linkage (if applicable). File specifications are available for all years that a State is available. (Text files)
/Description of Data Elements	Includes: <ul style="list-style-type: none"> • Availability of Data Elements by State in 1995-1997 (PDF file) • Availability of Data Elements by State in 1998-2002 (PDF file) • Description of all SID data elements, including uniform coding and state-specific information, in two files organized alphabetically by data element name <ul style="list-style-type: none"> • DataElements_SID_Volume1_A-M (PDF file) • DataElements_SID_Volume2_N-Z (PDF file)
/SAS Load Programs	SAS programming code to convert ASCII data files into SAS. One file per data file per year: Core, Charges (if applicable in 1998-2002), State-specific (if applicable in 1995-1997), and AHA Linkage (if applicable). SAS load programs are available for all years that a State is available. (Text files)
/SPSS Load Programs	SPSS programming code to convert ASCII data files into SPSS. One file per data file: Core, Charges (if applicable), and AHA Linkage (if applicable). SPSS load programs are only available for data files beginning in 2000. (Text files)
/HCUP Tools_Labels	Includes: <ul style="list-style-type: none"> • Label file for the Clinical Classifications Software (CCSs), 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)

OTHER HCUP PRODUCTS

Information on HCUP products and services is available on the World Wide Web on the AHRQ Website for HCUP User Support <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 1000 hospitals sampled to approximate a 20-percent stratified sample of U.S. community hospitals. The NIS is available for a 15-year time period, from 1988 to 2002, 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 85 percent of all U.S. community hospital discharges. 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. 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
Social & Scientific Systems, Inc.
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. HCUPnet can be found at: <http://www.ahrq.gov/data/hcup/hcupnet.htm>.

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/data/hcup/qinext.htm>.

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 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 http://www.hcup-us.ahrq.gov/tools_software.jsp.

Procedure Classes facilitate research on hospital services using administrative data by identifying whether a procedure is (a) diagnostic or therapeutic, and (b) minor or major in terms of invasiveness and/or resource use. More information and software can be downloaded from http://www.hcup-us.ahrq.gov/tools_software.jsp.

PUBLICATIONS

Information on HCUP reports including new findings, publications, research notes based on HCUP data, and technical reports about HCUP issues can be found at <http://www.hcup-us.ahrq.gov/reports.jsp>.

DATA USE AGREEMENT FOR THE STATE INPATIENT DATABASES

This agreement must be signed by anyone seeking to use data in the State Inpatient Databases (SID) maintained by the Center for Delivery, Organization, and Markets (CDOM), Agency for Healthcare Research and Quality (AHRQ) before access to such data can be granted. All data maintained by CDOM/AHRQ is confidential or proprietary except data specified for restricted access public release, or data authorized by AHRQ and the original data source for re-release.

Under section 924(c) of the Public Health Service Act (42 U.S.C. 299c-3(c)), data that identifies individuals or establishments collected by the Agency for Healthcare Research and Quality (AHRQ) may be used only for the purpose for which they were collected. Data supplied to AHRQ under the auspices of HCUP were provided by the data sources only for research, analysis, and aggregate statistical reporting.

No identification of persons--Any effort to determine the identity of any person contained in the databases (including but not limited to patients, physicians, and other health care providers) or to use the information for any purpose other than for research, analysis, and aggregate statistical reporting would violate the conditions of this data use agreement and therefore the above-referenced AHRQ confidentiality statute. Furthermore, under the statute, no identifying information may be published or released in any way without the consent of the person who supplied the information or who can be identified by the information. AHRQ omits from the data set all direct personal identifiers, as well as characteristics that might lead to identification of persons. It may be possible in rare instances, through complex analysis and with outside information, to ascertain from the data sets the identity of particular persons. Considerable harm could ensue if this were done. By virtue of this agreement, the undersigned agrees that such attempts will be prohibited and that information which could identify individuals directly or by inference will not be released or published. Because of these restrictions, users of the data must agree that they will not attempt to contact individuals for the purpose of verifying information supplied in the HCUP databases. Any questions about the data must be referred to AHRQ only.

Use of Establishment identifiers--Section 924(c) of the Public Health Service Act (42 U.S.C. 299c-3(c)) also restricts the use of any information that allows the identification of establishments to the purpose for which the information was collected. Permission was obtained from the data sources (state data organizations, hospital associations, and data consortia) to use the identification of hospitals (when such identification appears in the data sets) for the purpose of conducting research only. Such research purpose includes linking institutional information from outside data sets for analysis and aggregate statistical reporting. Such purpose does not include the use of information in the data sets concerning individual establishments for commercial or competitive purposes involving those individual establishments, or to determine the rights, benefits, or privileges of establishments. Users of the data must not identify establishments directly or by inference in disseminated material. In addition, users of the data must not contact establishments for the purpose of verifying information supplied in the HCUP databases. Any questions about the data must be referred to AHRQ only.

The undersigned gives the following assurances with respect to the AHRQ data sets.

- I will not use nor permit others to use the data in these sets in any way except for research, analysis, and aggregate statistical reporting;
- I will require others in the organization (specified below) who use the data to sign this agreement (specifically acknowledging their agreement to abide by its terms) and will submit those signed agreements to AHRQ;
- I will ensure that the data are kept in a secured environment and that only authorized users have access to the data;
- I will not release nor permit others to release any information that identifies persons, directly or indirectly; I will not release information where the number of observations (i.e., discharge records) in any given cell of tabulated data is less than or equal to 10;
- I will not release nor permit others to release the data sets or any part of them to any person who is not a member of the organization (specified below), except with the approval of AHRQ;
- I will not attempt to link nor permit others to attempt to link the hospital stay records of persons in this data set with personally identifiable records from any other source;

Data Use Agreement for HCUP State Inpatient Databases (continued)

- I will not attempt to use nor permit others to use the datasets to learn the identity of any person included in any set;
- I will not use nor permit others to use the data concerning individual establishments (1) for commercial or competitive purposes involving those individual establishments, (2) to determine the rights, benefits, or privileges of individual establishments nor (3) to report, through any medium, data that could identify, directly or by inference, individual establishments;
- When the identities of establishments are not provided on the data sets, I will not attempt to use nor permit others to use the data sets to learn the identity of any establishment in the data sets;
- I will not contact nor permit others to contact establishments or persons in the data sets to question, verify, or discuss data in the HCUP databases;
- I will indemnify, defend, and hold harmless the data sources and AHRQ from any or all claims and losses accruing to any person, organization, or other legal entity as a result of violation of this agreement. This provision applies only to the extent permitted by federal law and regulation (i.e., to the extent permitted by 31 United States Code Section 1341 (Subtitle II, Chapter 13, Subchapter III, "Limitations on Expending and Obligating Amounts."));
- I will make no statement nor permit others to make statements indicating or suggesting that interpretations drawn are those of data sources or AHRQ;
- I will provide an abstract and reference for any published research material resulting from the use of these HCUP State Inpatient Databases to the HCUP Central Distributor; and
- I will acknowledge in all reports based on these data that the source of the data is the specific state(s) or data organization(s) that submitted data to the HCUP (e.g., "Healthcare Cost and Utilization Project (HCUP), state name(s) State Inpatient Databases (SID), Agency for Healthcare Research and Quality").

I understand that these assurances are collected for the United States Agency for Healthcare Research and Quality to require compliance with its statutory confidentiality requirement. My signature indicates my agreement to comply with the above-stated requirements with the knowledge that any violation of this statute is subject to a civil penalty of up to \$10,000 under 42 U.S.C. 299c-3(d), and that deliberately making a false statement about this or any matter within the jurisdiction of any department or agency of the Federal Government violates 18 U.S.C. 1001 and is punishable by a fine of up to \$10,000 or up to five years in prison. Violators of this agreement may also be subject to penalties under state confidentiality statutes that apply to these data for particular states.

Signed: _____ Date: _____

Print or Type Name: _____

Title: _____

Organization: _____

Address: _____

City: _____ State: _____ Zip Code: _____

Phone Number: _____ Fax: _____ E-mail: _____

Note to Purchaser: Shipment of the data product will only be made to the person who signs this data use agreement.