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 STATE AMBULATORY SURGERY AND SERVICES DATABASES (SASD)

These pages provide only an introduction to the SASD package.

Full documentation is provided online at the HCUP User Support website:

www.hcup-us.ahrq.gov

Issued June 15, 2022

Agency for Healthcare Research and Quality
Healthcare Cost and Utilization Project (HCUP)
5600 Fishers Lane
Mail Stop 7W25B
Rockville, MD 20857

SASD Data and Documentation Distributed through the HCUP Central Distributor

Website: www.hcup-us.ahrq.gov
Phone: (866) 290-4287 (toll free)
Fax: (805) 979-3787
Email: HCUP@AHRQ.gov

Table of Contents

HCUP STATE AMBULATORY SURGERY AND SERVICES DATABASES (SASD)
SUMMARY OF DATA USE LIMITATIONS

******* REMINDER *******

All users of the SASD must take the online Data Use Agreement (DUA) training course, and read and sign a Data Use Agreement. Details and links may be found on the following page.

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

- Will not use the data for any purpose other than research, analysis, and aggregate statistical reporting.
- Will not re-release any data to unauthorized users.
- Will not redistribute HCUP data by posting on any website or publishing in any other publicly accessible online repository. If a journal or publication requests access to data or analytic files, I will cite restrictions on data sharing in the Data Use Agreement and direct them to AHRQ HCUP (www.hcup-us.ahrq.gov) for more information on accessing HCUP data.
- 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 report any statistics where the number of observations (i.e., individual discharge records) in any given cell of tabulated data is less than or equal to 10 (≤10).
- 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 affecting establishments, or to determine rights, benefits, or privileges of individual establishments.
- Will not use the data for criminal and civil litigation, including expert witness testimony or for law enforcement activities.
- 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.[b]

Any violation of the limitations in the data use agreement is punishable under Federal law by a fine, up to five years in prison, or both. Violations may also be subject to penalties under State statutes.

[a] This is a summary of key terms of the Data Use Agreement for HCUP State Databases, please refer to the DUA for full terms and conditions.
[b] Suggested citations for the HCUP databases are provided in the Requirements for Publishing with HCUP Data available at www.hcup-us.ahrq.gov/db/publishing.jsp.

HCUP DATA USE AGREEMENT REQUIREMENTS

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.

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

The **HCUP Data Use Agreement for the State Databases** is available on the HCUP User Support (HCUP-US) website.

- PDF file, 251 KB: www.hcup-us.ahrq.gov/team/StateDUA.pdf

HCUP CONTACT INFORMATION

HCUP Central Distributor and HCUP User Support

Information about the content of the HCUP databases is available on the HCUP User Support (HCUP-US) website (www.hcup-us.ahrq.gov).

If you have questions, please review the HCUP Frequently Asked Questions located at www.hcup-us.ahrq.gov/tech_assist/faq.jsp.

If you need further technical assistance, please contact the HCUP Central Distributor and User Support team at:

Phone: 866-290-HCUP (4287) (toll free)
Email: HCUP@AHRQ.gov
Fax: 805 979-3787

Mailing address:
HCUP Central Distributor
c/o IBM
5425 Hollister Avenue, Suite 140
Santa Barbara, CA 93111

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

Please send user feedback to 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 your interest in the HCUP State Ambulatory Surgery and Services Databases (SASD)

HCUP State Ambulatory Surgery and Services Databases (SASD)

ABSTRACT

The State Ambulatory Surgery and Services Databases (SASD) are part of the Healthcare Cost and Utilization Project (HCUP), sponsored by the Agency for Healthcare Research and Quality (AHRQ).

The HCUP State Ambulatory Surgery and Services Databases (SASD) are a powerful set of databases that include encounter-level data for ambulatory surgeries and may also include various types of outpatient services such as observation stays, lithotripsy, radiation therapy, imaging, chemotherapy, and labor and delivery. The specific types of ambulatory surgery and outpatient services included in each SASD vary by State and data year.

- The SASD include encounter-level data for ambulatory surgery and other outpatient services from hospital-owned facilities in participating States that are translated into a uniform format to facilitate multistate comparisons and analyses.
- All SASD include data from hospital-owned ambulatory surgery facilities. In addition, some States include data from facilities not owned by a hospital.
- The SASD contain a core set of clinical and nonclinical information on all visits, regardless of the expected payer, including but not limited to Medicare, Medicaid, private insurance, self-pay, or those billed as ‘no charge’.
- In addition to the core set of uniform data elements common to all SASD, some include other elements such as the patient's race.

Researchers and policymakers use the SASD to compare inpatient surgery data with ambulatory surgery data; identify State-specific trends in ambulatory surgery utilization, access, charges, and outcomes; and conduct market-area research and small-area variation analyses.

The individual State databases are in the same HCUP uniform format and represent 100 percent 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 1997–2021 SASD.

The SASD 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 States that do not allow the release of hospital identifiers.

Twenty-six of the data organizations participating in HCUP have agreed to release their SASD files through the HCUP Central Distributor under the auspices of AHRQ. Uses are limited to research and aggregate statistical reporting.

INTRODUCTION TO THE HCUP STATE AMBULATORY SURGERY AND SERVICES DATABASES (SASD)

OVERVIEW OF THE SASD

The Healthcare Cost and Utilization Project (HCUP) State Ambulatory Surgery and Services Databases (SASD) consist of individual data files from 36 participating data organizations. The SASD include encounter-level data for ambulatory surgeries and may also include various types of outpatient services such as observation stays, lithotripsy, radiation therapy, imaging, chemotherapy, and labor and delivery. The specific types of ambulatory surgery and outpatient services included in each SASD vary by State and data year. All SASD include data from hospital-owned ambulatory surgery facilities. In addition, some States include data from facilities not owned by a hospital. The designation of a facility as hospital-owned is specific to its financial relationship with a hospital that provides inpatient care and is not related to its physical location. Hospital-owned ambulatory surgery and other outpatient care facilities may be contained within the hospital, physically attached to the hospital, or located in a different geographic area.

The SASD are 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 SASD include State-specific data elements or data elements available only for a limited number of States. The uniform format of the SASD helps facilitate cross-State comparisons. In addition, the SASD are well suited for research that requires complete enumeration of hospital-based ambulatory surgery within market areas or States.

Twenty-six of the 36 participating data organizations have agreed to release their SASD files through the HCUP Central Distributor under the auspices of AHRQ. The individual State databases are in the same HCUP uniform format. In general, they represent 100 percent 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 1997–2021 SASD.

SASD data sets are currently available for multiple States and years. Each release of the SASD includes:

- Data in American Standard Code for Information Interchange (ASCII) format on a compact disc with read-only memory (CD-ROM).
- Encounter-level data for ambulatory surgery and other outpatient services from hospital-owned facilities in participating States. In addition, some States provide ambulatory surgery and outpatient services from nonhospital-owned facilities.
- American Hospital Association (AHA) Linkage File to link the SASD to data from the AHA Annual Survey of Hospitals. This is only available for States that allow the release of hospital identifiers.

The SASD are calendar year files for all data years except 2015. Because of the transition to ICD-10-CM/PCS[1] on October 1, 2015, the 2015 SASD are split into two parts. Nine months of the 2015 data with ICD-9-CM[2] codes (discharges from Jan 1, 2015 – September 30, 2015) are in one set of files labeled Q1Q3. Three months of 2015 data with ICD-10-CM/PCS codes (discharges from October 1, 2015 – December 31, 2015) are in a separate set of files labeled Q4. More information about the changes to the HCUP databases for ICD-10-CM/PCS and use of data across the two coding system may be found on the HCUP User Support website under ICD-10-CM/PCS Resources (www.hcup-us.ahrq.gov/datainnovations/icd10_resources.jsp).

[1] ICD-10-CM/PCS: International Classification of Diseases, 10th Edition, Clinical Modification/ Procedure Coding System

[2] ICD-9-CM: International Classification of Diseases, Ninth Edition, Clinical Modification

SASD documentation and tools—including file specifications, programming source code for loading ASCII data into SAS (SAS Institute Inc.; Cary, NC), SPSS (IBM Corp.; Somers, NY), and Stata (StataCorp; College Station, TX), and value labels—are available online at the HCUP User Support website (www.hcup-us.ahrq.gov).

Starting with the 2006 SASD, the AHA Linkage files are available via the HCUP User Support website (www.hcup-us.ahrq.gov). The AHA Linkage files may not be available when the encounter-level database is released.

How the HCUP SASD Differ from State Data Files

The SASD 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
- 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 SASD are a subset of the data collected by the corresponding data organizations. HCUP uniform coding is used on most data elements on the SASD. A few State-specific data elements retain the original values provided by the respective data organizations.

What Types of Facilities Are Included in the SASD?

All SASD include data from hospital-owned ambulatory surgery facilities. In addition, some States include data from facilities not owned by a hospital. The designation of a facility as hospital-owned is specific to its financial relationship with a hospital that provides inpatient care and is not related to its physical location. Hospital-owned ambulatory surgery and other outpatient care facilities may be contained within the hospital, physically attached to the hospital, or located in a different geographic area. The designation as hospital-owned means that HCUP can identify that the hospital is billing for this service. Table 1 lists the types of facilities by State.

Table 1. Types of Facilities in the SASD

State	Hospital-Owned Ambulatory Surgery Facilities	Nonhospital-Owned Ambulatory Surgery Facilities
California	Yes	Yes
Colorado	Yes	No
District of Columbia	Yes	No
Florida	Yes	Yes
Georgia	Yes	No
Hawaii	Yes	No
Indiana	Yes	No
Iowa	Yes	No
Kansas	Yes	No
Kentucky	Yes	Yes
Maine	Yes	No
Maryland	Yes	No
Michigan	Yes	Yes
Minnesota	Yes	No
Missouri	Yes	Yes
Nebraska	Yes	No
Nevada	Yes	Yes
New Jersey	Yes	No
New York	Yes	Yes
North Carolina	Yes	Yes
Oregon	Ýes	Yes
South Carolina	Yes	Yes
South Dakota	Yes	No
Utah	Yes	Yes
Vermont	Yes	No
Wisconsin	Yes	Yes

The HCUP data element HOSPITAL-OWNED (starting in data year 2018) and FREESTANDING (from data year 1998–2017) can be used to identify hospital-owned facilities with ambulatory surgery and possibly other outpatient care data in the SASD. A facility is considered hospital-owned (HOSPITAL-OWNED =1 or FREESTANDING = 0) if any one of the following is true:

- The facility is listed in the American Hospital Association (AHA) Annual Survey Database.
- The facility is not listed in the AHA Survey, but the facility provides inpatient discharge data to HCUP.
- Documentation provided by the data source clearly indicates that the facility is hospital-owned.

If the facility in the SASD does not meet any of the above criteria, it is marked as not being owned by a hospital (HOSPITAL-OWNED =0 or FREESTANDING = 1). Because not all hospitals report to the AHA, there is a possibility that some facilities marked with HOSPITAL-OWNED =0 or FREESTANDING=1 are hospital-owned.

What is the File Structure of the SASD in the 2019–2021 Files?

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

- Core file
- Charges file
- AHA Linkage file
- Diagnosis and Procedure Groups file

The **Core file** is an encounter-level file that contains:

- Core data elements that form the nucleus of the SASD
- 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
- Are needed for traditional 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
- Are not needed for traditional applications

The Core file is an encounter-level file with one observation per discharge abstract.

The **Charges file** contains detailed charge information. There are three kinds of Charges files:

1) *Line item detail* in which a submitted charge pertains to a specified revenue center, and there may be multiple charges reported for the same revenue center. This type of Charges file includes multiple records per discharge abstract. Each record includes the following information for <u>one</u> service:
    a. Revenue center (REVCODE)
    b. Charge (CHARGE)
    c. Unit of service (UNITS)
    d. Current Procedural Terminology (CPT®) and Healthcare Common Procedure Coding System (HCPCS) codes (CPTHCPCS)
    e. Day of service (SERVDAY) for some files

    For example, if a patient had five laboratory tests, there are five records in the Charges file with information on the charge for each laboratory test. Information from this type of Charges file may be combined with the Core file by the unique record identifier (KEY), but there is not a one-to-one correspondence of records.

2) *Summarized detail* in which charge information is summed within the revenue center. This type of Charges file includes one record per discharge abstract. Each record contains three corresponding arrays with the following information:
    a. Revenue center (REVCDn)
    b. Total charge for the revenue center (CHGn)
    c. Total units of service for the revenue center (UNITn)

    For example, if a patient had five laboratory tests, REVCD1 would include the revenue code for laboratory, CHG1 would include the total charge for the five tests, and UNIT1 would be five. To combine data elements between this type of Charges file and the Core file, merge

the files by the unique record identifier (KEY). There will be a one-to-one correspondence of records.

3) *Collapsed detail* in which charge information is summed across revenue centers. This type of Charges file includes one record per discharge abstract. Each record contains an array of collapsed charges (CHGn).

   Consider the example of a patient that had five laboratory tests from different revenue centers in the range of 300 to 319. CHG1, which was predefined as Laboratory Charges for revenue centers 300–319, would include the total charge for the five tests; however, there is no detail on which specific revenue centers were used. To combine data elements between this type of Charges file and the Core file, merge the files by the unique record identifier (KEY). There will be a one-to-one correspondence of records.

Refer to the Description of Data Elements online at the HCUP User Support website (www.hcup-us.ahrq.gov) for more information on the charge information from the different States.

The **AHA Linkage file** contains AHA linkage data elements that allow the SASD 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. Because the data organizations in participating States determine whether the AHA linkage data elements may be released through the HCUP Central Distributor with the SASD, not all SASD include AHA linkage data elements.

Starting with the 2006 SASD, the AHA Linkage files are available via the HCUP User Support website (www.hcup-us.ahrq.gov). The AHA Linkage files may not be available when the encounter-level database is released.

The AHA Linkage file is a hospital-level file with one observation per hospital or facility. To combine encounter-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."

Diagnosis and Procedure Groups File is an encounter-level file that contains data elements from AHRQ software tools. They are designed to facilitate the use of the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM/PCS) diagnostic and procedure information in the HCUP databases. The unit of observation is an ambulatory surgery encounter. The HCUP unique record identifier (KEY) provides the linkage between the Core files and the Diagnosis and Procedure Groups files. These files are available beginning with the 2005 SASD.

What is the File Structure of the SASD in the 2016–2018 Files?

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

- Core file
- Charges file
- AHA Linkage file

Unavailable with the 2016–2018 SASD is the Diagnosis and Procedure Groups file that had been included with the SASD in prior data years. The data elements included in that file were derived from AHRQ software tools. If you are interested in applying the AHRQ software tools to the ICD-10-CM/PCS data in the 2016–2018 SASD, beta versions of the AHRQ software tools are available on the HCUP User Support website at www.hcup-us.ahrq.gov//tools_software.jsp. Also available is a tutorial on how to apply the AHRQ software tools to the HCUP databases at www.hcup-us.ahrq.gov/tech_assist/tutorials.jsp.

The **Core file** is an encounter-level file that contains:

- Core data elements that form the nucleus of the SASD
- 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
- Are needed for traditional 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
- Are not needed for traditional applications

The Core file is an encounter-level file with one observation per discharge abstract.

The **Charges file** contains detailed charge information. There are three kinds of Charges files:

1) *Line item detail* in which a submitted charge pertains to a specified revenue center, and there may be multiple charges reported for the same revenue center. This type of Charges file includes multiple records per discharge abstract. Each record includes the following information for <u>one</u> service:
    a. Revenue center (REVCODE)
    b. Charge (CHARGE)
    c. Unit of service (UNITS)
    d. Current Procedural Terminology (CPT®) and Healthcare Common Procedure Coding System (HCPCS) codes (CPTHCPCS)
    e. Day of service (SERVDAY) for some files

    For example, if a patient had five laboratory tests, there are five records in the Charges file with information on the charge for each laboratory test. Information from this type of Charges file may be combined with the Core file by the unique record identifier (KEY), but there is not a one-to-one correspondence of records.

2) *Summarized detail* in which charge information is summed within the revenue center. This type of Charges file includes one record per discharge abstract. Each record contains three corresponding arrays with the following information:
    a. Revenue center (REVCDn)
    b. Total charge for the revenue center (CHGn)
    c. Total units of service for the revenue center (UNITn)

    For example, if a patient had five laboratory tests, REVCD1 would include the revenue code for laboratory, CHG1 would include the total charge for the five tests, and UNIT1 would be five. To combine data elements between this type of Charges file and the Core file, merge the files by the unique record identifier (KEY). There will be a one-to-one correspondence of records.

3) *Collapsed detail* in which charge information is summed across revenue centers. This type of Charges file includes one record per discharge abstract. Each record contains an array of collapsed charges (CHGn).

    Consider the example of a patient that had five laboratory tests from different revenue centers in the range of 300 to 319. CHG1, which was predefined as Laboratory Charges for revenue centers 300–319, would include the total charge for the five tests; however, there is no detail on which specific revenue centers were used. To combine data elements between this type of Charges file and the Core file, merge the files by the unique record identifier (KEY). There will be a one-to-one correspondence of records.

Refer to the Description of Data Elements online at the HCUP User Support website (www.hcup-us.ahrq.gov) for more information on the charge information from the different States.

The **AHA Linkage file** contains AHA linkage data elements that allow the SASD 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. Because the data organizations in participating States determine whether the AHA linkage data elements may be released through the HCUP Central Distributor with the SASD, not all SASD include AHA linkage data elements.

Starting with the 2006 SASD, the AHA Linkage files are available via the HCUP User Support website (www.hcup-us.ahrq.gov). The AHA Linkage files may not be available when the encounter-level database is released.

The AHA Linkage file is a hospital-level file with one observation per hospital or facility. To combine encounter-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 ambulatory surgery records for DSHOSPID "A," but the Hospital file contains only 1 record for DSHOSPID "A."

What is the File Structure of the SASD in the 2015 Files?

The file structure of the 2015 SASD is similar to previous years (and future years) in terms of how data elements are split across multiple data files, but differs from others years because the records within the 2015 files have been separated into two sets of files based on the discharge date because of the transition from reporting medical diagnoses and inpatient procedures using ICD-9-CM to the ICD-10-CM/PCS code sets.[3]

The 2015 SASD are split into two separate sets of files based on the discharge date and different coding schemes:

- If the record was discharged between January 1, 2015, and September 30, 2015, it is retained in the quarter 1 to quarter 3 (Q1–Q3) files and includes ICD-9-CM data
- If the record was discharged between October 1, 2015 and December 31, 2015, it is retained in the quarter 4 (Q4) files and includes ICD-10-CM/PCS data.

Almost all of the diagnosis and procedure-related data elements that are based on ICD-10-CM/PCS data have been renamed with the prefix of I10 to distinguish them from the ICD-9-CM-based data element.

Based on the availability of data elements across States, data elements included in the 2015 SASD are structured as follows:

- Core files, one for Q1–Q3 and one for Q4
- Charges files, one for Q1–Q3 and one for Q4
- AHA Linkage file (one file because this is a hospital-level file, instead of an encounter-level file)
- Diagnosis and Procedure Groups files, one for Q1–Q3 and one for Q4

The **Core file** is an encounter-level file that contains:

- Core data elements that form the nucleus of the SASD
- 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

[3] ICD-9-CM: International Classification of Diseases, Ninth Edition, Clinical Modification; ICD-10-CM/PCS: International Classification of Diseases, 10th Edition, Clinical Modification/ Procedure Coding System

- Lend themselves to uniform coding across sources
- Are needed for traditional 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
- Are not needed for traditional applications

The **Charges file** contains detailed charge information. There are three kinds of Charges files:

1) *Line item detail* in which a submitted charge pertains to a specified revenue center and there may be multiple charges reported for the same revenue center. This type of Charges file includes multiple records per discharge abstract. Each record includes the following information for one service:
    a. Revenue center (REVCODE)
    b. Charge (CHARGE)
    c. Unit of service (UNITS)
    d. Current Procedural Terminology (CPT®) and Healthcare Common Procedure Coding System (HCPCS) codes (CPTHCPCS)
    e. Day of service (SERVDAY) for some files

    For example, if a patient had five laboratory tests, there are five records in the Charges file with information on the charge for each laboratory test. Information from this type of Charges file may be combined with the Core file by the unique record identifier (KEY), but there is not a one-to-one correspondence of records.

2) *Summarized detail* in which charge information is summed within the revenue center. This type of Charges file includes one record per discharge abstract. Each record contains three corresponding arrays with the following information:
    a. Revenue center (REVCDn)
    b. Total charge for the revenue center (CHGn)
    c. Total units of service for the revenue center (UNITn)

    For example, if a patient had five laboratory tests, REVCD1 would include the revenue code for laboratory, CHG1 would include the total charge for the five tests, and UNIT1 would be five. To combine data elements between this type of Charges file and the Core file, merge the files by the unique record identifier (KEY). There will be a one-to-one correspondence of records.

3) *Collapsed detail* in which charge information is summed across revenue centers. This type of Charges file includes one record per discharge abstract. Each record contains an array of collapsed charges (CHGn).

    Consider the example of a patient that had five laboratory tests from different revenue centers in the range of 300 to 319. CHG1, which was predefined as Laboratory Charges for revenue centers 300–319, would include the total charge for the five tests, but there is no detail on which specific revenue centers were used. To combine data elements between this type of Charges file and the Core file, merge the files by the unique record identifier (KEY). There will be a one-to-one correspondence of records.

Refer to the Description of Data Elements online at the HCUP User Support website (www.hcup-us.ahrq.gov) for more information on the charge information from the different States.

The **AHA Linkage file** contains AHA linkage data elements that allow the SASD 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. Because the data organizations in participating States determine whether the AHA linkage data elements may be released through the HCUP Central Distributor with the SASD, not all SASD include AHA linkage data elements.

Starting with the 2006 SASD, the AHA Linkage files are available via the HCUP User Support website (www.hcup-us.ahrq.gov). The AHA Linkage files may not be available when the encounter-level database is released.

The AHA Linkage file is a hospital-level file with one observation per hospital or facility. To combine encounter-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."

Diagnosis and Procedure Groups File is an encounter-level file that contains data elements from AHRQ software tools. They are designed to facilitate the use of the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) diagnostic and procedure information in the HCUP databases. The unit of observation is an *outpatient record*. The HCUP unique record identifier (KEY) provides the linkage between the Core files and the Diagnosis and Procedure Groups files. These files are available beginning with the 2005 SASD.

What is the File Structure of the SASD in the 2005–2014 Files?

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

- Core file
- Charges file
- AHA Linkage file
- Diagnosis and Procedure Groups file

The **Core file** is an encounter-level file that contains:

- Core data elements that form the nucleus of the SASD
- 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
- Are needed for traditional 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
- Are not needed for traditional applications

The Core file is an encounter-level file with one observation per discharge abstract.

The **Charges file** contains detailed charge information. There are three kinds of Charges files:

1) *Line item detail* in which a submitted charge pertains to a specified revenue center, and there may be multiple charges reported for the same revenue center. This type of Charges file includes multiple records per discharge abstract. Each record includes the following information for one service:
    a. Revenue center (REVCODE)
    b. Charge (CHARGE)
    c. Unit of service (UNITS)
    d. Current Procedural Terminology (CPT®) and Healthcare Common Procedure Coding

System (HCPCS) codes (CPTHCPCS)
e. Day of service (SERVDAY) for some files

For example, if a patient had five laboratory tests, there are five records in the Charges file with information on the charge for each laboratory test. Information from this type of Charges file may be combined with the Core file by the unique record identifier (KEY), but there is not a one-to-one correspondence of records.

2) *Summarized detail* in which charge information is summed within the revenue center. This type of Charges file includes one record per discharge abstract. Each record contains three corresponding arrays with the following information:
   a. Revenue center (REVCDn)
   b. Total charge for the revenue center (CHGn)
   c. Total units of service for the revenue center (UNITn)

   For example, if a patient had five laboratory tests, REVCD1 would include the revenue code for laboratory, CHG1 would include the total charge for the five tests, and UNIT1 would be five. To combine data elements between this type of Charges file and the Core file, merge the files by the unique record identifier (KEY). There will be a one-to-one correspondence of records.

3) *Collapsed detail* in which charge information is summed across revenue centers. This type of Charges file includes one record per discharge abstract. Each record contains an array of collapsed charges (CHGn).

   Consider the example of a patient that had five laboratory tests from different revenue centers in the range of 300 to 319. CHG1, which was predefined as Laboratory Charges for revenue centers 300–319, would include the total charge for the five tests; however, there is no detail on which specific revenue centers were used. To combine data elements between this type of Charges file and the Core file, merge the files by the unique record identifier (KEY). There will be a one-to-one correspondence of records.

Refer to the Description of Data Elements online at the HCUP User Support website (www.hcup-us.ahrq.gov) for more information on the charge information from the different States.

The **AHA Linkage file** contains AHA linkage data elements that allow the SASD 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. Because the data organizations in participating States determine whether the AHA linkage data elements may be released through the HCUP Central Distributor with the SASD, not all SASD include AHA linkage data elements.

Starting with the 2006 SASD, the AHA Linkage files are available via the HCUP User Support website (www.hcup-us.ahrq.gov). The AHA Linkage files may not be available when the encounter-level database is released.

The AHA Linkage file is a hospital-level file with one observation per hospital or facility. To combine encounter-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 ambulatory surgery records for DSHOSPID "A," but the Hospital file contains only 1 record for DSHOSPID "A."

Diagnosis and Procedure Groups File is an encounter-level file that contains data elements from AHRQ software tools. They are designed to facilitate the use of the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) diagnostic and procedure information in the HCUP databases. The unit of observation is an *outpatient record*. The HCUP unique record identifier (KEY) provides the linkage between the Core files and the Diagnosis and Procedure Groups files. These files are available beginning with the 2005 SASD.

What is the File Structure of the SASD in the 1998–2004 Files?

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

- Core file
- Charges file
- AHA Linkage file

The **Core file** is an encounter-level file that contains:

- Core data elements that form the nucleus of the SASD
- 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
- Are needed for traditional 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
- Are not needed for traditional applications

The Core file is an encounter-level file with one observation per discharge abstract.

The **Charges file** contains detailed charge information. There are three kinds of Charges files:

1) *Line item detail* in which a submitted charge pertains to a specified revenue center. and there may be multiple charges reported for the same revenue center. This type of Charges file includes multiple records per discharge abstract. Each record includes the following information for <u>one</u> service:
    a. Revenue center (REVCODE)
    b. Charge (CHARGE)
    c. Unit of service (UNITS)
    d. CPT/HCPCS codes (CPTHCPCS)
    e. Day of service (SERVDAY) for some files

    For example, if a patient had five laboratory tests, there are five records in the Charges file with information on the charge for each laboratory test. Information from this type of Charges file may be combined with the Core file by the unique record identifier (KEY), but there is not a one-to-one correspondence of records.

2) *Summarized detail* in which charge information is summed within the revenue center. This type of Charges file includes one record per discharge abstract. Each record contains three corresponding arrays with the following information:
    a. Revenue center (REVCDn)
    b. Total charge for the revenue center (CHGn)
    c. Total units of service for the revenue center (UNITn)

    For example, if a patient had five laboratory tests, REVCD1 would include the revenue code for laboratory, CHG1 would include the total charge for the five tests, and UNIT1 would be five. To combine data elements between this type of Charges file and the Core file, merge the files by the unique record identifier (KEY). There will be a one-to-one correspondence of records.

3) *Collapsed detail* in which charge information is summed across revenue centers. This type of Charges file includes one record per discharge abstract. Each record contains an array of collapsed charges (CHGn).
Consider the example of a patient that had five laboratory tests from different revenue centers in the range of 300 to 319. CHG1, which was predefined as Laboratory Charges for revenue centers 300–319, would include the total charge for the five tests; however, there is no detail on which specific revenue centers were used. To combine data elements between this type of Charges file and the Core file, merge the files by the unique record identifier (KEY). There will be a one-to-one correspondence of records.

Refer to the Description of Data Elements online at the HCUP User Support website (www.hcup-us.ahrq.gov) for more information on the charge information from the different States.

The **AHA Linkage file** contains AHA linkage data elements that allow the SASD 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. Because the data organizations in participating States determine whether the AHA linkage data elements may be released through the HCUP Central Distributor with the SASD, not all SASD include AHA linkage data elements.

The AHA Linkage file is a hospital-level file with one observation per hospital or facility. To combine encounter-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 ambulatory surgery records for DSHOSPID "A," but the Hospital file contains only 1 record for DSHOSPID "A."

What is the File Structure of the SASD in the 1997 Files?

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

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

The **Core file** contains core data elements that form the nucleus of the SASD. 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
- Are needed for traditional 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
- Are not needed for traditional applications

The **AHA Linkage file** contains AHA linkage data elements that allow the SASD 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. Because the data organizations in participating States determine whether the AHA linkage data elements may be released through the HCUP Central Distributor with the SASD, not all SASD include AHA linkage data elements.

The Core and State-specific files are encounter-level files with one observation per abstract. The same record is represented in each file, but each contains different data elements. To combine data elements across encounter-level files, merge the files by the unique record identifier (SEQ_ASD). 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 encounter-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 ambulatory surgery records for DSHOSPID "A," but the AHA Linkage file contains only 1 record for DSHOSPID "A."

GETTING STARTED

SASD Data Files

SASD Data Files are provided on CD-ROMs. The number of CD-ROMs depends on the State and year of data.

To load SASD data onto your PC, you will need between one and four gigabytes of space available, depending on which SASD 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, follow these steps:

1. Create a directory for the State-specific SASD on your hard drive.
2. Copy the self-extracting data files from the SASD 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.

SASD Programs, Documentation, and Tools

The SASD programs, technical documentation files, and HCUP tools are available online via the Databases page at the HCUP User Support website (www.hcup-us.ahrq.gov/databases.jsp). The site provides important resources for SASD users, and all of the files may be downloaded free of charge. A summary is provided in Table 2.

The SASD programs include SAS, SPSS, and Stata load programs containing the programming code necessary to convert SASD ASCII files into SAS, SPSS, or Stata. Please note that for the 2015 SASD, there will be one set of load programs for the Q1–Q3 files and another set of load programs for the Q4 files.

The SASD technical documentation provides detailed descriptions of the structure and content of the SASD.

The HCUP tools include the Clinical Classifications Software (CCS) and general label and format information that are applicable to all HCUP databases.

Information intended to summarize key issues to be anticipated by researchers before analyzing health services outcomes in the HCUP databases that include ICD-10-CM/PCS coding is included on the HCUP User Support website (www.hcup-us.ahrq.gov/datainnovations/icd10_resources.jsp). The section discusses key differences in the structure of HCUP databases, presents preliminary coding differences that were observed in HCUP databases, and provides general guidance and forewarning to users interested in analyzing outcomes that are potentially impacted by the transition.

Table 2. SASD Database Documentation Available on HCUP-US

Description of SASD Database

- SASD Overview
- Introduction to the SASD (*this document*)
- SASD File Compositions—describes types of hospitals and types of records included in each SASD (e.g., number of visits by year)
- SASD-Related Reports

Restrictions on Use

- HCUP Data Use Agreement Training
- SASD Data Use Agreement
- Requirements for Publishing with HCUP Data

File Specifications and Load Programs

- File Specifications—details data file names, number of records, record length, and record layout (e.g., file size by year)
- SAS Load Programs
- SPSS Load Programs
- Stata Load Programs

Data Elements

- Availability of States Across All Years
- Availability of Data Elements by Year
- Availability of HCUP Revisit Variables across States and Years
- Summary Statistics—lists means and frequencies on nearly all data elements

Additional Resources for Data Elements

- HCUP Quality Control Procedures—describes procedures used to assess data quality
- HCUP Coding Practices—describes how HCUP data elements are coded
- HCUP Hospital Identifiers—explains data elements that characterize individual hospitals

ICD-10-CM/PCS Included in the SASD Starting With 2015

- 2016 State Databases Revised File Structure and New Data Elements
- Caution: 2015 SASD Includes ICD-9-CM and ICD-10-CM/PCS Data
  - 2015 State Databases Revised File Structure and New Data Elements
- Additional ICD-10-CM/PCS Resources
- Tutorial for Loading HCUP Software Tools for ICD-10-CM/PCS

Known Data Issues

- Includes State-specific information on databases that have been updated or have known data issues

HCUP Tools: Labels and Formats

- HCUP Formats Program—Creates SAS formats to label the values of selected categorical data elements in HCUP files
- HCUP Diagnosis and Procedure Groups Formats Program—Creates SAS formats to label the values of HCUP Diagnosis and Procedure Groups data elements, including Clinical Classifications Software Refined (CCSR) data elements
- ICD-9-CM Formats Program—Creates SAS formats to label the values of ICD-9-CM Diagnoses and Procedures
- ICD-10-CM Formats Program—Creates SAS formats to label the values of ICD-10-CM Diagnoses and Procedures

HCUP Supplemental Files

- American Hospital Association Linkage Files
- HCUP Variables for Revisit Analysis

Obtaining HCUP Data

- Purchase HCUP data from the HCUP Central Distributor