INTRODUCTION TO

THE HCUP STATE EMERGENCY DEPARTMENT DATABASES (SEDD)

These pages provide only an introduction to the SEDD package.

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

www.hcup-us.ahrq.gov.

Issued June 15, 2022

SEDD Data and Documentation Distributed through the HCUP Central Distributor

Table of Contents

HCUP STATE EMERGENCY DEPARTMENT DATABASES (SEDD)
SUMMARY OF DATA USE LIMITATIONS

***** REMINDER *****

All users of the SEDD 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 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 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: 866-290-4287 (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.

**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 Emergency Department Databases (SEDD)**

HCUP State Emergency Department Databases (SEDD)

ABSTRACT

The State Emergency Department Databases (SEDD) are part of the Healthcare Cost and Utilization Project (HCUP), sponsored by the Agency for Healthcare Research and Quality (AHRQ).

The HCUP State Emergency Department Databases (SEDD) are a powerful set of databases that capture discharge information on all emergency department visits that do not result in an admission. Information on patients initially seen in the emergency room and then admitted to the hospital is included in the HCUP State Inpatient Databases (SID). To enumerate all emergency department visits, SEDD discharges should be combined with SID discharges that originate in the emergency department.

- The SEDD contain the emergency department encounter abstracts from hospital-affiliated emergency departments in participating States that are translated into a uniform format to facilitate multistate comparisons and analyses.

- The SEDD 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 SEDD, some State data include other elements such as the patient's race.

Researchers and policymakers use the SEDD to investigate access to health care in a changing health care marketplace; identify State-specific trends in emergency department 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 1999–2021 SEDD.

The SEDD can be linked to hospital-level data from the American Hospital Association (AHA) Annual Survey of Hospitals and county-level data from the Bureau of Health Professions' Area Resource File (ARF), except in States that do not allow the release of hospital identifiers.

Thirty of the data organizations participating in HCUP have agreed to release their SEDD files through the HCUP Central Distributor under the auspices of AHRQ. Uses are limited to research, analysis, and aggregate statistical reporting.

INTRODUCTION TO THE HCUP STATE EMERGENCY DEPARTMENT DATABASES (SEDD)

OVERVIEW OF THE SEDD

The Healthcare Cost and Utilization Project (HCUP) State Emergency Department Databases (SEDD) consist of individual data files from 42 participating data organizations. The SEDD contain information on emergency department (ED) visits to hospital-affiliated emergency departments that do not result in admissions. Information on patients initially seen in the emergency room and then admitted to the hospital is included in the HCUP State Inpatient Databases (SID). Composition and completeness of data files may vary from State to State. The SEDD 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 SEDD include State-specific data elements or data elements available only for a limited number of States. The uniform format of the SEDD helps facilitate cross-State comparisons. In addition, the SEDD, in combination with the SID, are well suited for research that requires complete enumeration of hospitals and emergency room encounters within market areas or States.

Thirty of the 42 participating data organizations have agreed to release their SEDD 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 1999–2021 SEDD.

SEDD data sets are currently available for multiple States and years. Each release of the SEDD includes:
- Data in American Standard Code for Information Interchange (ASCII) format on a compact disc with read-only memory CD-ROM.

- Patient-level discharge abstract data in participating States for 100 percent of discharges from hospital-affiliated emergency departments that do not result in admissions.

- American Hospital Association (AHA) Linkage File to link the SEDD to data from the AHA Annual Survey of Hospitals. This is only available for States that allow the release of hospital identifiers.

The SEDD 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 SEDD 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).

SEDD 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 SEDD, 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 discharge-level database is released.

How the HCUP SEDD Differ from State Data Files

The SEDD available through the HCUP Central Distributor differ from the data files available from the data organizations in the following ways:

---

[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

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

What Types of Facilities Are Included in the SEDD?

The types of hospitals included in the SEDD depend on the information provided by the data organizations. 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. All hospitals reported by the data organizations are retained in the SEDD files.

Tables showing the number of hospitals in the SEDD can be found online at the HCUP User Support website (www.hcup-us.ahrq.gov). The tables present the hospitals by the number of:

- Community hospitals
- Hospitals not classified as community
- Community hospitals not in the SEDD

The only reason a hospital is not included in the SEDD is if the data organization did not supply the necessary data.

Information contained in the AHA Annual Survey of Hospitals was used to determine if a hospital was a community hospital. Community hospitals, as defined by AHA, include "all nonfederal, short-term, general and other specialty hospitals, excluding hospital units of institutions." Included among community hospitals are academic medical centers and specialty hospitals such as obstetrics, gynecology, otolaryngology, short-term rehabilitation, orthopedic, and pediatric hospitals. Noncommunity hospitals include Federal hospitals (e.g., Veterans Affairs, Department of Defense, and Indian Health Service hospitals), long-term hospitals, psychiatric hospitals, alcohol/chemical dependency treatment facilities, and hospital units within institutions such as prisons.

Some hospitals could not be categorized as community or noncommunity hospitals because they could not be matched with AHA information. This occurs when a hospital was closed in a previous year or when the hospital does not report to the AHA.

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

Based on the availability of data elements across States, data elements included in the SEDD 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 SEDD
- 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., patient age, sex)

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 a discharge-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) that are predefined by the data organization that provided the data.

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 SEDD 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 SEDD, not all SEDD include AHA linkage data elements.

Starting with the 2006 SEDD, 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 discharge-level database is released.

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 one 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, Tenth Revision, Clinical Modification (ICD-10-CM/PCS) diagnostic and procedure information in the HCUP databases. The unit of observation is an emergency department visit. 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 SEDD.

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

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

- Core file
- Charges file
- AHA Linkage file

Unavailable with the 2016-2018 SEDD is the Diagnosis and Procedure Groups file that had been included with the SEDD 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 SEDD, 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 SEDD
- 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., patient age, sex)

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 a discharge-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) that are predefined by the data organization that provided the data.

   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 SEDD 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 SEDD, not all SEDD include AHA linkage data elements.

Starting with the 2006 SEDD, 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 discharge-level database is released.

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 one record for DSHOSPID "A."

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

The file structure of the 2015 SEDD 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 SEDD 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 SEDD 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 SEDD
- State-specific data elements intended for limited use

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

---

[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

- Are available from all or nearly all data sources
- Lend themselves to uniform coding across sources
- Are needed for traditional applications (e.g., patient age, sex)

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 <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) that are predefined by the data organization that provided the data.

   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 SEDD 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 SEDD, not all SEDD include AHA linkage data elements.

Starting with the 2006 SEDD, 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 discharge-level database is released.

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

Diagnosis and Procedure Groups File is an encounter-level file that contains data elements from AHRQ software tools 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 *inpatient stay 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 SEDD.

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

Based on the availability of data elements across States, data elements included in the SEDD 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 SEDD
- 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., patient age, sex)

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 a discharge-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) that are predefined by the data organization that provided the data.

   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 SEDD 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 SEDD, not all SEDD include AHA linkage data elements.

Starting with the 2006 SEDD, 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 discharge-level database is released.

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 one 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 *inpatient stay 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 SEDD.

What is the File Structure of the SEDD in the 1999–2004 Files?

Based on the availability of data elements across States, data elements included in the SEDD 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 SEDD
- 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., patient age, sex).

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 a discharge-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) that are predefined by the data organization that provided the data.

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 SEDD 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 SEDD, not all SEDD include AHA linkage data elements.

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 one record for DSHOSPID "A."

GETTING STARTED

SEDD Data Files

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

To load SEDD data onto your PC, you will need between one and four gigabytes of space available, depending on which SEDD database you are using. Because of the size of the files, the data are distributed as self-extracting PKZIP compressed files. To uncompress the data, follow these steps:

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

SEDD Programs, Documentation, and Tools

The SEDD 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 SEDD users, and all of the files may be downloaded free of charge. A summary is provided in Table 1.

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

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

The HCUP tools include 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 1. SEDD Database Documentation Available on HCUP-US

Description of the SEDD Files

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

Restrictions on Use

- HCUP Data Use Agreement Training
- SEDD 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 for All States and Years—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 Data Included in the SEDD Starting With 2015

- 2016 State Databases Revised File Structure and New Data Elements
- Caution: 2015 SEDD 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

- Documents that describe known data issues or corrections made to data files

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