Cost-to-Charge Ratio Files:
User Guide for Central Distributor State Inpatient Database (CD-SID) CCRs

The Healthcare Cost and Utilization Project (HCUP) Cost-to-Charge Ratio Files (CCR Files) are hospital-level files that facilitate the conversion of total charges into hospital costs (expenses) for providing care, which can be linked to HCUP inpatient and emergency department databases.

This user guide describes the 2001-2018 CCR for Central Distributor State Inpatient Databases (CCR for CD-SID) Files. The user guide was updated to reflect the 2018 CCR for CD-SID version 2 file, which was revised to include two additional States (New York and Utah).

1. Overview of Methodology

   The CCR Files are constructed from appropriate cost centers in the hospital cost reports obtained from the Centers for Medicare and Medicaid Services (CMS) Healthcare Cost Report Information System (HCRIS). The HCUP CCR Files are annual datasets that provide hospital-specific cost-to-charge ratios based on all-payer inpatient costs for nearly every hospital in each year's collection of Central Distributor SID.

   The CCR for Inpatient Files are used to estimate the resource cost of inpatient care and its variation across hospitals and conditions. The files are designed to supplement the data elements in the HCUP inpatient databases which contain data on total charges for each hospital stay. *Charges* represent the amount a hospital billed for the case; c*osts* reflect the actual expenses incurred in the production of hospital services, such as wages, supplies, and utility costs. The charges (costs) do not reflect the specific amounts that hospitals receive in payment.

   HCUP utilizes the CMS fiscal year files, such as "hosp10_2018_NMRC.csv," also referred to as Prospective Payment System (PPS) records, for hospital data submitted through March 31st, approximately 18 months after the close of a fiscal year.

2. Description of CCR for CD-SID Files

   The HCUP CCR for CD-SID files provide an estimate of all-payer, inpatient cost-to-charge ratios for hospitals in states that participate in the Central Distributor SID for data years 2001-2018. The files are provided as CSV (comma-separated value) text files that use a comma to separate values on each record. Records are included for all community hospitals from the HCUP CD-SID that have "clean" matches – meaning the HCUP SID hospital has a match with both the American Hospital Association (AHA) Annual Survey Database and the CMS HCRIS file for the corresponding fiscal year. All HCUP hospitals in the CCR for CD-SID files are in the AHA Annual Survey Database.

   Separate CCR for CD-SID files are released for each data year and should be used with the corresponding year of the SID. Three states release state-specific files that are separate from the CCR for CD-SID file. For certain years, Iowa, Minnesota, and Nebraska CCR files are released as separate, state-specific files available by request from the HCUP Central Distributor to purchasers whose organizational affiliation and ownership meet the Partner's eligibility criteria. All other CD-SID states that permit release of the cost-to-charge ratio measures are included in the CCR for CD-SID file.

The CCR for CD-SID files can be linked to discharge records in the SID using the HCUP hospital identification number, HOSPID, which is a unique hospital number exclusive to the HCUP data.

- For States that release an HCUP AHA Linkage File, this is achieved in two steps, first by linking records from the CCR for CD-SID file to the HCUP AHA Linkage File by the data element HOSPID. Then by linking the resulting file to the Central Distributor SID by DSHOSPID.

- For states that do not release an HCUP AHA Linkage File, HOSPID is included directly on their Central Distributor SID file. For these states, the data elements from the Cost-to-Charge file can be merged onto the Central Distributor SID by HOSPID.

The AHA Linkage files can be downloaded from the HCUP AHA Linkage Files page on the HCUP User Support (HCUP-US) website.

The cost of inpatient care for a discharge is estimated by multiplying TOTCHG (total charges reported on the discharge record) by either the hospital-specific all-payer inpatient cost-to-charge ratio, APICC, or the group average all-payer inpatient cost-to-charge ratio, GAPICC.

Note: HOSPID on the CCR CSV (comma-separated value) text file is enclosed in quotations in order to preserve leading zeros. As a result, some software applications may interpret HOSPID as a character variable, which in turn would not match the numeric version of HOSPID on the SID. Users should load the HOSPID data element on the CCR file as numeric or convert it to numeric prior to merging with the CD-SID files.

3. Records in the CCR CD-SID Files

The datasets contain a record for each hospital (unduplicated HOSPIDs) in states that release both their Central Distributor SID and CCR measures. The CCR for CD-SID files include only the Central Distributor states that were available at the time the files were created; states that subsequently agreed to participate in the Central Distributor and release their CD-SID were not included. For certain years, the CCR Files have been updated to add States not available in earlier versions. Refer to *Appendix A, Updates to Annual CCR for CD-SID Files* for additional information about states added to the revised files.

Table 1 provides the count of records (hospitals) included in the CCR for CD-SID files each year compared with the count of all hospitals in the SID maintained by AHRQ (including SID that are not released through the HCUP Central Distributor).

Table 1: CCR for CD-SID Record Counts, By Year

Year	Number of Records (Hospitals)	
	CCR for CD-SID[a]	Total CD-SID[b]
2018 v2	2,690	3,023
2018 v1	2,464	2,797
2017	2,318	2,651
2016	2,297	2,629
2015	2,178	2,514
2014	2,078	2,415

Year	Number of Records (Hospitals)	
	CCR for CD-SID[a]	Total CD-SID[b]
2013	1,876	2,081
2012	2,313	2,517
2011	2,412	2,558
2010	2,519	2,666
2009	2,402	2,548
2008	2,367	2,514
2007	2,305	2,451
2006	2,313	2,463
2005	1,691	1,830
2004	1,613	1,771
2003	1,626	2,257
2002	1,602	2,234
2001	1,578	2,210

Notes:
[a] Excludes States that did not permit AHRQ to release their cost-to-charge measures and three States that release separate state-specific CCR files (Iowa, Minnesota, and Nebraska).
[b] This represents the count of all hospitals in the SID maintained by AHRQ (including SID that are not released through the HCUP Central Distributor).

Table 2 provides the count of records with hospital-specific and group-average cost ratios. Where permitted by HCUP Partner organizations, the CCR for CD-SID file includes a hospital-specific all-payer inpatient cost-to-charge ratio, APICC. For all hospitals, there is also a weighted group average, GAPICC. Analysts can use the APICC, when available, and otherwise use the weighted group average, GAPICC.

Table 2: Number of Records (Hospitals) in the CCR for CD-SID, by Year and Presence of APICC and GAPICC

Year	Records (Hospitals) with APICC	Percent with APICC, %	Records (Hospitals) with GAPICC only
2018 v2	1,751	65	939
2018 v1	1,555	63	909
2017	1,486	64	832
2016	1,481	64	816
2015	1,395	67	680
2014	1,409	68	669
2013	1,391	74	485
2012	1,679	73	634
2011	1,799	75	613
2010	1,971	78	548
2009	1,749	73	653
2008	1,865	78	502
2007	1,705	74	600
2006	1,654	72	659
2005	1,112	66	579
2004	1,091	68	518
2003	1,084	66	542
2002	972	61	630
2001	984	62	594

4. Participating States

Almost all states participating in the Central Distributor are included in the CCR for CD-SID file. Table 3 lists the states included in each year's file. Three states – Iowa, Minnesota, and Nebraska – release their cost-to-charge measures in separate, state-specific files for certain years. One State, South Carolina, does not release HOSPID and consequently CCR measures for its CD-SID. States added to revision files are bolded.

Table 3: States Included in the CCR for CD-SID, by Year

Year	States in CCR for CD-SID file	State-Specific Files
2018 v2	AK AZ AR CA CO DC DE FL GA KS KY MA MD ME MI MS NC NJ NM NV **NY** OR RI SD **UT** VT WA WI WV (29)	IA MN NE
2018 v1	AK AZ AR CA CO DC DE FL GA KS KY MA MD ME MI MS NC NJ NM NV OR RI SD VT WA WI WV (27)	IA MN NE
2017	AK AZ AR CO DC DE FL GA KS KY MA MD ME MI MS NC NJ NM NV NY OR RI SD UT VT WA WI WV	IA MN NE
2016	AZ AR CO DC FL GA HI KS KY MA MD ME MI MS NC NJ NM NV NY OR RI SD UT VT WA WI WV	IA MN NE
2015	AZ AR CO DC FL GA HI KY MA MD ME MI MS NC NJ NM NV NY OR RI SD UT VT WA WI WV	IA MN NE
2014	AZ AR CO DC FL GA HI KY MA MD ME MI NC NJ NM NV NY OR RI SD UT VT WA WI WV	IA MN NE
2013	AZ AR CO FL HI KY MA MD MI NC NJ NM NV NY OR RI SD UT VT WA WI WV	IA NE
2012	AZ AR CA CO FL HI KY MD MA ME MI NC NJ NM NV NY OR RI SD UT VT WA WI WV	IA NE
2011	AZ AR CA CO FL HI KY MD MA ME MI MS NC NJ NM NV NY OR RI SD UT VT WA WI WV	IA NE
2010	AZ AR CA CO FL HI IA KY MD MA ME MI MS NC NJ NM NV NY OR RI SD UT VT WA WI WV	NE
2009	AZ AR CA CO FL HI IA KY MD MA ME MI NC NJ NM NV NY OR RI SD UT VT WA WI WV	NE
2008	AZ AR CA CO FL HI IA KY MD MA ME MI NC NJ NV NY OR RI SD UT VT WA WI WV	NE
2007	AZ AR CA CO FL HI IA KY MD MA ME MI NC NJ NV NY OR RI UT VT WA WI WV	none
2006	AZ AR CA CO FL HI IA KY MD MA ME MI NC NJ NV NY OR RI UT VT WA WI WV	none
2005	AZ FL HI IA KY MD MA MI NV NJ NY NC OR RI UT VT WA WI WV	none
2004	AZ FL IA KY MD MA MI NV NJ NY NC OR RI UT VT WA WI WV	none
2003	AZ FL IA KY ME MD MA MI NV NJ NY NC OR RI UT WA WI WV	none
2002	AZ FL IA KY ME MD MA MI NV NJ NY NC OR UT WA WI WV	none
2001	AZ FL IA KY ME MD MA MI NV NJ NY NC OR UT WA WI WV	none

Notes: States listed in bold text were added to the revised annual CCR for CD-SID file. For 2018, the initial CCR for CD-SID was released without two pending states (NY, UT). For 2010-2017, one or more of the current CD-SID states were not available at the time the CCR files were created.

5. Hospital-Specific Cost-to-Charge Ratio (APICC)

   The hospital-specific all-payer inpatient cost-to-charge ratio (APICC) is created by dividing the inpatient costs by the inpatient charges. These values are obtained from the CMS Healthcare Cost Reporting Information System (HCRIS) reports, also known as PPS data.

   APICC is present on the CCR for CD-SID file for hospitals that meet these criteria:
   - APICC is kept for HCUP SID hospitals that have a matching record in both the HCRIS (PPS) file and the AHA Annual Survey of Hospitals data.
   - APICC is missing when there is no cost information in the HCRIS (PPS) data or the calculated cost-to-charge value is deemed an outlier.

   Several adjustments are made to costs and charges before they are usable in this generalized formula, the most important being the assignment of a portion of ancillary costs to inpatient routine and acute cost centers. Both operating costs and capital-related costs are included in the calculation of APICC.

6. Group Average Cost-to-Charge Ratio (GAPICC)

   The group average all-payer inpatient cost-to-charge ratio (GAPICC) is a weighted average for the hospitals in peer groups (defined by state, urban/rural, investor-owned/other, and bed size), using the proportion of each hospital's beds relative to their peer group as the weight for each hospital.

   These averages are based on all hospitals in the SID maintained by AHRQ (including SID that are not released through the HCUP Central Distributor), which amounts to approximately 4,900 hospitals participating in HCUP each year. *Clean records* are defined as HCUP hospitals that also have records in both the AHA and CMS data as of March 31st, when the CMS files are acquired. These records have a matching hospital in the CMS cost report, have availability of certain completed data items in the report, and pass certain edit checks. Note that a group average can be based on only 1 hospital in the peer group (defined by state and hospital type). The group average may incorporate non-HCUP hospitals. Both operating costs and capital-related costs are included in the calculation of GAPICC.

7. Hospital Type for Grouping (HTYPE)

   The hospital type for grouping peer hospitals (HTYPE) is calculated within state, using hospital characteristics obtained from the AHA Annual Survey. These include bedsize, type of ownership/control, and hospital urban/rural location.The GAPICC is calculated within state and each of these groupings.

   The following are values for the HTYPE variable:

   1= investor-owned, under 100 beds
   2= investor-owned, 100 or more beds
   3= not-for-profit, rural, under 100 beds
   4= not-for-profit, rural, 100 or more beds
   5= not-for-profit, urban, under 100 beds
   6= not-for-profit, urban, 100-299 beds
   7= not-for-profit, urban, 300 or more beds.

State and local nongovernment hospitals are included in the *not-for-profit* categories. *Urban* is defined as being part of a Metropolitan Statistical Area (MSA); *beds* are the total hospital beds set up (as reported in each year's AHA Annual Survey Database). *Teaching status*, which is often used for grouping HCUP hospitals was not incorporated into the definition of HTYPE. This indicator was not present on the CMS hospital cost reports. A proxy measure, the ratio of interns and residents per bed was tested, in regression analyses, with the result that the cost ratios by the proxy for teaching status were not significantly different. Therefore, only ownership and bed size were used for defining HTYPE.

8. Area Wage Index (WI_X)

   The Area Wage Index is computed by CMS to measure the relative hospital wage level in a geographic area compared to the national average hospital wage level. It is provided on the file to allow researchers to analyze cost differences geographically or to control for price factors beyond the hospital's control. The area wage index measure (WI_X) was obtained from the CMS wage index files available online. Hospital cost variation has a 0.8 elasticity with the area wage index in some AHRQ published studies, meaning that variation in the hospital cost is roughly proportional to the variation in overall hospital costs. Multivariate studies should not assume strict proportionality.

   The index is computed for each urban Core-Based Statistical Area (CBSA) and then linked with the AHA before it is added to the file. If the AHA-reported CBSA does not match the CMS hospital area, then the Area Health Resources Files (AHRF) and other hospitals in the same county are used to find a matching CBSA. All rural areas in each state are combined for a single wage index. This information is available for download from CMS. For the HCUP hospitals in each year all were matched to an area wage index using CMS files and the AHA survey.

9. Geographic Adjustment Factor (GAF)

   The Geographic Adjustment Factor is a capital cost adjustment index for Core Based Statistical Areas and is included on the file beginning with the 2009 data year. It is used in calculating the Medicare reimbursement payments for capital costs. The geographic adjustment factor measure (GAF) was obtained from the CMS wage index files available online. This data element may prove useful in regression calculations. Some states restrict the release of GAF on the CD-SID CCR file.

10. State Postal Code (Z013)

    The State Code (AHA element Z013) is the two-character state postal code (e.g., "AZ" for Arizona) for hospitals included in the CCR for CD-SID file.

11. Internal Validation Studies

    A regression analysis of the all-payer inpatient CCR was performed in earlier years. This analysis used all clean HCUP and non-HCUP records with both AHA and CMS data. This was a weighted OLS regression using acute medical-surgical beds as the weighting variable, with separate state constant terms. Factors leading to significant differences in the CCR were: investor-ownership, rural location, large size (more than 300 beds), and a high ratio of interns and residents per bed (top 5%). Several of the state constant terms

were also significant. The results tended to validate the "peer-grouping" method used here to create weighted group averages for each HCUP record.

In 2001 a study was performed for two states where hospital-based and department-based methods of calculating cost by Diagnosis Related Groups (DRG) were compared. This study concluded that hospital-wide CCRs as provided in the CCR for Central Distributor SID file, although not as accurate as department-based CCRs, are more accurate than gross charges in estimating relative cost by DRG. Subsequent studies were conducted involving a dozen states that report detailed charges. These studies produced more accurate cost estimates using departmental CCRs instead of hospital-wide CCRs. Users interested in quantifying potential biases due to use of the hospital-wide CCRs should contact HCUP User Support (hcup@ahrq.gov).

12. Tools for More Accurate Cost Estimates

HCUP periodically evaluates the differences in cost estimates by hospital and by cost centers (departments) and individual services. In general, department-specific CCRs are more accurate for deriving the cost of a hospital stay than hospital-wide CCRs. However, not all of the HCUP Partner organizations request that hospitals report detailed charges for every discharge, and not all hospitals have usable CMS accounting reports.

Based on these evaluations, HCUP has produced two sets of cost adjustment factors, for data years 2006 and 2009. The adjustment factors are contained in the appendices of the methods reports available at https://www.hcup-us.ahrq.gov/reports/methods/methods.jsp.

The initial report, conducted with 2006 data, provides adjustment factors by Clinical Classifications Software (CCS) categories and All-Patient Refined Diagnosis Related Groups (APR-DRG). The adjustment factors allow the analyst, after calculating costs based on hospital-wide CCRs, to correct the estimates. Such adjustments increase the estimated costs in some APR-DRG and CCS categories and reduce them in others. For more information about the approach, please see HCUP Methods Series Report # 2008-04. Song, X, Friedman, B. *Calculate Cost Adjustment Factors by APR-DRG and CCS Using Selected States with Detailed Charges*. Online October 8, 2008. U.S. Agency for Healthcare Research and Quality.

An updated report, conducted in 2012, used a more extensive methodology to develop adjustment factors for 2009 data for each Medicare-Diagnosis Related Group (MS-DRG) and each CCS category. These adjustment factors address a shortcoming of the hospital-wide CCR – that it does not account for variations among service departments in the hospital. This year's report created 13 cost-center clusters that take into account the higher markup (the inverse of CCR) for ancillary services as a whole than for routine bed-unit services. These adjustment factors allow the analyst, after calculating costs based on hospital-wide CCRs, to correct the estimates by MS-DRG or CCS category, producing a more accurate CCR and, hence, a more accurate cost estimate. For more information about the approach, please see HCUP Methods Series Report # 2011-04. Sun Y, Friedman B. *Tools for More Accurate Inpatient Cost Estimates with HCUP Databases, 2009.* Errata added October 25, 2012. 2012. ONLINE October 29, 2012. U.S. Agency for Healthcare Research and Quality.

13. Data Elements

The following tables list the data elements (and their respective labels) included in the Cost-to-Charge Ratio files for the Central Distributor SID and the separate, state-specific CCR files for Iowa, Minnesota, and Nebraska.

Table 4. Data Elements on CCR for CD-SID Files: 2009-2018

Data Element	Description
HOSPID	HCUP hospital identification number
APICC	All-payer inpatient CCR, hosp-specific
GAPICC	Group average all-payer inpatient CCR
GAF	Capital cost adjustment index for Core Based Statistical Areas
HTYPE	Hospital type used for grouping
WI_X	Wage Index, source CMS, edited
YEAR	Year for linking to HCUP records
Z013	State postal code

Table 5. Data Elements on CCR for CD-SID Files: 2001-2008

Data Element	Description
HOSPID	HCUP hospital identification number
APICC	All-payer inpatient CCR, hosp-specific
GAPICC	Group average all-payer inpatient CCR
HTYPE	Hospital type used for grouping
WI_X	Wage Index, source CMS, edited
YEAR	Year for linking to HCUP records
Z013	State postal code

REFERENCES

Friedman B, De La Mare J, Andrews R, McKenzie DH. Practical options for estimating cost of hospital inpatient stays. *J Health Care Finance*. 2002 Fall;29(1):1-13.

Song, X, Friedman, B. Calculate Cost Adjustment Factors by APR-DRG and CCS Using Selected States with Detailed Charges. HCUP Methods Series Report # 2008-04. U.S. Agency for Healthcare Research and Quality. Available: http://www.hcup-us.ahrq.gov/reports/methods.jsp.

Sun Y, Friedman B. Tools for More Accurate Inpatient Cost Estimates with HCUP Databases, 2009. [Errata added October 25, 2012.] HCUP Methods Series Report # 2011-04. Online October 29, 2012. U.S. Agency for Healthcare Research and Quality. Available: http://www.hcup-us.ahrq.gov/reports/methods/methods.jsp.

APPENDIX A: UPDATES TO ANNUAL CCR FOR CD-SID FILES

2018 CCR for CD-SID File

The 2018 CCR for CD-SID file was updated in December 2020 to add two states: New York and Utah. The initial file released in June 2020 did not include these states.

2010 CCR for CD-SID File

Please be aware that AHRQ released a revised version of the 2010 CCR files in August 2013. At the time the initial files were created, CMS had recently revised its standard accounting forms for hospitals, which apparently affected the timeliness of reporting for data year 2010. As of June 30, 2012, the CMS files used for the initial version of the CCRs contained usable 2010 accounting reports for only 61.5% of HCUP hospitals. For hospitals with no usable report, the CCR was imputed from a weighted average for a peer group within the state (the variable name is GAPICC). Several HCUP states had a particularly high proportion of hospitals with missing reports in 2010, which results in a smaller number of hospitals used for imputation. Hospitals with missing accounting reports in the initial files can be identified by the variables APICC and CLEANCC having missing values.

In the Spring of 2013, AHRQ obtained an updated file of 2010 accounting reports from CMS. As of May 2013, the CMS files used for the revised 2010 CCR files contained usable 2010 accounting reports for 89% of HCUP hospitals. For hospitals that were missing accounting reports in the initial files, the APICC was calculated from the updated reports, where permitted by HCUP Partner organizations. GAPICC was recalculated using the updated weighted average for a peer group within the state. The values of GAPICC in the revised CCR files may differ from the initial version because a larger number of hospitals was used for imputation.