Cost-to-Charge Ratio Files:

User Guide for Central Distributor State Inpatient Database (CD-SID) CCRs

1. Purpose

The purpose of this data file is to provide Healthcare Cost and Utilization Project (HCUP) users of the Central Distributor State Inpatient Databases (CD-SID) with ratios that will allow the conversion of charge data to cost estimates. The annual Cost-to-Charge Ratio (CCR) files are constructed using information from the Healthcare Cost Report Information System (HCRIS) files submitted by hospitals to the Centers for Medicare & Medicaid Services (CMS). Records are included for all community hospitals from the HCUP CD-SID that have "clean" matches with both the AHA Annual Survey Database and the CMS hospital cost file for the corresponding fiscal year.[1]

The HCUP CCR 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 2001-2016. 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 instead of including their data in the CCR for CD-SID.[2]

The CCR for CD-SID files can be linked to records in the SID using the data element HOSPID, which is a unique hospital number exclusive to the HCUP data. This is achieved in two steps, first linking the Cost-to-Charge file to the HCUP AHA Linkage File (provided with the Central Distributor SID) by HOSPID and then linking the resulting file to the Central Distributor SID by DSHOSPID. Some states will include HOSPID directly on their Central Distributor SID file (instead of releasing an AHA Linkage File). For these states, the Cost-to-Charge file can be merged directly onto the Central Distributor SID by HOSPID.

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

Note: HOSPID on the CCR CSV text file is enclosed in quotations, so it should be loaded as numeric or converted to numeric prior to merging with the CD-SID.

---

[1] HCUP uses the CMS fiscal year files like "hosp10_2016_NMRC.csv," also referred to as PPS records, for hospital data submitted through March 31st, approximately 18 months after the close of a fiscal year.

[2] The CCR for CD-SID files for certain years exclude records for Iowa, Minnesota, and Nebraska, which are released as separate, state-specific files available by request from the HCUP Central Distributor to purchasers whose primary affiliation is with a college/university/government and whose intended use of the data does not involve product development, market research, or commercial applications.

2. File Format

The datasets contain a record for each hospital (unduplicated HOSPIDs)[3] in states that release both their Central Distributor SID[4] and CCR measures:

Year	Record Counts	
	CCR for CD-SID	Total CD-SID
2016	2,297	2,629 hospitals
2015	2,178	2,514 hospitals
2014	2,078	2,415 hospitals
2013	1,876	2,081 hospitals
2012	2,313	2,517 hospitals
2011	2,412	2,558 hospitals
2010	2,519	2,666 hospitals
2009	2,402	2,548 hospitals
2008	2,367	2,514 hospitals
2007	2,305	2,451 hospitals
2006	2,313	2,463 hospitals
2005	1,691	1,830 hospitals
2004	1,613	1,771 hospitals
2003	1,626	2,257 hospitals
2002	1,602	2,234 hospitals
2001	1,578	2,210 hospitals

All HCUP hospitals in the CCR for CD-SID files are in the American Hospital Association (AHA) Annual Survey.

Where permitted by HCUP Partner organizations, the dataset provides a hospital-specific all-payer inpatient cost/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.

Year	Records with APICC	Percent	Records with GAPICC only
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

[3] The column labeled “Total CD-SID” in the table includes records for Iowa, Minnesota, and Nebraska. The column labeled “CCR for CD-SID” excludes these three States, and represents the record count for the CCR for CD-SID file.

[4] The CCR for CD-SID files include only the Central Distributor states that were available at the time the files were created; states that became Central Distributor participants at a later time are not included.

Year	Records with APICC	Percent	Records with GAPICC only
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

3. Participating States

Almost all of the states participating in the Central Distributor are included in the CCR for CD-SID file. [5] Three states – Iowa, Minnesota, and Nebraska – release their cost-to-charge measures in state-specific skinny files for certain years.

Year	States in CCR for CD-SID file	State-Specific Files
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

[5] For 2010-2014, one or more of the current CD-SID states were not available at the time the CCR files were created. For all years, South Carolina does not release its CCRs for CD-SID.

Year	States in CCR for CD-SID file	State-Specific Files
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

4. Hospital-Specific Cost/Charge Ratio (APICC)

   The all-payer inpatient cost-to-charge ratio (APICC) is created by dividing the inpatient costs by the inpatient charges. Both of these values are found on the CMS Healthcare Cost Reporting Information System (HCRIS) reports, or PPS data. APICC is kept for HCUP SID hospitals that have a matching record in both the PPS and the AHA data. APICC is missing when there is no cost information in the PPS data or the calculated cost/charge values were considered outliers. 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.

5. Weighted Group Average (GAPICC)

   The group average 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 the total set of hospitals from the full collection of HCUP states in each year, regardless of participation in releasing their SID through the Central Distributor. *Clean records* are defined as HCUP hospitals that also have records in both the AHA and CMS data as of the March 31st date 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 group averages can be based on only 1 hospital in the peer group (defined by state and hospital type). The group average may be associated with a non-HCUP hospital. Both operating costs and capital-related costs are included in the calculation of GAPICC.

   For each HCUP data year, there are up to 4,900 hospitals for all states participating in HCUP.

6. Hospital Type for Grouping (HTYPE)

   The hospital type for grouping peer hospitals (HTYPE) is available on the CCR for Central Distributor SID file. It is helpful to know how this variable is defined to create peer groups from all hospitals within each state – not only those participating in the Central Distributor SID. Some researchers will find the

information below useful with respect to replicability, and reviewers for journal articles might find this more detailed description especially valuable.

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 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 defined in each year's AHA Annual Survey Database). Teaching status, which is customarily used for grouping HCUP hospitals was not assigned; interns and residents per bed are not available on the AHA survey so a high value of this indicator of teaching status could not be used for grouping.

7. 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. 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.

8. Geographic Adjustment Factor (GAF)

   The Capital cost adjustment index for Core Based Statistical Areas is included on the file beginning with the 2009 data year. It is used in calculating the Medicare reimbursement payments for capital costs. This data element may prove useful in regression calculations. Some states restrict the release of GAF on the CD-SID CCR file.

9. State 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.

10. 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 different methods of calculating cost by DRG were compared. Hospital-wide CCRs as provided here, although not as accurate as department-based CCRs, are more accurate than gross charges in estimating relative cost by DRG. In more recent years, studies involving a dozen states that report their detailed charges have been done. These studies produced more accurate CCRs because they use departmental CCRs as opposed to 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).

11. Tools for More Accurate Cost Estimates

    HCUP periodically evaluates the differences in cost estimates by hospital and by cost centers (departments) and individual services.[6] There are two sets of cost adjustment factors available, for data years 2006 and 2009. The adjustment factors are contained in the appendices of the following methods reports available at https://www.hcup-us.ahrq.gov/reports/methods/methods.jsp.

    An 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 an analyst to correct cost estimates based on hospital-wide CCRs. Such adjustments will increase the estimated costs for patients in some APR-DRG and CCS categories and reduce the estimated costs for patients in other APR-DRG and CCS categories. For more information about the approach, please see HCUP Methods Series Report # 2008-04. Song, X, Friedman, B. *Calculate Cost Adjustment*

[6] 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 ask hospitals to report detailed charges for every discharge, and not all hospitals have usable CMS accounting reports.

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 correction factors for 2009 data for each Medicare-Diagnosis Related Group (MS-DRG) and each CCS category. This addresses an issue with the hospital-wide CCR in 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. The cost-center specific and hospital-wide CCRs were applied to SID discharges for each MS-DRG or CCS category. These adjustment factors allow an analyst to correct cost estimates based on hospital-wide CCRs for the patient's MS-DRG or CCS category to get 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.

12. Variable Lists

The following tables summarize the variables (and their respective labels) included in the Cost-to-Charge files for the Central Distributor SID.

CCR for CD-SID: 2009-2016

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

CCR for CD-SID: 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

Appendix A: Special note for the 2010 SID CCR 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 as a larger number of hospitals were used for imputation.