

## Cost-to-Charge Ratio Files:

### User Guide for National Inpatient Sample (NIS) CCRs

#### 1. Purpose

The purpose of this data file is to provide Healthcare Cost and Utilization Project (HCUP) users of the National (Nationwide) Inpatient Sample (NIS) 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). Separate CCR for NIS files are released for each data year and should be used with the corresponding year of the NIS.

*Beginning with the 2012 data, the CCR for NIS files were revised to reflect the redesign of the National Inpatient Sample. These versions of the file include HOSP\_NIS, which is a unique hospital number exclusive to the NIS. The CCR variables were also renamed to indicate they are designed to be used exclusively with the 2012 and later NIS.*

The HCUP CCR files provide an estimate of all-payer inpatient cost-to-charge ratios for hospitals in the 2001-2016 NIS. Records are included for all community hospitals from the HCUP NIS that have “clean” matches with both the AHA Annual Survey Database and the CMS hospital cost file for the corresponding fiscal year.<sup>1</sup>

The CCR for NIS files can be linked to records in the NIS using the data element HOSPID, through 2011, and the data element HOSP\_NIS, for 2012 and later years. HOSP\_NIS is reassigned each year and does not link to other HCUP databases or to external databases, or track hospitals over years.

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 or the group average all-payer inpatient cost/charge ratio. The hospital-specific cost/charge data element is named APICC, through 2011, and is named CCR\_NIS, for 2012 and later years. CCR\_NIS is based on the hospital-specific all-payer inpatient cost/charge ratio (APICC) when available, or the group average all-payer inpatient cost/charge ratio (GAPICC) otherwise.

---

<sup>1</sup> HCUP uses the CMS fiscal year files like “hosp10\_2015\_NMRC.csv,” also referred to as PPS records, for hospital data submitted through March 31<sup>st</sup>, approximately 18 months after the close of a fiscal year.

## 2. File Format

*Beginning with the 2012 data, the cost-to-charge ratio has been modified to enhance confidentiality of the NIS. Statistical reliability of the estimates is not affected.*

For 2012-2016, the datasets contain a record for each hospital (unduplicated HOSP\_NIS) in the NIS:

- 4,575 hospitals in the 2016 CCR-NIS
- 4,573 hospitals in the 2015 CCR-NIS
- 4,411 hospitals in the 2014 CCR-NIS
- 4,363 hospitals in the 2013 CCR-NIS
- 4,378 hospitals in the 2012 CCR-NIS.

For 2001-2011, the datasets contain a record for each hospital (unduplicated HOSPIDs) in the NIS for states that permitted release of their cost-to-charge ratios:<sup>2</sup>

- 1,008 of 1,049 hospitals in the 2011 CCR-NIS
- 1,010 of 1,051 hospitals in the 2010 CCR-NIS
- 1,013 of 1,050 hospitals in the 2009 CCR-NIS
- 1,017 of 1,056 hospitals in the 2008 CCR-NIS
- 1,044 of 1,044 hospitals in the 2007 CCR-NIS
- 1,045 of 1,045 hospitals in the 2006 CCR-NIS
- 942 of 1,054 hospitals in the 2005 CCR-NIS
- 908 of 1,004 hospitals in the 2004 CCR-NIS
- 855 of 994 hospitals in the 2003 CCR-NIS
- 855 of 995 hospitals in the 2002 CCR-NIS
- 851 of 986 hospitals in the 2001 CCR-NIS.

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

*Note: The HCUP hospital identifier, HOSP\_NIS or HOSPID, on the CCR CSV (comma-separated value) text file is enclosed in quotation marks in order to preserve leading zeros in Excel. As a result, some software applications may interpret HOSP\_NIS (HOSPID) as a character variable, which in turn would not match the numeric version of HOSP\_NIS (HOSPID) on the NIS. This data element should be loaded as numeric or converted to numeric prior to merging with the NIS.*

---

<sup>2</sup> Through 2011, one or more states did not provide permission to participate in the NIS CCR.

### 3. Linkage to the NIS

Where permitted by HCUP Partner organizations, the files provide a hospital-specific all-payer inpatient cost-to-charge ratio, APICC (renamed as CCR\_NIS). For all hospitals, there is also a weighted group average, GAPICC. Analysts can use APICC, when available, and otherwise use the weighted group average, GAPICC. For 2012 and later files, the cost-to-charge ratio element CCR\_NIS can be used for all hospitals.

Year	Records with APICC	Percent	Records with GAPICC only
2011	889	88%	119
2010	920	91%	90
2009	906	89%	107
2008	888	87%	129
2007	818	78%	226
2006	807	77%	238
2005	657	70%	285
2004	669	74%	239
2003	641	75%	214
2002	571	67%	284
2001	623	73%	228

### 4. NIS-Specific Cost-to-Charge Ratio—CCR\_NIS

The cost-to-charge ratio element CCR\_NIS, provided for linkage to the 2012 and later NIS, is populated with the all-payer inpatient cost-to-charge ratio (APICC), when available. The hospital group average CCR (GAPICC) is used to populate CCR\_NIS when the APICC is not available. The construction of the all-payer inpatient and group average CCR are described in the next sections.

### 5. 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 NIS 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 are 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.

### 6. 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

bedsize), using the proportion of each hospital's beds relative to their peer group as the weight for each hospital.

These averages are based on clean observations, meaning the HCUP hospitals that also have records in both the AHA Annual Survey and CMS cost 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 it is possible for group averages to be based on only one hospital in the peer group (defined by state and hospital type). The group average may be associated with a non-HCUP hospital.

## 7. Hospital Type for Grouping—HTYPE

The hospital type (HTYPE) is utilized for grouping peer hospitals. Although HTYPE is not provided on the CCR for NIS files, it is helpful to know how this variable is defined to create peer groups using all hospitals within each state. 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.

## 8. Area Wage Index—WI\_X / WAGEINDEX

The Area Wage Index is an index 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 data 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 NIS hospitals in each year, all hospitals were matched to an area wage index using CMS files, the AHA survey, and the Area Resource File in cases where the AHA survey was incomplete.

Through 2011, this data element is called WI\_X. Beginning in 2012, the wage index in the CCR file has been modified to enhance confidentiality of the NIS and renamed as WAGEINDEX. Statistical reliability of the estimates is not affected.

#### 9. Geographic Adjustment Factor—GAF

The Capital cost adjustment index for Core Based Statistical Areas is included on the file in earlier years. It is used in calculating the Medicare reimbursement payments for capital costs. This data element may prove useful in regression calculations. However, analysts should note that for a number of states contributing hospital data in the NIS, permission was not provided to release values of GAF.

#### 10. State Code—Z013

The State Code (AHA element Z013), provided in earlier years, is the two-character state postal code (e.g., “AZ” for Arizona) for hospitals included in the CCR for NIS 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 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](mailto: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.<sup>3</sup> 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 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.

---

<sup>3</sup> 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.

### 13. Variable Lists

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

#### CCR for NIS: 2012-2016

Data Element	Description
HOSP_NIS	NIS hospital number
CCR_NIS	All-payer inpatient or group average CCR
WAGEINDEX	Wage Index, source CMS, edited
YEAR	Year for linking to HCUP records

#### CCR for NIS: 2001-2011

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

## Appendix A: Special Notes for Year-Specific Files

### Revised 2010 CCR-NIS File

Please be aware that AHRQ released a revised version of the 2010 CCR for NIS file 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 amount of missing data for APICC.

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.

### Adjustments for 2004 Cost Estimates for the Northeast

Calculations of costs for the 2004 NIS in the Northeast (HOSP\_REGION=1) will tend to be slightly overestimated due to the absence this year of one large state, Pennsylvania. An example of this is found in a comparison of the HCUPnet tables of average cost by region for 2004 versus 2003.

To obtain a more accurate cost-to-charge ratio for the Northeast region, users may want to multiply the total cost of care in each stratum (NIS\_STRATUM) shown below, where the first digit ('1') represents the Northeast, by the following adjustment factors:

NIS_STRATUM	Adjustment Factor
1011	1.057
1012	1.071
1013	1.105
1021	1.078
1022	1.099
1023	1.003

NIS_STRATUM	Adjustment Factor
1031	.845
1032	.863
1033	.777

For national studies where users are comparing 2004 versus 2003 national estimates and are including the Northeast, a downward adjustment can be made to *all* CCRs by 3% to correspond with other national findings.

### Changes in 2005 Cost Estimates from Previous Years

The cost/charge ratios in the 2005 CCR for NIS file, when applied to the 2005 Nationwide Inpatient Sample, result in somewhat higher mean cost than anticipated. This atypical increase in the national CCR can be associated with a change in the hospital universe for the NIS. This phenomenon can happen in any year, but in 2005 it had the effect of increasing the national CCR.

The now-archived annual Facts and Figures Report at [http://www.hcup-us.ahrq.gov/reports/factsandfigures/HAR\\_2005.pdf](http://www.hcup-us.ahrq.gov/reports/factsandfigures/HAR_2005.pdf) was prepared before the CMS 2005 data were assembled. It is based on preliminary assumptions that give a lower overall CCR than we find in the more recent data. As a result, the mean cost for 2005 now appears to be \$8079 rather than the \$7900 estimated in the annual report, a difference of 2%.

### Adjustments for National Cost Estimates in Specific Years

To obtain national cost estimates for a set of cases, for analyses through 2011, it is recommended that users re-weight all discharges to account for cases where cost estimates are missing.

The original discharge weight (DISCWT) should be multiplied by the following: Total weight of original cases divided by total weights, after excluding cases with missing cost. By performing these calculations, the weights for remaining cases are increased.

For several years, the national cost estimates may be affected in particular by missing states or states that provide only group average cost/charge ratios.

**2008-2011 CCR-NIS File:** One state was dropped from the file (PA). One state, ME, only includes group average.

**2007 CCR-NIS File:** No states were dropped from the file. One state, ME, only includes group average.

**2006 CCR-NIS File:** No states were dropped from the file. No states include only the group average.

**2005 CCR-NIS File:** One state was dropped from the file (TX). Three states—NE, OK, and OR—only include the group average.

**2004 CCR-NIS File:** One state was dropped from the file (TX). Two states, NE and OR, only include the group average.

**2003 CCR-NIS File:** Two states were dropped from the file (PA and TX). One state, NE, only includes group average.

**2002 CCR-NIS File:** Three states were dropped from the file (HI, PA, and TX). Two states, NE and OR, only include group average.

**2001 CCR-NIS File:** Three states were dropped from the file (HI, PA, and TX). Two states, NE and OR, only include the group average.