Cost-to-Charge Ratio Files:

2001 Central Distributor State Inpatient Database (CD SID) User Guide

1. Purpose

   The purpose of this data file is to provide Healthcare Cost and Utilization Project (HCUP) data users with ratios that will allow the conversion of charge data to cost estimates. The file is constructed using all-payer, inpatient cost and charge information from the detailed reports by hospitals to the Centers for Medicare and Medicaid Services (CMS). It provides an estimate of all-payer inpatient cost-to-charge (CCR) for hospitals in states that participate in the 2001 SID Cost-to-Charge Central Distributor. The participating states are: AZ, FL, IA, KY, ME, MD, MA, MI, NV, NJ, NY, NC, OR, UT, WA, WI, and WV. Where permitted by HCUP State Partners, the dataset provides a hospital-specific CCR and a weighted group average.

   The file can be linked to participating 2001 Central Distributor SID files by using the HOSPID variable (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). This is achieved by first linking the Cost-to-Charge file to the hospital linkage file (that comes with the Central Distributor SID) by HOSPID and then linking the result to the Central Distributor SID file by DSHOSPID. Some states will include HOSPID directly on the CD SID file because they do not release AHAID on the hospital linkage file. For these states, the Cost-to-Charge file can be merged directly onto the CD SID file by HOSPID. The cost of inpatient care for a discharge can then be estimated by multiplying TOTCHG (from the discharge record) by either the hospital-specific cost-to-charge ratio (APICC), or the group weighted average cost-to-charge ratio (GAPICC).

2. File Format

   The dataset contains one record for each of the 1578 of 2210 Central Distributor SID hospitals in 2001 (unduplicated HOSPIDs). All HCUP hospitals in the file are also in the American Hospital Association (AHA) 2001 survey.

   Analysts might want to use the hospital-specific cost-to-charge when available (984 cases approximating 62%) and the weighted group average when the hospital-specific CCR is not available (594 cases). Alternatively, one might use the group average in all cases.

Four states were dropped from the file (CA, CO, NE, and SC). Two states, MI and OR, only include the group average.

3. Internal Validation Studies

   A regression analysis of the all-payer inpatient CCR was performed this year and in earlier years. This analysis used all clean HCUP and non-HCUP records with both AHA and PPS data. (Clean records are defined as having complete CMS schedules and worksheets, containing key variables within an acceptable range.) 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.

   A second type of validation study was performed for two states. In one case, the state accounting system by department was taken as the "gold standard" for cost estimation. Three alternatives were compared as predictors of differences in cost by DRG: Centers for Medicare and Medicaid Services (CMS) departmental cost-to-charge ratios, CMS hospital-wide inpatient cost-to-charge, and raw charges. The mean-squared-error criterion was used. The CMS departmental cost-to-charge ratios applied to detailed charges are somewhat more accurate in predicting the "gold standard" costs than are the hospital-wide inpatient cost-to-charge. The latter is substantially more accurate as a predictor than the raw charges. Unfortunately, detailed charges are not available for all HCUP states, so we can only use the hospital-wide inpatient cost-to-charge for cost estimation with the CD SID.

4. Weighted Group Average—GAPICC

   The group average CCR (GAPICC) is a weighted average for the hospitals in the group (defined by state, urban/rural, investor-owned/other, and number of beds), using the proportion of group beds as the weight for each hospital. The groups are defined based on all clean HCUP and non-HCUP records for community hospitals with matching AHA 2001 Annual Survey data and for CMS accounting database records as of December 31, 2003. Both operating costs and capital-related costs are included.

5. Hospital Type for Grouping—HTYPE

   HTYPE is available on the Central Distributor SID Cost-to-Charge file. It is helpful to know how this variable is defined to create peer groups within each state using all hospitals – 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.

Unfortunately, data about the ratio of 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. *Urban* is defined as being part of a Metropolitan Statistical Area (MSA); *beds* are the total hospital beds set up (2001 AHA survey).

6. Area Wage Index—WI_X

The area wage index is computed by CMS for each urban MSA. All rural areas in a state are combined for a single wage index. This information is available for download from CMS. For the HCUP hospitals in 2001, all but 256 were matched to an area wage index using CMS and the AHA survey. Editing was performed because of missing or inaccurate MSAs, or because of special numbering for New England cities and towns. After editing, only 24 hospitals have a missing wage index. One caution is that some urban hospitals have been allowed higher area wage indexes in federal regulations than found in the file. Sub-MSA special wage indexes were assigned to 2.5% of urban MSAs.

7. Variable List

There are seven variables in the Central Distributor SID Cost-to-Charge file. The following list summarizes the variables (and their respective labels) included in the Cost-to-Charge data file.

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

YEAR	Year for linking to HCUP records