Cost-to-Charge Ratio Files:

2005 Nationwide Inpatient Sample (NIS) 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 nearly every HCUP NIS hospital in 2005. Where permitted by HCUP State Partners, the dataset provides a hospital-specific CCR and a weighted group average.

   The file can be linked to the 2005 file of NIS charges using the HOSPID variable. The HOSPID variable on the CCR CSV text file is enclosed in quotations in order to preserve leading zeros in Excel. 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 NIS. This data element should be loaded as numeric or converted to numeric prior to merging with the NIS.

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

2. File Format

   The dataset contains one record for each of 942 of 1054 total HCUP NIS hospitals in 2005 (unduplicated HOSPIDs). All HCUP hospitals in the file are also in the American Hospital Association (AHA) 2005 survey.

   Analysts might want to use the hospital-specific cost-to-charge when available (657 cases approximating 70%) and the weighted group average when the hospital-specific CCR is not available (285 cases). Alternatively, one might use the group average in all cases.

   One state was dropped from the file (TX). Three states, NE, OK, and OR, only include the group average. To obtain national cost estimates for a set of cases, users will need to re-weight all discharges to account for cases where cost estimates are missing. The original case 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.

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 CMS 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 NIS.

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 2005 Annual Survey data and CMS accounting database records as of June 30, 2007. Both operating costs and capital-related costs are included.

5. Hospital Type for Grouping—HTYPE

   Although HTYPE is not provided on the NIS 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 selected for the NIS. 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 (2005 AHA survey). State and local hospitals are included in the not-for-profit categories.

6. Area Wage Index—WI_X

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 .8 elasticity with the area wage index in some AHRQ published studies, meaning that changes in the hospital wages are closely linked to changes in overall hospital costs. The index is computed for each urban Core-Based Statistical Area (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 2005, all hospitals were matched to an area wage index using CMS and the AHA survey.

7. Changes from Previous Years

The CCRs in this 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 Annual Report at 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%.

8. Variable List

There are six variables in the NIS Cost-to-Charge data file in 2005. The following list summarizes the variables (and their respective labels) included in this 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 avg. all-payer inpatient CCR
YEAR	Year for linking to HCUP records