

Calculating Community-Level Statistics for HCUPnet: Methods

This document provides details on the methods used to develop community-level statistics based on Healthcare Cost and Utilization Project (HCUP) data for HCUPnet.

Purpose of Community-Level Statistics

The Agency for Healthcare Research and Quality (AHRQ) has developed county-level information to be used by local communities, State and Federal agencies, hospitals, and health care researchers. County-level health data provide the focused view necessary to support health policy and improvements in the health care system. Community-level statistics represent measures created at the county-level or county equivalent-level. The term county-level statistics used here or under the Community-Level Statistics link on the HCUPnet website should be understood as representing county-level statistics or county-equivalent statistics (e.g. boroughs or parishes)

County-level statistics will be included on HCUPnet as a new drill-down category. Users can query volume, rates, and costs for all inpatient discharges in the county and by selected diagnosis and procedure categories. Information can be subdivided further by demographic characteristics such as sex, age group, and payer type, when possible. State-level and national benchmarks are also presented.

Metrics

The community-level metrics capture various measures of hospital utilization and expenditures (Table 1).

Table 1. Metrics for Reporting

Total number of discharges
Rate of discharges per 100,000 population
Mean length of stay, days
Aggregate number of days in the hospital
Number of inpatient days per 100,000 population
Mean cost per stay, \$
Aggregate costs for all hospital stays, \$
Cost for inpatient stays per capita, \$

The metrics are reported at the county level for all discharges, all major diagnostic categories (MDCs), selected Clinical Classifications Software (CCS) principal diagnoses, and selected CCS all-listed procedures. The CCS categories were chosen based on volume and procedures were limited to major operating room procedures (PCLASSn = 3 [major diagnostic] or 4 [major therapeutic]).

Stratification

Users can obtain county-level metrics stratified by sex, age, and expected payer.

The categories include:

- Sex: Male, Female, Missing
- Age Groups, Year: <1, 1–17, 18–44, 45–64, 65–84, 85+, Missing
- Expected Payer: Medicare, Medicaid, Private Insurance, Uninsured, Other, Missing

At present, there are no AHRQ-endorsed estimates of payer-specific denominators that can be used to compute payer-specific per capita rates, so population-based statistics by payer are not available in the community-level statistics.

Hospital Selection

The HCUP State Inpatient Databases (SID) are the primary data sources for the county-level statistics. All discharges from community hospitals, as defined by the American Hospital Association (AHA), were included, minus long term acute-care facilities.

Community hospitals are defined by the AHA as “all non-Federal, short-term, general, and other specialty hospitals, excluding hospital units of institutions.” Discharges from long-term acute care facilities were specifically excluded from the county-level statistics because evaluation of discharges from these hospitals revealed significantly longer lengths of stay and higher mortality rates than other community hospitals. Also, diagnoses, treatment, and procedures tend to be different from long-term acute care facilities than from other community hospitals.

County Selection

Community hospitals may be missing from the SID because some statewide data organizations exempt certain types of hospitals (e.g., small rural hospitals) from reporting or reporting is voluntary in some States, thus not all hospitals are included from the data sources. Missing hospitals may have small discharge volumes or be geographically concentrated. Alternatively, missing hospitals may have large volumes and be geographically dispersed. The Medicare Hospital Service Area File (HSAF) was used to estimate the impact of missing hospitals on HCUP community statistics and, therefore, to identify counties with incomplete data.¹ The HSAF provides the universe of Medicare discharges in the United States and contains the patient’s ZIP Code, Medicare provider identification number (ID), and a sum of patient discharges, days, and charges for all Medicare patients. This deidentified file is available to the public.

Capture rates computed from the HSAF and SID allowed us to examine several thresholds for suppression of county information that is due to missing hospitals in the SID. As a result of this investigation, counties where the capture rate was less than 98 percent were suppressed (the actual percentages were rounded).

¹ See Centers for Medicare & Medicaid Services. Hospital Service Area File. Last modified July 5, 2013. <http://www.cms.gov/Research-Statistics-Data-and-Systems/Files-for-Order/NonIdentifiableDataFiles/HospitalServiceAreaFile.html>. Accessed August 12, 2013.

In addition, counties were excluded from States that did not contribute SID data in 2011 at the time of the development of these statistics. US territory counties or counties with invalid county information and States that do not participate in HCUP were also excluded. County-level statistics were produced for all contributing States. County-level statistics are published on HCUPnet only after each State gives written permission to publish their statistics.

Suppression

All metrics based on fewer than 11 observations were suppressed, which is consistent with the terms of the HCUP Data Use Agreement (DUA). Results that could indirectly identify a hospital were also suppressed (i.e., at least two hospitals were represented in all cells). These suppression rules applied to the measure numerator.

Population Estimates and Assignment of Patient County

All states collect zip code data for patients. The county was identified for each discharge in the SID using the patient's ZIP Code and the Nielsen ZIP Code to cross reference the county. Nielsen is a vendor that compiles and adds value to the U.S. Bureau of Census data. Patients with missing or invalid counties or patients from U.S. or foreign territories were excluded.

County-level population estimates by sex and age were used for the per capita measures. Two sources of population estimates were considered. The first source was the Population File for Use with AHRQ Quality Indicators, which is based on actual county-level data available from the Census Bureau.² The second source was ZIP Code-level population estimates from Nielsen. Nielsen uses intracensus methods to estimate household and demographic statistics for geographic areas.³ Nielsen was used rather than the Population File because the census does not provide intracensus ZIP Code-level population estimates. Also, using the ZIP Code-based proximal county population estimates is consistent with the SID patient residence information, which is based on ZIP Code of the patient's residence.

In 2010, the Census Bureau made changes to counties in Alaska.⁴ These changes have not yet been incorporated in the Nielsen ZIP-to-county crosswalk. Therefore, pre-2010 county boundaries in Alaska were used in community-level statistics.

The county population data from Nielsen includes a 0–4 year age group. For community-level statistics, the population aged less than 1 year in each county was estimated by dividing the 0–4 age group by 5. This assumed a uniform distribution by age in the population. After subtracting the <1 estimates, the remainder of the 0–4 age group was combined with the older group (ages 1–17).

² See AHRQ. 2012 Population File for Use with AHRQ Quality Indicators. Version 4.4. March 2012. <http://www.qualityindicators.ahrq.gov/Downloads/Software/SAS/V44/AHRQ%20QI%20Population%20File%20V4.4.pdf>. Accessed August 12, 2013.

³ For a description of the Nielsen methodology, see Nielsen Pop-Facts™ Methodology. July 2012. <http://www.tetrad.com/pub/documents/popfactsmeth.pdf>. Accessed August 12, 2013.

⁴ See United States Census Bureau. Last modified December 5, 2012. <http://www.census.gov/geo/reference/county-changes.html>. Accessed August 12, 2013.

Reporting Cell Decision Rules and Handling Missing Data

HCUPnet cell suppression rules were applied. These rules require the exclusion of a reporting cell (i.e., a combination of a metric and stratification variable level for a given county) that draws from fewer than 2 hospitals or contains less than 11 discharges. Data from counties that did not meet minimum reporting rules (i.e., less than 11 discharges or fewer than 2 hospitals) were suppressed and not released on HCUPnet.

The HCUP data used in the production of the statistics included discharge counts, length of stay, charges, CCS principal diagnosis, MDC, CCS all-listed procedures category, sex, age, and expected payer. Missing charges and length of stay were imputed by assigning the average charges for the patient's State and DRG.⁵ Missing values for the patient's age, sex, county, or expected primary payer were included in the stratification analyses as a missing category. One exception was made for expected primary payer, if the payer was missing and the patient was aged 65 years or older; in those cases, Medicare was assumed as the primary expected payer.

The cost of inpatient care for a discharge was estimated by multiplying total charges by the all-payer inpatient cost-to-charge ratio or by the group average all-payer inpatient cost-to-charge ratio based on data from Medicare Cost Reports from the Center for Medicare and Medicaid Services (CMS).⁶

National and State Comparisons

The community statistics reports include State and national values as benchmarks. Computation of the national and State-level statistics followed slightly different procedures when compared to the county-level statistics.

The HCUP Nationwide Inpatient Sample (NIS) is the data source for national benchmark values. Long-term acute care facilities were excluded from the NIS to be consistent with the hospital selections used for county-level reporting. As a result, the national statistics developed as benchmarks for the community statistics project differ from NIS statistics reported elsewhere in HCUPnet.

State-level benchmarks involved creation of discharge-level completion weights, which were functions of hospital strata (i.e., ownership, location, teaching status, and bed size). The completion weight was the quotient of stratum-specific total discharges reported in the 2011 AHA survey and the total discharges in the SID for the same stratum. It is important to note that these calculations used patient residence rather than hospital location to identify the State, which is the standard with the NHQR/DR, and thus will differ from statistics presented in the NHQR/DR.

⁵ Technically, the data element DRG_NoPOA was used to classify the patient.

⁶ See "Costs" in HCUPnet definitions:

<http://hcupnet.ahrq.gov/HCUPnet.jsp?Id=2D287D7DB2D69367&Form=MAINSEL&JS=Y&Action=%3E%3ENext%3E%3E&HCUPnet%20definitions.%20x=1>.