Estimating Total Emergency Department Visits: A Comparison of Two Data Sources

Steven Machlin, MS and Ryan Mutter, PhD
Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850

1. Introduction[1]

There are many data sources that provide national estimates of the number of emergency department (ED) encounters. These sources vary in the methodologies they employ. Among the approaches used are the collection of billing data, surveys of hospitals and households, and medical record abstraction. The use of different data collection strategies can affect the number of ED encounters estimated. For example, whereas survey data are prone to non-response and measurement error, administrative billing data can be affected by payer billing practices. Owens et al. (2010) provide a profile of eight different national ED data sources and an overview of the issues associated with the various estimation approaches.

This analysis compares ED visit counts from two sources: the Healthcare Cost and Utilization Project (HCUP) and the American Hospital Association Annual Survey (AHA). In particular, counts for 2007 are compared at the total, State, and institution levels for the 27 States where ED-level visit counts are available from both data sources. It also examines which hospital characteristics are associated with agreement and directional differences in estimates between the two data sources.

2. Data and Methods

2.1 HCUP Data Background

Healthcare Cost and Utilization Project (HCUP) data are an example of administrative data based on bills generated by the hospital. HCUP is a Federal-State-Industry partnership sponsored by the Agency for Healthcare Research and Quality (AHRQ).[2] ED data in HCUP are provided from community, short-term, non-federal, non-rehabilitation

[1] The views expressed in this paper are those of the authors and no official endorsement by the Department of Health and Human Services (DHHS) or the Agency for Healthcare Research and Quality (AHRQ) are intended or should be inferred. The authors wish to thank Joel Cohen, Pamela Owens, and Trena Ezzati-Rice for their helpful reviews of this paper. The authors gratefully acknowledge the data organizations in participating states that contributed data to the Healthcare Cost and Utilization Project (HCUP) and that we used in this study: Arizona Department of Health Services, California Office of Statewide Health Planning & Development, Connecticut Hospital Association, Florida Agency for Health Care Administration, Georgia Hospital Association, Hawaii Health Information Corporation, Indiana Hospital Association, Iowa Hospital Association, Kansas Hospital Association, Maine Health Data Organization, Maryland Health Services Cost Review Commission, Massachusetts Division of Health Care Finance and Policy, Minnesota Hospital Association, Missouri Hospital Industry Data Institute, Nebraska Hospital Association, New Hampshire Department of Health & Human Services, New Jersey Department of Health & Senior Services, New York State Department of Health, North Carolina Department of Health and Human Services, Ohio Hospital Association, Rhode Island Department of Health, South Carolina State Budget & Control Board, South Dakota Association of Healthcare Organizations, Tennessee Hospital Association, Utah Department of Health, Vermont Association of Hospitals and Health Systems, and Wisconsin Department of Health Services.
[2] For more information, see http://www.hcup-us.ahrq.gov/overview.jsp.

hospitals with EDs by State data organizations, hospital associations, and private data organizations. The universe of HCUP ED encounters in a State is comprised of treat-and-release ED encounters (i.e., visits to the ED that do not result in admission to the same hospital) and ED encounters that result in admission to the same hospital. Treat-and-release ED encounters account for approximately 84 percent of all ED visits and are captured in the State Emergency Department Databases (SEDD) for participating States.[3] ED encounters resulting in admission to the same hospital comprise approximately 16 percent of all ED visits and are captured in the State Inpatient Databases (SID) for participating States.[4] In 2007, 27 States provided both a SEDD and a SID to HCUP.[5] These data sets captured nearly 60 percent of all ED visits in the United States. HCUP uses SEDD and SID data with ED counts from the American Hospital Association Annual Survey as control totals to create the Nationwide Emergency Department Sample (NEDS), which yields national estimates and is the largest publicly available all-payer ED database in the U.S.[6]

2.2 AHA Data Background

The AHA Annual Survey is sent out to over 6,300 hospitals in the U.S. and its territories. These hospitals are asked about their organizational structure, personnel, financial performance, services offered, and utilization, including the number of ED visits. The average response rate for the 2007 survey is 85 percent. The AHA imputes some missing data in the survey, based on previous responses from hospitals with similar size, ownership, services, length of stay and geography, or estimates from regression models (Health Forum, 2008). There were 4,809 EDs in community, non-rehabilitation U.S. hospitals that reported total ED visits in the 2007 AHA Annual Survey.

2.3 Analytic Sample

We based our analysis on 2,455 hospital-based EDs from 27 States that could be identified in both HCUP and AHA data in 2007. For each of these hospital-based EDs, we compared the total number of ED encounters from HCUP (i.e., the sum of SEDD records and the SID records with evidence of ED services) to the corresponding AHA survey total (hospitals provide the total number of ED visits in response to survey question D.1.g).[7] Other hospital characteristics (described in section 2.5 below) were obtained from the 2007 AHA.

2.4 Analysis

We compared descriptive statistics from the 2 sources on: 1) total ED counts for the 27 States as a whole, 2) distributions of ED counts at the hospital-level for the 27 States as a whole, and 3) relative differences in State-level estimates of total ED visits. In addition, we used logistic regression models to examine variation across hospital-based EDs in the relative odds that: 1) ED visit counts differ substantially between the 2 sources and 2) ED visit counts are higher/lower for one source than the other.

---

[3] For more information, see http://www.hcup-us.ahrq.gov/seddoverview.jsp.
[4] For more information, see http://www.hcup-us.ahrq.gov/sidoverview.jsp.
[5] The 27 states are AZ, CA, CT, FL, GA, HI, IA, IN, KS, MA, MD, ME, MN, MO, NC, NE, NH, NJ, NY, OH, RI, SC, SD, TN, UT, VT, and WI.
[6] For more information, see http://www.hcup-us.ahrq.gov/nedsoverview.jsp.
[7] The 2007 survey instrument is available online at http://www.ahadata.com/ahadata/PDFs/2008/2007AHAAnnualSurvey.pdf.

2.5 Logistic Regression Models

Using a criterion of 10% to identify noteworthy differences, the dependent variable for the first model was set to 1 if the ratio of the HCUP ED visit count to the AHA visit count was less than 0.90 or greater than 1.10 (about 42.5% of the observations). All other observations were assigned a value of 0. We fit another logistic regression model to examine variation in directional differences between the HCUP and AHA data. In particular, because aggregate counts based on HCUP were slightly lower than based on AHA (see sections 3.1-3.2 below), in the second model the dependent variable was assigned a 1 if the HCUP count was less than the AHA count (about 60% of the observations) versus a 0 if the HCUP count was equal or greater than the AHA count. The following variables that are commonly used to explain/control for hospital behavior were included as predictors in the models:

- State (26 indicators reflecting the 27 HCUP States),
- Teaching status indicator[8]
- Urban/rural indicator[9]
- Ownership status (2 indicators that distinguish for-profit, non-profit and government hospitals),
- Number of hospital beds (2 indicators that distinguish 3 bed size categories), and
- Volume of ED visits (5 indicators that distinguish 6 categories based on AHA data).

In addition, we included an indicator for critical access hospital (CAH) status. CAHs comprise about 27% of the sample. This hospital designation was created by the Balanced Budget Act (BBA) of 1997 with the intention of enhancing the financial viability of small, isolated rural and "necessary provider" hospitals by paying them on a cost basis instead of prospectively (Rosko and Mutter, 2010). This variable was included because the Centers for Medicare & Medicaid Services (CMS) requires that inpatient and outpatient services be billed separately for Medicare beneficiaries treated in CAHs (Medicare Learning Network, 2005). As a result, a Medicare beneficiary who visits the ED of a CAH before being admitted may generate both a SEDD and a SID record in HCUP. Therefore, HCUP may overcount ED encounters for Medicare beneficiaries visiting CAHs.

3. Results

3.1 Aggregate Comparisons and by ED Volume

Table 1 shows the total number of ED visits in HCUP and the AHA in the 27 States (row 1) and also compares the percentage distribution of EDs by ED volume category for the 2 data sources. While these distributions are generally quite similar, the slightly higher overall total visits for the AHA (69.8 million versus 67.4 million for HCUP) is mainly driven by a slightly higher proportion of hospitals reporting 50,000 or more encounters (16.9 vs. 15.9 percent).

[8] A hospital is considered to be a teaching hospital if it has an American Medical Association (AMA)-approved residency program, is a member of the Council of Teaching Hospitals (COTH), or has a ratio of full-time equivalent interns and residents to beds of .25 or higher.

[9] A Core Based Statistical Area (CBSA) was considered urban, and a non-metropolitan statistical area was classified as rural.

Table 1: Total Number of ED Encounters and Percentage Distribution of EDs by ED Volume Category in HCUP and the AHA for 27 States, 2007

	HCUP	AHA
Total ED Encounters	67.4 million	69.8 million
ED Volume (# of Encounters in 1000s)	Distribution of Hospital-based EDs	
< 10	29.4%	29.7%
10-19	18.6%	18.5%
20 – 29	15.0%	15.4%
30 – 39	12.3%	11.4%
40– 49,	8.7%	8.2%
50+	15.9%	16.9%

Table 2 shows the extent of concordance between HCUP and AHA data by ED volume category for the analytic sample of hospital-based EDs. The diagonal of this table (highlighted numbers) represents the EDs in the same category for each data source and together comprise 77.7 percent of all EDs in the analysis. An additional 17.1 percent of the hospitals are classified by the sources in adjacent categories while the remaining 5.2 percent differ by 2 or more categories.

Table 2: Distribution of Hospital-based EDs by HCUP and AHA ED Volume Categories, 2007

		AHA ED Volume Category (1000s)					
		< 10	10 – 19	20 – 29	30 – 39	40 – 49	50+
HCUP ED Volume Category (1000s)	< 10	637	57	9	10	5	3
	10 – 19	63	342	36	6	2	8
	20 – 29	16	36	260	42	6	9
	30 – 39	8	6	49	192	34	14
	40 – 49	3	7	12	24	122	46
	50+	1	5	11	6	33	355

3.2 State-level Count Comparisons

Table 3 summarizes State-level differences in relative ED counts between the 2 sources. Among the 27 States, the ratio of HCUP to AHA ED counts ranged from 0.76 to 1.10 (i.e., from HCUP being 24% lower to 10% higher than the AHA estimate). State-level totals for 15 States were lower in HCUP than the AHA while the converse was true for the remaining 12 States. However, HCUP State-level estimates differed by less than 10 percent from the AHA estimate for 19 of the 27 States. It should also be noted that the number of hospital-based EDs on which State-level estimates are based varies widely across the 27 States in this analysis, ranging from only 10 in the State with the smallest number of institutions to 314 in the State with the most.

Table 3: Distribution of 27 States by Ratio of HCUP to AHA ED Totals, 2007

Ratio of HCUP to AHA ED Count	Number of States
.76	1
.82-.87	6
.91-.98	8
1.0-1.07	11
1.10	1

3.3 Logistic Regression Model 1 Results: *Variation in Likelihood of Agreement (10% criterion) by Hospital Characteristics*

Table 4 contains bivariate statistics on the proportion of observations with a value of 1 on the dependent variable for Model 1 (i.e., proportion where HCUP ED count differs from AHA count by more than 10%) as well as the estimated odds ratios (OR) for each independent variable in the logistic regression model. These ORs reflect the relative odds of an observation having a value of 1 versus a 0 on the dependent variable, controlling for the effects of the other independent variables in the model.

The significant predictors (at the .05 level) in this logistic regression model were State (not shown in table), hospital ownership, number of hospital beds, volume of ED visits (based on AHA data) and the urban hospital indicator. More specifically, controlling for other factors in the model, the following hospital characteristics were associated with greater odds of HCUP versus AHA differences of 10% or more: for-profit hospital status, higher hospital bed size, lower volume of ED visits, and urban hospital location.

3.4 Logistic Regression Model 2 Results: *Variation in Likelihood of HCUP<AHA by Hospital Characteristics*

Table 5 contains bivariate statistics on the proportion of observations with a value of 1 on the dependent variable for Model 2 (i.e., proportion where HCUP ED count is less than the AHA count) as well as the estimated odds ratios for each independent variable in the logistic regression model. These ORs reflect the relative odds of an observation being a 1 versus a 0 on the dependent variable, controlling for the effects of the other independent variables in the model.

The highly significant predictors in this logistic regression model were State (not shown in table), number of hospital beds, volume of ED visits (based on AHA data) and urban hospital indicator. In addition, the CAH indicator was also significant at the .05 level (p=.04). More specifically, controlling for other factors in the model, the following hospital characteristics were associated with greater odds of HCUP counts being lower than the AHA count: State (not shown), fewer hospital beds, more ED visits, non-urban hospital location, and classification as a CAH.

4. Summary/Discussion

This analysis compares ED visit counts in 2007 from HCUP and the AHA at the total, State, and institutional levels for the 27 States where ED-level visit counts are available from both data sources. Although there is a tendency for HCUP counts to be lower than AHA across most hospital characteristics, the aggregate estimates for the 27 States are relatively close: HCUP is 3.5% lower than the AHA. This finding corroborates the use of the AHA as a reasonable source for control totals for the NEDS, a sample of hospital-based EDs from HCUP States with both SID and SEDD data designed to yield national estimates of ED utilization (HCUP, 2010).

For the institutional-level comparison, we considered several characteristics that are commonly associated with differences in hospital behavior without clear *a priori* hypotheses about whether the HCUP and AHA counts would be similar (and the direction of the difference if they were not) for most of these characteristics. The analysis indicates that for-profit hospital status, larger hospital bed size, lower volume of ED encounters, and urban hospital location are associated with a greater odds of HCUP and AHA counts being more than 10% different. State effects were also significant. Further analysis showed that fewer hospital beds, more ED visits, rural hospital location, and CAH status were associated with greater odds of the HCUP count being lower than the AHA count. State effects were also significant. Our analysis did not support the hypothesis that HCUP ED counts may be higher than AHA counts for CAHs due to Medicare reporting requirements (see section 2.5 above). In fact, the model produced a result that was opposite of the hypothesized direction. Further research and sensitivity analysis is needed to better understand the reasons for these results.

This analysis was conducted on two of the data sources available for estimating the number of ED counts in the U.S. In general, HCUP and AHA produce fairly similar estimates of ED visits. However, the analysis illustrates that there are some hospital characteristics associated with the extent that ED counts for the two data sources diverge. In conclusion, researchers should consider the strengths and weaknesses of the resources available to them when analyzing ED visits, and they should select the data source that is most appropriate to address their particular research question.

5. References

Healthcare Cost and Utilization Project (2010). Introduction to the HCUP Nationwide Emergency Department Sample (NEDS), 2007. Rockville, MD: Agency for Healthcare Research and Quality. Available at http://www.hcup-us.ahrq.gov.

Health Forum LLC. (2008). *AHA Hospital Statistics*. Chicago, IL: Health Forum.

Medicare Learning Network. (2005). Billing Requirements for Physician Services Rendered in Method II Critical Access Hospitals (CAHs). *MLN Matters*. MLN Matters Number: MM3800. Related Change Request (CR) #: 3800. Implementation Date: July 5, 2005. Baltimore, MD: Centers for Medicare & Medicaid Services. Available at http://www.cms.gov/MLNMattersArticles/2005MMA.

Owens, P., Barrett, M., Gibson, T., Andrews, R., Weinick, R., and Mutter, R. (2010). Emergency Department Care in the United States: A Profile of National Data Sources. *Annals of Emergency Medicine* 56(2): 150-165.

Rosko, M. and Mutter, R. (2010). Inefficiency Differences between Critical Access Hospitals and Prospectively Paid Rural Hospitals. *Journal of Health Politics, Policy, and Law* 35(1): 95-126.

Table 4: Descriptive statistics and logistic model results (Model 1: Y variable indicates HCUP differs from AHA by >10%)

		Bivariate stats		Logistic regression model			
	n	Proportion differ by 10% or more	Std. Dev	OR	95% Wald Confidence Limits		P-Value for t-test of OR=1 (ref)
Overall	2,455	0.425	0.49				
Teaching Hospital							
No	2,000	0.440	0.50	0.93	0.70	1.24	0.6297
Yes	455	0.358	0.48	1.00			
Ownership							
Nonprofit	1,590	0.387	0.49	0.55	0.41	0.74	<.0001
For profit	306	0.513	0.50	1.00			
Government	559	0.483	0.50	0.66	0.47	0.92	0.0137
Number of hospital beds							
<100	1,027	0.479	0.50	1.00			
100-299	946	0.406	0.49	1.36	1.04	1.77	0.0252
300+	482	0.346	0.48	1.47	0.98	2.20	0.0622
Volume of ED visits (1000s)							
<10	728	0.602	0.49	6.50	4.08	10.34	<.0001
10-19	453	0.404	0.49	2.23	1.49	3.34	<.0001
20-29	377	0.385	0.49	1.59	1.09	2.31	0.016
30-39	280	0.300	0.46	1.08	0.74	1.59	0.6835
40-49	202	0.327	0.47	1.11	0.75	1.64	0.593
50+	415	0.306	0.46	1.00			

Table 4: Descriptive statistics and logistic model results (Model 1: Y variable indicates HCUP differs from AHA by >10%)

		Bivariate stats		Logistic regression model			
	n	Proportion differ by 10% or more	Std. Dev	OR	95% Wald Confidence Limits		P-Value for t-test of OR=1 (ref)
CAH hospital							
No	1,789	0.383	0.49	1.00			
Yes	666	0.536	0.50	0.89	0.65	1.22	0.4649
Urban hospital							
No	967	0.462	0.50	0.68	0.53	0.86	0.0015
Yes	1,488	0.401	0.49	1.00			

[1]Model included a set of 26 State dummy variables to allow for the effects of unmeasured State characteristics. Coefficients associated with the dummy variables not included in tables to simplify presentation.

Table 5: Descriptive statistics and logistic model results (Model 2: Y variable indicates HCUP < AHA)

		Bivariate stats		Logistic regression model			
	n	Proportion HCUP<AHA	Std. Dev	OR	95% Wald Confidence Limits		P-Value for t-test of OR=1
Overall	2,455	0.602	0.49				
Teaching Hospital							
No	2,000	0.596	0.49	1.18	0.88	1.59	0.2458
Yes	455	0.633	0.48	1.00			
Ownership							
Nonprofit	1,590	0.618	0.49	1.19	0.88	1.61	0.4125
For profit	306	0.533	0.50	1.00			
Government	559	0.596	0.49	1.18	0.84	1.66	0.5140
Number of hospital beds							
<100	1,027	0.581	0.49	1.00			
100-299	946	0.606	0.49	0.71	0.54	0.93	0.1499
300+	482	0.641	0.48	0.37	0.25	0.56	<.0001
Volume of ED visits (1000s)							
<10	728	0.503	0.50	0.06	0.04	0.10	<.0001
10-19	453	0.587	0.49	0.18	0.12	0.27	<.0001
20-29	377	0.560	0.50	0.22	0.15	0.32	0.0029
30-39	280	0.707	0.46	0.54	0.36	0.81	<.0001
40-49	202	0.668	0.47	0.56	0.38	0.85	<.0001
50+	415	0.730	0.44	1.00			

Table 5: Descriptive statistics and logistic model results (Model 2: Y variable indicates HCUP < AHA)

	Bivariate stats			Logistic regression model			
	n	Proportion HCUP<AHA	Std. Dev	OR	95% Wald Confidence Limits		P-Value for t-test of OR=1
CAH hospital							
No	1,789	0.612	0.49	1.00			
Yes	666	0.577	0.49	1.40	1.02	1.93	0.0358
Urban hospital							
No	967	0.614	0.49	1.47	1.15	1.88	0.0023
Yes	1,488	0.595	0.49	1.00			

[1]Model included a set of 26 State dummy variables to allow for the effects of unmeasured State characteristics. Coefficients associated with the dummy variables not included in tables to simplify presentation.