Emergency Department Care in California: Who Uses It and Why?
Technical Appendix

Shannon McConville
Helen Lee

Supported with funding from the California Program on Access to Care

Description

In this document, we describe the data sources used in our *California Counts* report and provide additional explanation, by section.

Contents

PPIC | PUBLIC POLICY INSTITUTE OF CALIFORNIA

I. Emergency Department Capacity and Use

The primary dataset used in this section is the Hospital Annual Utilization Data (HAUD) collected by the California Office of Statewide Health Planning and Development (OSHPD). The data are collected annually from all hospitals licensed in the state and include several measures related to emergency department (ED) services: total number of ED visits, ED visits that result in hospital admissions, number of ED treatment stations (beds), ED license status and level. We use HAUD data for all years between 1997 and 2006.

We use HAUD information for all general acute-care hospitals to assess trends in capacity and use. Every year, HAUD has non-responding hospitals.[1] Although we include these non-responding hospitals in our dataset to count the total number of operating EDs, the information in HAUD is very limited for non-responders and does not include total number of ED beds or ED visits.[2] To the extent possible, we impute missing values for ED visits and ED beds for these non-responding hospitals that had an ED in operation for the entire year and had non-missing values for ED visits and ED beds variables in the immediately previous and subsequent years. In the case of ED visits, we average the previous and subsequent years; for ED beds, we simply use the previous year's number of beds if that figure is available and if not, we use information from the subsequent year. In some cases, the hospital is missing data for several years in a row, even though other indicators show they are licensed and in operation for

[1] Missing hospital information in HAUD ranges from a low of 1 percent of hospitals missing data in 2006 to a high of 9 percent of hospitals missing data in 2002. For most of the other years, about 4 percent of hospitals are missing information. However, once we perform the imputations, the number of hospitals missing information in any given year ranges from less than 1 percent to 3 percent.

[2] Hospitals that were open and licensed throughout the year and operated an ED but were non-respondents are included in the aggregate ED count. HAUD provides information for non-respondents on presence and type of ED, but no information on ED beds and visits. Hospitals that were closed or had a suspended license during the year are not included. Hospitals that closed or were suspended during the year but submitted data are included in the ED visit counts.

those years, and these hospitals do not contribute to the aggregate ED visit and bed counts. Thus, our estimates of aggregate ED visits and ED beds are slight underestimates.[3]

II. Overcrowding Analysis

Again, we use HAUD as our primary data source to examine ED overcrowding. Starting in 2002, HAUD began collecting measures of overcrowding, including counting patients who register but leave before receiving a medical evaluation, and ambulance diversion hours. [4] We also use HAUD to determine other hospital characteristics, including hospital size (number of beds), ownership status, and occupancy rates. Occupancy rates are calculated as the percentage of patient census-days divided by licensed bed-days.

We do not impute the two overcrowding measures because these may vary more across years, nor do we use the data collected in 2002 because of many missing values. Instead, to develop a consistent and conservative estimate for our overcrowding measures, we average non-missing values across hospitals over four years (2003 – 2006) and only include hospitals that had valid measures for at least two of the four years.[5]

We also incorporate neighborhood characteristics based on the hospital's zip code from the 2000 Census SF3 data for zip code tabulation areas (ZCTA).[6] The Census ZCTA boundaries

[3] Comparing the total number of ED visits derived from HAUD in 2005 with the total number of ED visits derived from the ED encounter and patient discharge data suggests that visits are underestimated by about 2 percent, although this will fluctuate based on the total number of non-responding hospitals and those with missing data for ED visits.

[4] OSHPD data may underestimate the amount of ambulance diversion compared to data collected by Local Emergency Medical Service (LEMS) agencies, which are responsible for the regional coordination of ambulance transport to hospitals. A recent study on ambulance diversions in California reports diversion hours reported by LEMS agencies are about 10-15 percent higher than corresponding OSHPD diversion hours. (The Abaris Group, 2007.)

[5] Ninety-six percent of responding hospitals had valid measures for at least two of the four years. We do not use 2002 data in the overcrowding analysis because there is some concern about potential reporting errors and several missing values. In 2002, a new online system of reporting the HAUD was implemented by OSHPD and seems to have resulted in some data inconsistencies and a much greater number of non-responding hospitals.

[6] To assess whether hospital zip codes are a good proxy for the patient population of an ED, we used the 2005 ED encounter data, which provides patient-level zip code information. On average, 24 percent of ED visits were by patients who lived in the same zip code as the hospital.

are designed to be as consistent as possible with zip codes, and we matched these to ZCTAs to zip codes, using geographic information system software, ArcGIS 9, to append aggregate population characteristics. [7] The neighborhood zip code characteristics included are total population, percent of population below the federal poverty level, percent of the population 65 years of age or over, percent of the population that is foreign born, and percent of population that is foreign born and arrived in the United States within the previous five years. We also include counts of Medi-Cal beneficiaries at the zip code level, which are based on the Medi-Cal Beneficiary by Zip Code files for July 2005. These are updated on a quarterly basis using the Medi-Cal Eligibility Determination System (MEDS) data file and are publicly available online.

We only include large, urban counties in the overcrowding analysis: Alameda, Contra Costa, Fresno, Kern, Los Angeles, Orange, Riverside, Sacramento, San Bernardino, San Diego, San Francisco, San Joaquin, San Mateo, Santa Clara, Stanislaus and Ventura, which contain about 85 percent of the total state population and almost three-quarters of all hospitals with EDs. The vast majority of hospitals with overcrowded conditions based on the two measures we analyze are located in these 16 counties. Performing the overcrowding quartile analysis on all hospitals showed that 85 percent of hospitals in the top quartile of patients leaving without being seen were located in these 16 urban counties, as were 98 percent of hospitals in the top quartile of ambulance diversions. The regions discussed in Table 1 are defined by the county in which the hospital is located. The Los Angeles region is Los Angeles County; Bay Area includes Alameda, Contra Costa, San Francisco, San Mateo, and Santa Clara counties; Central Valley includes Fresno, Kern, San Joaquin, Sacramento, and Stanislaus counties; Other Southern California includes Orange, Riverside, San Bernardino, San Diego and Ventura counties.

[7] Although designed to be compatible, there is not a perfect one-to-one match between zip codes and ZCTAs. In addition, zip codes are updated and revised by the postal service over time and thus matching with Census 2000 data will result in some inconsistencies between zip codes in later years and those in the year 2000. We did perform some manual corrections of missing zip code information using online mapping services, but there are still some inconsistencies between ZCTA and zip code matches.

We also discuss differences in overcrowding measures based on whether hospitals receive funds from the Disproportionate Share Hospitals (DSH) program. [8] We use the OSHPD Hospital Annual Financial Data files from 2003 to 2006 to identify DSH hospitals. Hospitals that received DSH funds for at least two of the four years are included.

In addition to the bivariate results on overcrowding presented in Table 1, we also discuss results from a regression analysis of overcrowding. We performed ordinary least squares regression where the dependent variable was the percentage of patients leaving the ED without being seen by a physician, and included all 250 hospitals located in urban counties, rather than the 125 hospitals that fell into the top and bottom 25 percent of the distribution. Because of multicollinearity among some of the hospital and neighborhood characteristics, we do not include all variables displayed in Table 1 as independent variables in the regression analysis. The independent variables included are region, hospital ownership, hospital size, occupancy rate of medical-surgical beds, percent of ED visits that result in a hospital admission (ED admission rate), total population, poverty rate, percent foreign born, and percent over age 65. Full regression results are shown below.

[8] Disproportionate Share Hospital (DSH) funds are additional bulk payments made to hospitals by the Medicaid program and are administered in California through the SB855 program.

Overcrowding Regression Analysis, Urban Counties, 2003 - 2006

	Coefficients	Standard Errors	
Region (Los Angeles)			
Bay Area	0.0000	0.005	
Central Valley	0.0146	0.006	**
Other Southern California	-0.0063	0.005	
Hospital Ownership (Non-profit)			
Public	0.0502	0.007	***
For profit	0.0057	0.004	
District	-0.0044	0.006	
Hospital Size - per 100 beds	0.0040	0.001	***
Occupancy Rate, Medical-Surgical beds	0.0194	0.010	**
ED Admission Rate	-0.0172	0.030	
Neighborhood Characteristics			
Total Population - per 1,000 population	0.0001	0.000	
Poverty Rate	0.0712	0.024	***
Percent Foreign Born	-0.0065	0.017	
Percent Age 65 and over	-0.0301	0.034	
R-squared	0.4060		
Sample Size	248		

Notes: Reference category is provided in parentheses.
* p < 0.1, ** p < 0.05, *** p < 0.01

III. Non-Urgent Conditions Analysis

The other main dataset used in this report is the 2005 OSHPD Emergency Department encounter data, which includes a record for every ED visit that did not result in a hospital admission. We combine this dataset with information from the OSHPD Patient Discharge data on inpatient visits, where the source of admission was the hospital emergency department. We then construct a database of all ED visits that occurred in the state in 2005. These data include information on patient age, race/ethnicity, insurance status, county and zip code of residence, primary diagnosis, and procedures performed. Because some of this patient information is masked for patient confidentiality (e.g., race/ethnic background of the patient is masked in 20

percent of the cases), we examine only patient age and insurance status[9], because these characteristics are most complete. However, about 5 percent of visits have masked patient age and are not included in any of our analyses by age category.

Because the ED encounter data is visit-level, rather than patient-level, we cannot identify frequency of use by individual patients.[10] Instead, we calculate visit rates by various population groups that require population estimates for denominators. We use a variety of sources for these population estimates. California Department of Finance demographic data, 2005 is used for all age and county population estimates. We use the California Health Interview Survey (CHIS), 2005 for estimates of the population by insurance coverage. We also use Medi-Cal Beneficiary by Zip Code data to estimate the Medi-Cal population by county.

To examine the seriousness of ED patient health conditions, previous studies have examined the severity of ED visits in the state. But there are questions surrounding the reliability of the data they analyze because definitions of urgent and critical patients may not have been consistently applied (Bamezai and Melnick, 2006).[11] We employ an algorithm created by a team of health service researchers and ED and primary care physicians (Billings et al., 2003), which categorizes ED encounters that did not result in a hospital admission by principal diagnosis (ICD9 codes). The algorithm was developed based on detailed examination of 6,000

[9] Insurance status is derived from the expected payer source variable contained in the emergency department encounter data, which as the name suggests, indicates the expected, not actual, payer for the visit. Other studies have assessed the accuracy of this variable using OSHPD patient discharge data and found that it may undercount Medi-Cal admissions by assigning them to other coverage sources, most often Medicare in the case of patients with dual eligibility for both programs. Also, because Medi-Cal has retroactive eligibility, meaning that if an uninsured patient is determined eligible for Medi-Cal coverage after receiving treatment, that treatment will still be covered by the Medi-Cal program. Thus, some patients will become covered by Medi-Cal as a result of their visit. This may also result in uninsured patients being misclassified as covered by Medi-Cal and vice-versa. Another analysis found that about 5 percent of hospitalizations coded as Medi-Cal were among patients who gained Medi-Cal coverage as a result of the hospitalization and about 2 percent of hospitalizations coded as uninsured were covered by Medi-Cal (Chattopadhyay and Bindman, 2005).

[10] Other research indicates that there are a relatively small group of patients that are heavy users of EDs. Surprisingly, these heavy users tend to be insured and have a usual source of care, but are also in poor physical and mental health (Hunt et al., 2006)

[11] These studies have used HAUD, which includes hospital-level ED visit counts by severity based on Current Procedural Terminology (CPT) codes, not information on individual ED encounters. We also analyzed these variables, but found large inconsistencies across years that suggested they may not be consistently reported.

full ED records that included initial complaint, presenting symptoms, vital signs, medical history, patient demographics, diagnoses, procedures, and resources used. For every visit, the algorithm assigns a percentage of the visit to one of four categories of urgency, or else flags it as an injury, psychiatric condition, or drug- and alcohol-related.[12] Based on this algorithm and ED and inpatient data, we classify all ED visits into six main types:

1) visits that result in a hospital admission

2) non-avoidable visits that require ED care, such as appendicitis or heart attack

3) avoidable visits, which includes visits that require ED care but could have been avoided with timely and effective primary care (e.g., asthma); visits that required care within 12 hours, but could have been treated in primary care; and visits that are non-urgent and did not require care within 12 hours

4) injury-related visits

5) psychiatric, drug-, or alcohol-related visits

6) unclassified visits

IV. Characteristics of Emergency Department Users Analysis

Finally, we use 2005 data from the California Health Interview Survey (CHIS) to examine predictors of recent ED use among a sample of adults in the state. The CHIS is a representative, statewide telephone survey of more than 40,000 households conducted by the UCLA Center for Health Policy Research that covers many health-related topics, including use of physician services and use of emergency departments. For more information on CHIS, we refer the reader to the survey's website: http://www.chis.ucla.edu/.

We use the CHIS dataset for adults and include only non-elderly adults age 18 to 64 in the analytic sample. We exclude from our analysis non-elderly adults who have Medicare

[12] For example, all visits for urinary tract infections (ICD9 599.0) are assigned the following percentages: 66 percent "non-urgent," 17 percent "emergent/primary care treatable," and 17 percent "emergent - ED care needed - preventable/avoidable."

coverage (they are primarily permanently disabled), and other cases with missing values. The final analytic sample includes 31,224 adults.

We run logistic regression analyses where the dependent variable is whether the respondent had at least one emergency department visit in the previous year. Results are weighted using replicate weights and jackknife variance estimation in STATA 10.0 (CHIS, 2006) and have been adjusted for age, sex, race/ethnicity, immigrant status, marital status, education, poverty level, employment, self-reported health status, indicators for a usual source of care, the presence of serious health conditions (such as asthma, diabetes, or hypertension), number of doctor visits in past year, and urbanization of residence. Regression coefficients for select variables are included in the table below. For complete results, please contact the authors.

We present the results of this analysis in Figures 8 and 9. The likelihood of an emergency department visit presented in both of these figures is calculated as the predicted probability for the group of interest, holding all other explanatory variables at their means.

Logistic Regression Analysis of Recent ED Visit, Nonelderly Adults, 2005

	Coefficients	Standard Errors	
Race/Ethnicity (NH White, US Born)			
NH White, Foreign Born	0.0362	0.118	
Black, US Born	0.1512	0.107	
Hispanic, US Born	-0.0848	0.077	
Hispanic, Naturalized Citizen	-0.2367	0.120	**
Hispanic, Non-Citizen	-0.2442	0.113	**
Asian, US Born	-0.5344	0.196	***
Asian, Naturalized Citizen	-0.6302	0.138	***
Asian, Non-Citizen	-0.6674	0.165	***
All other [1]	0.1997	0.114	*
Insurance (Privately insured)			
Uninsured	-0.0281	0.084	
Medi-Cal	0.4830	0.094	***
Usual source of care other than ER (Yes)	0.1315	0.095	
Self reported health (Excellent/Very Good)			
Good	0.1318	0.058	**
Fair	0.4801	0.074	***
Poor	1.0868	0.115	***
# doctor visits in past year (3+ visits)			
None	-2.1851	0.118	***
1 visit	-1.0048	0.070	***
2 visits	-0.6376	0.067	***
Serious health conditions (None)			
Asthma	0.4609	0.075	***
Diabetes, high blood pressure, or heart disease	0.1287	0.061	**
Sample Size	31,224		

[1] All other category includes foreign-born Blacks and Native Hawaiian/Pacific Islanders, American Indian/Alaskan Native, and Multi-Race groups regardless of nativity and citizenship status.

Notes: Reference category is provided in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

This model also controls for age, sex, marital status, education, poverty level, employment, and urbanization of residence.

References

The Abaris Group, "California Emergency Department Diversion Project: Report One," Oakland, California, 2007.

Bamazei, Anil and Glenn Melnick, "Marginal Cost of Emergency Department Outpatient Visits: An Update Using California Data," Medical Care, Vol. 44, No. 9, 2006, pp. 835-841.

Chattopadhyay, Arpita and Andrew B. Bindman, "Accuracy of Medicaid Payer Coding in Hospital Patient Discharge Data: Implications for Medicaid Policy Evaluation," *Medical Care*, Vol. 43, No. 6, 2005, pp. 586-591.

Hunt, Kelly A., Ellen J. Weber, Jonathan A. Showstack, David C. Colby, and Michael L. Callaham, "Characteristics of Frequent Users of Emergency Departments," Annals of Emergency Medicine, Vol. 48, No. 1, 2006, pp. 1 -8.

UCLA Center for Health Policy and Research, "Weighting and Estimation of Variance in the CHIS Public Use Files," Los Angeles, California, 2006. Accessed at http://www.chis.ucla.edu/pdf/weighting_var_chis_010106.pdf on June 28, 2008.