Emergency departments (EDs) are an important source of care for a large segment of the population of the United States. There are more than 131 million visits to the ED each year, and two-thirds of hospital admissions are via the ED.1,2 In spite of its importance, the ED is widely considered by patients and providers as a place of inefficient, unsatisfactory, and often unsafe health care encounters.3,4 Almost 40% of patients wait more than an hour to see a physician, and a quarter of the patients seen in the ED spend more than four hours there.1 Waits and delays lead to inefficient patient flow and crowding, factors linked to adverse clinical outcomes.5–11 This ED milieu also negatively affects access to care and is associated with 2% of patients (2.6 million persons annually) leaving the ED without being seen by a physician.1

Measurement and monitoring of emergency department performance has been prompted by The Joint Commission’s patient flow standards,12,13 and the Centers for Medicare and Medicaid Services (CMS) performance metrics.14,15 Data involving length of stay (LOS), left without being seen (LWBS), and door-to-physician times are currently being reported to CMS and will be factored into Value-Based Purchasing payment reform for the first time in 2013 and 2014.16

However, in contrast to typical evidence-based clinical benchmarks, comparison strategies for ED operational performance have been harder to establish because of substantial differences between EDs in terms of volume and acuity.

EDs exhibit and differ across an array of operating characteristics including those which reflect acuity (such as admission rate, high Current Procedural Terminology [CPT] acuity rate, and pediatric patient percentage). Leaders and managers of EDs are concerned that comparisons made between EDs do not take into account such differences. To address this issue, stakeholders from the emergency medicine and quality improvement communities have suggested comparison schemes developed by expert consensus, although data to support such schemes have been lacking.17–21 Performance capability of large groups of EDs have been...
tracked and the EDs characterized across many parameters.\textsuperscript{22–27} A small Canadian study published in 2010 correlated performance with rural versus urban location.\textsuperscript{28} Another recent study correlated LOS with safety-net status.\textsuperscript{29} We are not aware of a previous study that correlated performance on ED metrics with both annual volume and the other operating characteristics presented in this study.

The Emergency Department Benchmarking Alliance (EDBA), a collaborative of EDs located in the United States,\textsuperscript{30} has been collecting data from its membership involving operating characteristics, and performance since 2004. In 2009 its board of directors observed that simple scatterplots suggested trends regarding the performance of EDs on the basis of volume and operating characteristics (Appendix 1, available in online article). In view of these preliminary survey findings, we hypothesized that performance is volume dependent and influenced by acuity. Fair performance comparisons of EDs, policies, and even optimum design standards could be developed on the basis of the data.

**Methods**

**STUDY DESIGN AND DATA COLLECTION**

We performed a retrospective analysis of EDBA annual ED survey data for the most recent year for which data were available (2009) to explore observed patterns in ED performance relative to size and operating characteristics. The survey was based on 14.6 million ED visits in 358 hospitals of varying sizes across the United States, with an ED size representation (sampling) approximating that of the Emergency Medicine Network (EMNet; also known as the National ED Inventory, a national ED survey database begun by researchers at Mass General Hospital in 2003 to catalog comprehensively all EDs).\textsuperscript{31} The data were aggregated, with the unit of analysis being a summary over a large population of cases within each individual ED. This work was deemed Institutional Review Board (IRB) exempt by the Intermountain Healthcare (Salt Lake City) Institutional Research Board.

EDBA, a not-for-profit organization, has conducted comprehensive annual data surveys since 2004. Participation in the survey is a mandatory requirement for membership. Table 1 (above) displays a breakdown of the participating EDs by annual volume and as a percentage of the total membership. Although the breakdown of EDs into volume bands of 20,000 annual visits, like the other survey data, is based on self-report, EDBA has used this breakdown from the first surveys because it is thought by members to make sense operationally and functionally. In 2009 the median ED annual volume hovered at 40,000 visits, offering a natural breakpoint for the five EDBA volume bands.

**MEASURES**

Table 2 (page 397) shows a summary of the comparison measures commonly in use in emergency medicine and broken into the structure-process-outcome paradigm that is being used increasingly as a framework for such comparative analyses.\textsuperscript{32,33} These measures include operating characteristics, time, and proportion performance measures. Some operating characteristics create a profile of an ED that suggest the acuity of care provided. Process and outcome measures of ED performance are listed in Table 2 and include continuous-variable measures such as LOS and door-to-physician time, typically reported as medians. Proportion measures such as “walkaway” categories: Against medical advice (AMA), LWBS, and left before treatment complete (LBTC), are also listed. EDBA previously standardized all data elements through a summit of expert opinion and the publication of corresponding data definitions.\textsuperscript{34} These definitions were updated at a second consensus summit in 2010.\textsuperscript{35,36}

**Operating Characteristics.** Operating characteristics such as patients per day, high CPT acuity rate, pediatric patient percentage, admission rate, percentage admitted to the hospital, transfer rate, emergency medical services (EMS) arrivals, EMS arrival admission rates, and visits per care space have been defined in

<table>
<thead>
<tr>
<th>ED Annual Volume</th>
<th>Frequency</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 20,000 visits</td>
<td>63</td>
<td>17.8</td>
</tr>
<tr>
<td>20,000–39,999</td>
<td>146</td>
<td>40.7</td>
</tr>
<tr>
<td>40,000–59,999</td>
<td>92</td>
<td>25.6</td>
</tr>
<tr>
<td>60,000–79,999</td>
<td>35</td>
<td>9.7</td>
</tr>
<tr>
<td>80,000+</td>
<td>22</td>
<td>6.1</td>
</tr>
<tr>
<td>Total</td>
<td>358</td>
<td>100.0</td>
</tr>
</tbody>
</table>

* EDBA, Emergency Department Benchmarking Alliance.
two new consensus documents.\textsuperscript{35,36} High CPT acuity is defined as patients with CPT codes 99284 and 99285 for the ED visit. These codes are applied to the sickest and most complicated patients, and the survey identifies the percentage of ED visits assigned these codes.

**Utilization Rates.** Triggered by the National Quality Forum endorsement of CMS plans to begin tracking ED imaging utilization in its outpatient prospective payment system (OPPS) measures,\textsuperscript{37} some utilization rates are also reported. On the basis of the specified definitions,\textsuperscript{35,36} survey participants standardized utilization rates as procedures per 100 ED visits. Rates of electrocardiography (ECG), plain radiology (x-ray studies), and computerized tomography/magnetic resonance imaging (computerized tomography/magnetic resonance imaging [CT/MRI]) were reported for each volume band.

**Performance Metrics.** We used the following two primary outcome measures of performance:

- **LOS:** Time interval between arrival time and discharge time. This term indicates the number of minutes for all patients that spent time in the ED and is tracked as the median time. It can be stratified by subsets of patients as well (for example, admitted, discharged, observed).

- **LBTC:** Percentage of patients identified as seeking care but leaving before completion of treatment. This measure is reported as a percentage of overall ED visits and for EDBA has been a global measure of patients leaving before they received a disposition by the provider. The EDBA definition includes LWBS and left AMA.

We used one secondary outcome measure of performance, as follows:

- **Door to physician time:** The median time interval in minutes between arrival time and the time of first contact with the physician or provider.

**DATA ANALYSIS**

We used descriptive statistics to summarize operating characteristics, utilization rates, and performance metrics across the five EDBA volume bands. The data were an aggregate of all ED patients within each individual ED over the study period. Because of the large number of cases within each ED and using the Central Limited Theorem, the observed ED operating characteristics, utilization rates, and performance metrics were assumed to be approximately normally distributed. Analysis of Variance (ANOVA) was used to compare the differences in operating characteristics, utilization rates, and performance metrics across the volume bands. Linear regression analysis (using linear, quadratic, and Loess smooth line fit) was used to evaluate the correlation of ED LOS with census category. Univariate logistic regression analysis was used to assess the association between LBTC and door-to-physician time at different volumes. Odds ratios (OR) with corresponding 95% confidence intervals (CIs) were calculated and reported. All statistical analyses were performed using R (R Foundation for Statistical Computing, Vienna). Statistical significance level, $\alpha$, was defined at .05.

**Results**

Analyses of the five EDBA volume bands are presented in Table 3 (page 398). EDs were stratified by annual volume and placed in categories of 20,000-volume increments. Differences between these bands were readily apparent in operating characteristics, utilization rates, and performance on metrics.

<table>
<thead>
<tr>
<th>Type</th>
<th>Measure</th>
<th>Metric</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structure (Operating</td>
<td>Annual Volume</td>
<td>Count</td>
</tr>
<tr>
<td>Characteristics)</td>
<td>High CPT Acuity</td>
<td>Rate</td>
</tr>
<tr>
<td></td>
<td>Pediatric Patient</td>
<td>Rate</td>
</tr>
<tr>
<td></td>
<td>Admission rate</td>
<td>Rate</td>
</tr>
<tr>
<td></td>
<td>Transfer rate</td>
<td>Rate</td>
</tr>
<tr>
<td></td>
<td>EMS Arrival rate</td>
<td>Rate</td>
</tr>
<tr>
<td></td>
<td>Patients per care space</td>
<td>Count</td>
</tr>
<tr>
<td>Process (Time)</td>
<td>LOS</td>
<td>Median</td>
</tr>
<tr>
<td></td>
<td>Door-to-Physician Time</td>
<td>Median</td>
</tr>
<tr>
<td>Outcome (Proportion)</td>
<td>AMA</td>
<td>Proportion</td>
</tr>
<tr>
<td></td>
<td>LWBS</td>
<td>Proportion</td>
</tr>
<tr>
<td></td>
<td>LBTC</td>
<td>Proportion</td>
</tr>
</tbody>
</table>

* ED, emergency department; CPT, Current Procedural Terminology; EMS, emergency medical services; LOS, length of stay; AMA, against medical advice; LWBS, left without being seen; LBTC, left before treatment completed.
Compared with the lower volume departments, higher-volume EDs had higher admission rates ($p < .001$), higher EMS arrival rates ($p < .0001$) and lower transfer rates ($p < 0.015$). Other noteworthy findings are apparent in Table 3. Lower-volume EDs had the highest percentage of transfers to other facilities. Transfer rate varied by volume ($p = .0151$), and transfer rates correlated inversely with admission rates. The lower-volume EDs also had the distinction of claiming more inpatient admissions as having originated in the ED, although not at a statistically significant level. The proportion of high-CPT-acuity patients and the pediatric population percentage were not significantly different across the volume bands. Finally, the data indicated about 1,600 visits per ED treatment space occurred, with the highest patient/treatment space ratios in the moderate volume category.

## Utilization Rates

The higher-volume EDs also had higher utilization rates for EKG ($p = 0.011$) and CT/MRI ($p = .037$) than did the lower-volume EDs. There was no difference in x-ray study utilization between the higher- and lower-volume EDs.

## Performance Metrics

Finally, the higher-volume EDs had longer LOSs ($p < .001$), higher LBTC rates ($p < .0001$) and longer door-to-physician times ($p < .0012$) compared with the lower volume EDs.

Table 4 (page 399) and Table 5 (page 400) summarize the statistical analyses for the primary outcome measures. LOS and LBTC varied by volume and characteristics indicative of acuity (admission rate, high CPT rate, and pediatric patient percentage). After controlling for pediatric percentage, generalized linear regression analysis showed that the volume and acuity markers (admission rate and acuity rate by CPT code) were in-
Table 4. Linear Regression for Median Emergency Department Length of Stay by Volume and Acuity Markers*

<table>
<thead>
<tr>
<th>Volume</th>
<th>Mean Median LOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 20K</td>
<td>129.3</td>
</tr>
<tr>
<td>20K to 40K</td>
<td>159.5</td>
</tr>
<tr>
<td>40K to 60K</td>
<td>183.3</td>
</tr>
<tr>
<td>60K to 80K</td>
<td>193.4</td>
</tr>
<tr>
<td>&gt; 80K</td>
<td>219.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Acuity Markers</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ped %</td>
<td>.6898</td>
</tr>
<tr>
<td>Admit %</td>
<td>&lt; .0001</td>
</tr>
<tr>
<td>High CPT Acuity</td>
<td>&lt; .0001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Regression Terms</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear Only</td>
<td>&lt; .0001</td>
</tr>
<tr>
<td>Quadratic versus Linear only</td>
<td>.0298</td>
</tr>
<tr>
<td>Cubic, Quartic, Quintic versus Quadratic and Linear</td>
<td>.7876</td>
</tr>
</tbody>
</table>

* The mean and standard deviation for the EDBA volume bands and acuity markers are included in Table 3. LOS, length of stay; Ped, pediatric patient percentage; CPT, Clinical Procedural Terminology.

dependently associated with ED LOS. Table 4 shows the mean of the median LOS for each of the EDBA volume bands, as well as the $p$ values for comparing regression models using (ungrouped) volume to predict median LOS. The LOS increased as the ED volume increased ($p < .0001$).

The median LOS in large departments with more than 80,000 ED visits per year was 90 minutes longer than in departments with < 20,000 ED visits per year (219 minutes versus 129 minutes). Higher CPT acuity rate ($p < .0001$) and higher admission rate ($p < .0001$) also independently increased ED LOS. Higher pediatric patient percentage, widely considered a marker for lower acuity, was associated with decreased LOS but was not independently correlated with LOS and was not statistically significant (Table 4). High CPT acuity rate had a greater impact on LOS than either admission rate or pediatric patient percentage.

In both univariate and multivariate logistic regression analyses (Table 5), the likelihood of LBTC increased as the ED volume increased. Univariate analysis showed that patients arriving at a greater than –80K ED were more than twice as likely to LBTC than those arriving at a less than –20K ED (OR = 2.51, 95% CI 1.64, 3.83). In addition, EDs with longer average door-to-physician times (aggregate data) were associated with higher proportions of LBTC patients. A one-minute increase in average door-to-physician time increased the likelihood that a patient would LBTC by a factor of almost 4 (OR 3.62, 95% CI 3.04,4.27)

### Discussion

We have shown a clear relationship between ED volume and ED performance as measured by LOS and LBTC. We have also identified an array of operating characteristics of EDs that vary on the basis of census. We also found that those operating characteristics indicative of acuity correlate with performance. Heretofore these relationships have not been established.

Two previous studies stemming from the CDC National Hospital Ambulatory Medical Care Survey (NHAMCS) database examined utilization and some performance data but did not stratify the data by ED volume. A third, recent study, also based on NHAMCS data, correlated ED volume and other exogenous factors (such as case mix, age mix, teaching status and metropolitan statistical area status) with performance on four National Quality Forum–approved ED performance measures. The Pines et al team, grouping EDs in 20,000-visit volume bands, found that the higher-volume EDs performed poorly relative to the lower-volume EDs on time metrics and LWBS. These findings are consistent with the results for our study. The median LOS for a < 20,000 visits per year ED was 129 minutes, with a LBTC approaching 1% (Table 3). In contrast, EDs seeing more than 80,000 patients a year had a median LOS of 219 minutes and a LBTC rate approaching 4%.

By analyzing other operating characteristics captured in our data survey, a picture of the differences between lower- and higher-volume EDs begins to emerge. The smaller EDs typically featured the lowest clinical intensity, as reflected in their reported operating characteristics. They had lower CPT acuity, lower admission rates, and higher transfer rates (off-loading the sickest patients to tertiary care centers). These EDs also moved fewer patients through their treatment spaces per year, probably because they were typically not operating over capacity. Performance on time and proportion measures was best in these lower-volume EDs, suggesting an optimum size for maximal efficiency. These findings agree with the Canadian study, which demonstrated the efficiency and performance capability of rural EDs serving lower patient volumes.

In addition, we have identified operating characteristics indicative of acuity (such as admission rate, high CPT acuity rate and pediatric patient percentage) that correlate with performance. Although research in pediatric ED care is limited, our findings relative to pediatric patients dovetail with previously reported studies. It is well known that pediatric EDs have less boarding of admitted patients, lower patient acuity, and lower admission rates. It is likely that pediatric patients are moved more quickly through the ED visit, resulting in improved performance metrics. Similarly, it has long been understood that
higher-acuity adult patients use more diagnostic resources, have higher admission rates, and longer LOSs. Not surprisingly, then, higher acuity in our study, as particularly reflected in admission rates, were associated with a worsening of ED performance. Pines et al correlated acuity using different markers (case mix index and patient age) with performance, and the findings were analogous. Taken together, these two studies make a strong argument that ED patient acuity affects ED performance.

We recognize that explaining the different operating characteristics across the various ED volume bands may depend on demographics. Hospitals with the largest EDs are typically located in areas with dense populations. As those hospitals increase in overall patient volume, they tend to house more services and provide tertiary care for the region and thus become the repository for the sickest and most complex patients. Pines et al also noted that hospitals in metropolitan service areas underperformed relative to those not located in such areas. Meanwhile, the hospitals with lower-volume EDs, which are typically situated in locations with lower population densities, will transfer the sickest patients to the larger tertiary care hospitals. Thus, the differences in operating characteristics by volume should make sense in terms of geographic location and the services being provided to the community. The impact of these factors should be further studied in subsequent analyses.

These findings may have implications for ED design and operations. Improving the efficiency of an ED may depend on its ability to reduce the volume of a larger ED by creating a subset of smaller EDs or clinical Microsystems. Some EDs have been redesigned in terms of pods or work zones, in which portions of the ED volume are segmented to create smaller functional units,41,42 concepts borrowed from operations research.43 Smaller work zones quite simply have fewer patients, information, diagnostic procedures and therapeutic interventions to manage in one microsystem.

In addition, the overall low transfer rates observed in the survey suggest that patients could be treated in a timely fashion in the smaller community hospital with a low risk of being transferred to a tertiary care center. This information could serve to “load level” in the community, that is, to distribute patients more evenly around a region by safely increasing the utilization of lower-volume EDs. This would result in off-loading the tertiary care centers, which are often operating in over-capacity conditions.44

Our findings have important implications for health care policy and administration. Beginning in October 2010 CMS began tracking performance data for EDs, including clinical and operational measures. On the basis of recommendations from the National Quality Forum, LOS, LWBS, and door-to-physician time will be factored into the Value Based Purchasing Plan.16 If all EDs do not share the same performance norms and if an ED’s volume and patient acuity are critical determinants of performance capability, as both our study and the Pines et al38 study suggest, then such findings should inform how administrators and policy makers compare hospital performance.

Further, while one study found no difference in LOS performance between safety-net and non-safety net hospitals,29 our findings suggest that the high-volume, urban safety-net hospitals will be particularly vulnerable in the currently evolving Pay for Performance scheme. In agreement with Pines et al, comparing the totality of emergency departments without stratification by volume and adjustments for other factors, particularly those reflecting acuity, is neither fair nor meaningful. It would serve to mislead the public trying to identify institutions delivering “quality care” and penalize larger institutions serving the highest-acuity patients. Caution is warranted to avoid poorly informed comparisons of ED performance across the range of ED volumes.

**LIMITATIONS**

Our study is not population based. The CDC NHAMCS database has informed a number of studies that have been extrapolated to reflect the entire population of EDs in the United States, and this study does not represent an attempt to duplicate...
that work. We also acknowledge that the data used in our analysis are self-reported by the various EDs and that the method of collection varied from facility to facility. Larger EDs and EDs that are part of a larger hospital system were more likely to have information technology support for data collection, while many smaller EDs collected data by hand and from patient logs; some of the smaller sites were not capable of capturing all data elements. Although the definitions of the data elements used in our study have been standardized by EDBA, individual sites were not audited to gauge compliance with these definitions. We note that the ED is a complex system, and there are confounding variables that are also likely to influence performance, such as rural versus urban, teaching versus nonteaching, case mix, payor mix, Emergency Severity Index breakdown, age of patient population. Another limitation in our work is that membership in EDBA is voluntary and data reporting is mandatory so that the data set may be skewed toward better performance. Finally, given that EDBA represents a unique set of EDs, we cannot go so far as to claim that our figures are the benchmarks for the specialty.

Conclusion

Higher-volume EDs have longer patient LOSs and a higher percentage of patients LBTC, and these performance parameters are influenced by patient acuity. Performance measurement schemes should take ED volume, acuity, and other characteristics into account, given the implications for health care administrators, policy makers, and patients.

References

2. Owens P, Elixhauser A. Hospital admissions that began in the ED. AHRQ HCUP (Healthcare Cost and Utilization Project), Statistical Brief #1, Feb. 2006.

Online-Only Content

See the online version of this article for Appendix 1. Length of Stay as a Function of ED Volume and Acuity Markers

Shari Welch, MD, is an Emergency Physicians, Utah Emergency Physicians, Salt Lake City; a Fellow, Institute for Health Care Delivery Research, Intermountain Healthcare, Salt Lake City; and a Board Member, Emergency Department Benchmarking Alliance, Dayton, Ohio. James J. Augustine, MD, is Director of Clinical Operations and Vice President, Emergency Department Benchmarking Alliance. Gregory Snow, PhD, is a Statistician, Statistical Data Center, Intermountain Healthcare. Lucy A. Savitz, PhD, MBA, is Director of Research and Education, Institute for Health Care Delivery Research. Li Dong, MD, MS, is a Statistician and Senior Outcome Analyst, Intermountain Healthcare. Brent C. James, MD, Mstat, is Chief Quality Officer and Executive Director, Institute for Health Care Delivery Research. Please address correspondence to Shari J. Welch, sharijwelch@gmail.com.
Appendix 1. Length of Stay as a Function of Emergency Department Volume and Operating Characteristics