CLINICAL PRACTICE

Estimating the Degree of Emergency Department Overcrowding in Academic Medical Centers: Results of the National ED Overcrowding Study (NEDOCS)

Steven J. Weiss, MD, Robert Derlet, MD, Jeanine Arndahl, MD, Amy A. Ernst, MD, John Richards, MD, Madonna Fernández-Frankelton, MD, Robert Schwab, MD, Thomas O. Stair, MD, Peter Vicellio, MD, David Levy, MD, Mark Brautigan, MD, Ashira Johnson, MD, Todd G. Nick, PhD

Abstract

Objectives: No single universal definition of emergency department (ED) overcrowding exists. The authors hypothesize that a previously developed site-sampling form for academic ED overcrowding is a valid model to quantify overcrowding in academic institutions and can be used to develop a validated short form that correlates with overcrowding. **Methods:** A 23-question site-sampling form was designed based on input from academic physicians at eight medical schools representative of academic EDs nationwide. A total of 336 site-samplings at eight academic medical centers were conducted at 42 computer-generated random times over a three-week period by independent observers at each site. These sampling times ranged from very slow to severely overcrowded. The outcome variable was the degree of overcrowding as assessed by the charge nurse and ED physicians. The full model consisted of objective data that were obtained by counting the number of patients, determining patients' waiting times, and obtaining information from registration, triage, and ancillary services. Specific objective data were indexed to site-specific demographics. The outcome and objective data were compared using a multiple linear regression to determine predictive validity of the full model. A five-question reduced model was calculated using a backward stepdown procedure. Predictive validity and relationships between the outcome and objective data were assessed using a mixed-effects linear regression model, treating center as random effect. **Results:** Overcrowding occurred 12% to 73% of the time (mean, 35%), with two hospitals being overcrowded more than 50% of the time. Comparison of objective and outcome data resulted in an R^2 of 0.49 ($p < 0.001$), indicating a good degree of predictive validity. A reduced five-question model predicted the full model with 88% accuracy. **Conclusions:** Overcrowding varied widely between academic centers during the study period. Results of a five-question reduced model are valid and accurate in predicting the degree of overcrowding in academic centers. **Key words:** overcrowding; ED management. ACADEMIC EMERGENCY MEDICINE 2004; 11:38–50.

Emergency departments (EDs) provide an important public service by providing emergency care 24 hours a day, 365 days per year, without discrimination by social or economic status. One of the key expectations of EDs is the ability to provide immediate access and stabilization for those patients who have an emergency medical condition.[1–7] Overcrowding diminishes the capability of the ED to manage these emergencies effectively.

Over 20 years ago, overcrowding in EDs was described in some metropolitan academic centers. Throughout the 1980s and 1990s, a number of articles in the lay press and academic journals documented the problems related to providing adequate or even basic care to patients.[5–13] In 1990, *Time* magazine focused on overcrowded EDs in a detailed cover story.[14] Few of the issues have improved over the past decade, illustrated by one of *Time* magazine's main headlines in 2000 that once again focused on the problem of overcrowding.[15] During 2001, a series of newspaper, news journals, and academic articles have attempted to describe the problem of overcrowding in American EDs (as an example, *USA Today* April 2002, "Health and Science Section": "Overcrowding in Emergency Rooms Common"). Although news media interest has been cyclic, the problem in many areas has been persistent and chronic.

From the University of California Davis Medical Center, Sacramento, CA (SJW, RD, JA, AAE, JR); Harbor/UCLA, Los Angeles, CA (MF); Truman Medical Center, Kansas City, MO (RS); Brigham and Women's Hospital, Boston, MA (TOS); SUNY Stonybrook, Stony Brook, NY (PV); Temple University, Philadelphia, PA (DL); Sinai-Grace Hospital, Detroit, MI (MB); Jackson Memorial Hospital, Miami, FL (AJ); and the University of Mississippi Medical Center, Jackson, MS (TGN).

Received March 31, 2003; revisions received July 10 and 22, 2003; accepted July 23, 2003.

Address for correspondence and reprints: Steven Weiss, MD, Division of Emergency Medicine, 2315 Stockton Boulevard, PSSB 2100, Sacramento, CA 95817. Fax: 916-734-7950; e-mail: sweiss52@aol.com.

Presented at the SAEM annual meeting, St. Louis, MO, May 2002.
doi:10.1197/S1069-6563(03)00583-9

Overcrowding discussions lack a standardized scale or definition. Because no criterion standard defining "overcrowding" exists, no two hospitals are talking about the same entity when they discuss overcrowding. We set out to develop a simple tool for objectively assessing the degree of overcrowding in a hospital ED. In this way, EDs can respond to administrative questions about the degree of ED overcrowding in a standardized manner. We chose academic centers with high volumes because they are affected by overcrowding, they generally supply the inner city indigent populations, and they are often described as overcrowded.[16] The specific institutions taking part in the study were selected to represent all areas of the country and all programs types.

The purpose of the National ED Overcrowding Study (NEDOCS) was to develop a simple screening tool that can be used easily and quickly to determine the degree of ED overcrowding at an academic institution. The hypothesis of this study was that a previously developed 23-question site-sampling form for academic ED overcrowding is a valid model to quantify the degree of overcrowding at academic institutions and can be used to develop a validated simple short form that correlates with the degree of overcrowding.

METHODS

Study Design. This was a survey study designed to validate a model to predict ED overcrowding. The institutional review board at each facility approved the study.

Study Setting and Population. This study was divided into two phases and conducted at eight sites. The eight sites represent a convenience sample of institutions committed to taking part in the entire study sample. Because pediatric EDs have their own characteristics and patient volume–time distributions, they were excluded from this study unless they were inseparable from the adult ED.

Survey Development. During the first phase, site investigators completed a site information form that included facts and opinions on their facility's state of overcrowding. Then, at random sampling times, each site investigator completed a sampling form. This form collected objective data representing a snapshot of the ED at that time and subjective data on the attending physicians' and charge nurse's opinion of the degree of overcrowding. From these data we developed a composite outcome variable that represented the degree of overcrowding as assessed by the charge nurse and the ED physicians. Next, a full model was built using objective data to reflect the outcome variable. Finally, through appropriate statistical techniques, a reduced model was created that reflected the full model, and its validity was tested using a bootstrap technique.

Site-information Form. During the first phase of the study, the primary site investigator was asked to complete a survey requesting site-specific information about the total yearly numbers of patients in the ED and the hospital. This was primarily an opinion survey of the primary investigator and was not connected to the data collection portion of study. ED and hospital beds actively in use were also tallied. Questions were asked about how each hospital handled the problem of overcrowding, such as patients in hallways, overcrowding and diversion policies, and patient flow to the floors. This information was used to characterize the sites and was compared with academic emergency medical departments nationally, based on information available on the Society for Academic Emergency Medicine (SAEM) website.

Site-sampling Form. Based on the literature currently available on ED overcrowding,[2,6,13,17–19] we developed four subscales. Questions on each subscale were required to meet the following criteria: 1) reflect various components of ED patient management (i.e., triage, treatment, and disposition), 2) readily available and easily quantifiable, 3) reproducible between observers, 4) represent a snapshot of the ED, and 5) based on consistent definitions at all institutions. The core investigators developed subscales with input from all sites. The following four subscales were considered to fit these criteria;

1. Number of patients at various steps in ED management. The number of people in the waiting room was indexed to the number of ED beds (wtroom index, see below). The number of patients in triage and the number at registration reflected the process for patients waiting to be seen (triage, registration). The number of full rooms, hallway patients, and doubled-up patients (extra patients filling beds placed in rooms that are beyond that room's normal capacity) reflected patients waiting for ED management (full rooms, hallways, doubled). These values were summed and indexed to number of ED beds (patindex, see below). The number awaiting computed tomographic (CT) scans and radiographs reflected patients waiting further tests (CT, x-ray). Patients who had been managed by the ED and were awaiting consults (consults), those awaiting admission (admit index, see below), and those waiting to be transferred to another facility (transfers out) represented the flow of patients out of the ED. Total patients registered were counted (total reg) to reflect all patients within the ED. Finally, patients on ventilators reflected the number of intubated severely ill patients who are admitted but were not yet moved to intensive care beds (respirators).

All variables were collected in the simplest manner possible. Patients waiting for two different tests would have been counted under both categories. Fast tracks and observation areas were considered parts of the ED when present.

As mentioned above, we indexed three of these questions to site-relevant demographics according to the numbers obtained on Form One by the site investigators to produce results that could be compared among sites. These three questions were as follows: 1) "number of people in the waiting room" was divided by "number of ED beds" (wtroom index); 2) "number of patients in ED beds" was the sum of full rooms, hallways, and doubled, which was then divided by "number of ED beds" (patindex); and 3) "number of patients awaiting admission" was divided by "number of hospital beds" (admit index).

2. Times needed for various steps in ED management. Waiting times for the individual portions of ED care were recorded. Times from registration and triage until patients were called from the waiting room reflected times to be seen (reg time, triage time). These were obtained by determining the time the last patient called from the waiting room was registered and triaged. Times awaiting x-ray and laboratory tests reflected times to complete tests (x-ray time, lab time). These were obtained by finding the last laboratory test or radiograph that was being done at the moment and determining the time at which it was ordered. Times from registration to the time of management and admission reflected total ED times (ED time). And finally, the longest time that an admitted patient was being held in the ED was determined (admit time).

3. Staffing in the ED. ED staffing is critical to ED patient management and can cause a significant bottleneck if the support staff is not available. Staffing in the ED was determined for three groups of health care professionals: nurses, clerical personnel, and physicians (nurses, clerks, physicians). All staffing data were collected as absolute numbers. These variables reflected the hospital's ability to fully staff the ED.

4. Diversion status. We asked how many hours out of the last 24 the ED had been on diversion to get a sense of the problem on that given day (24hrdiversion). In addition, determination of diversion status at the sampling time was recorded for each institution (diversion). Community diversion status (community) was also recorded, because each institution and its status affects the others.

Community diversion was not used in the model for numerous reasons. First, only two of the hospitals had a community plan. Second, it was difficult to define the variable, because it was used differently in different settings. Finally, diversion of specific patient types could not be reconciled with a particular community diversion status.

In addition, it was difficult to clearly define diversion status. Some hospitals never went on diversion as a matter of policy, whereas others used in-hospital factors to determine diversion. Therefore, the variable had diverse meanings in each different environment. We decided not to use the diversion in the building of a reduced model because it was more of an outcome variable, occurring when the ED was thought to be overcrowded.

Opinions on Level of Overcrowding—Outcome Variable. The outcome variable was the degree of overcrowding as assessed by the charge nurse and the emergency physicians. Because no absolute quantitative measure reflecting ED overcrowding exists, we decided on the use of a set of opinion questions as our outcome variable. Therefore, the object of this study was to quantitatively describe the staff's sense of overcrowding. The development of a quantitative predictor of opinion is well described in other areas of research, such as ED patient satisfaction.[20,21] We believe that, administratively, the ability to reflect the degree of perceived overcrowding in the ED by an objective number allows for interventions that cannot be based on opinion alone. Three responses were obtained from the ED staff members by Likert-like scale scoring. First, we asked all of the attending ED physicians to quantify the degree of overcrowding in each coverage area. Second, we asked the physicians whether they felt "rushed." These results were averaged for all attending physicians working on that shift. Finally, the ED charge nurses were asked their opinions on the degree of overcrowding. Physicians and nurses were presented with the six-point scale. No instructions or definitions other than the terms presented were given. Results were tallied on an even Likert-like scale with the following meanings: 1 = not busy, 2 = busy, 3 = extremely busy but not overcrowded, 4 = overcrowded, 5 = severely overcrowded, 6 = dangerously overcrowded. The even Likert was used so that a breakpoint for overcrowded versus not overcrowded fell between 3 and 4.

The three questions were averaged to form a composite score. The composite score was used as the outcome or dependent variable to validate the overcrowding tool developed from the objective question set. For interpretation purposes, the 6-point scale was converted linearly to a scale ranging from 0 to 200, where a NEDOCS score of 100 represented both an ED that was at capacity and the cutoff for overcrowding (0 = not busy, 40 = busy, 80 = extremely busy but not overcrowded, 120 = overcrowded, 160 = severely overcrowded, 200 = dangerously overcrowded).

Survey Content and Administration. Site sampling data were obtained for numbers, times, staffing, and diversion. Number data were obtained by counting patients in the waiting room, ED rooms, and ED halls. Time data were obtained from registration/triage, ancillary services (laboratory and x-ray), the charge nurse, and the attending physicians. Staffing and diversion status were obtained from the staff overseeing the ED. For determination of the outcome variable, the charge nurse and the attending ED physicians rated the degree of overcrowding.

The study was conducted at eight academic medical centers, each with an ED census of over 40,000 adult patient visits per year. The sites and the comparison between them and all academic centers are shown in Table 1. Samplings occurred over the period from February 5 to February 25, 2002.

We chose every-four-hour sampling times (9 AM, 1 PM, 5 PM, 9 PM, and 1 AM) and randomized these over a three-week period, with each time represented at least twice in the sampling for 42 total samplings at each of the nationally representative sites (a total of 336 samplings for all institutions and times). We skipped the 5 AM sampling, because our historical data indicated that this time was uniformly slower than the other times. Sampling times ranged from very slow periods to severely overcrowded periods at all sites. The sampling form containing all of the variables was completed by trained research associates at all sites.

Data Analysis. The eight sites included in the study were compared for demographic information to all 121 academic EDs with residency programs based on statistics incorporated in the SAEM website (www.saem.org). Results were compared between the two groups using confidence interval analysis.

Site information was tallied into descriptive information about the eight sites that took part in the study. The study group proved to be diverse. We included sites with diverse philosophies toward fast tracks, observation units, overcrowding policies, diversion policies, patient holding policies, and triage policies.

Missing Data. Site-specific data were available and complete on the sampling form in over 90% of the response forms. Missing sampling times were replaced with equivalent times and days within four weeks of the end of the study period. We developed a substitution and imputation strategy for missing individual values. When values were missing on a sampling time, instead of deleting the entire sampling time, individual predictive models, using the interrelationships among the other predictor variables, were used to impute the missing values for each predictor variable. Single model–based imputations were used, because minimal data were missing.[22] All imputations were performed using the "Hmisc transcan" function packaged with S-Plus (Insightful Corp., Seattle, WA).[23,24]

The Predictive Model. Selection of the predictor variables for the site-sampling form was based on review of the literature and input from all eight site coordinators. We conducted a pilot study using the site-sampling form in one ED to determine the feasibility and accuracy of data gathering.[25] In that study, we measured correlations on questions between two investigators operating simultaneously and found that correlations for 17 of 22 variables was over 0.75, which we considered good (unpublished results, 2001). A mixed-effects linear regression model was used to determine relationships between predictors and the outcome. Center effect was treated as a random variable to account for correlation among samplings within a center.[26,27]

Some data were combined to form new variables that were more practical. A composite variable representing "total ED patients" was created with the variables full rooms, hallways, and doubled, which was indexed to total patient beds and entered into the model (patindex). Variables were omitted from the full model either because of too much overlap, collinearity with other variables, or because of too many missing values.

"Diversions status" could be considered either as a predictor variable or as a result of overcrowding. Members of our group expressed the need to develop a tool to determine whether or not to go on diversion that did not include "diversion status" as a predictor variable even if it was a good predictor variable. With this in mind, the reduced model excluding diversion status was developed.

A total of 19 prespecified variables were entered into the model (see Table 2). The full-model fit was used because it provides meaningful confidence intervals using standard regression formulas as opposed to models that select variables based on stepwise and p-value–based selection models. For all ordinal and continuous variables, nonlinearity was assessed using restricted cubic splines.[28] The

TABLE 1. Comparison of Study Sites with All Academic Hospitals and with Those Academic Hospitals with Patient Volumes of >40,000 per Year

	Study Sites	Academic EDs	Difference (95% CI)
Number	8	121	
Hospital beds	569 ± 282 (308–1200)	588 ± 221 (150–1200)	19 (−135, 183)
ED beds	48 ± 32 (22–123)	35 ± 15 (16–123)	13 (0, 24)
Adult pts/yr (1,000s)	57 ± 15 (40–83)	56 ± 26 (26–196)	1 (−13, 13)
Adult admits (1,000s)	21 ± 5 (12–28)	21 ± 7 (6–60)	0 (−5, 5)

CI = confidence interval; ED = emergency department; pts = patients.

TABLE 2. Multivariable Prognostic Importance of the 19 Items on Overcrowding Outcome Using Mixed-effects Linear Regression Models

Variable	Comparison*	Effect*	95% CI†	p-value
Center	—	—	—	**0.13**
Wtroom index	**30% vs. 10%**	**6.2**	**3.4, 9.0**	**<0.0001**
Triage	1 vs. 0	−0.31	−1.5, 0.85	0.60
Reg	1 vs. 0	0.51	−0.85, 1.89	0.46
Patindex	**70% vs. 50%**	**14**	**10, 19**	**<0.0001**
CT	1 vs. 0	1.2	−0.54, 3.0	0.17
X-ray	1 vs. 0	1.3	−0.47, 3.1	0.15
Consults	2 vs. 1	2.5	−0.2, 5.0	0.06
Admit index	**2 vs. 0**	**12**	**3, 21**	**0.01**
Transfer out	1 vs. 0	0.6	−6.6, 7.7	0.87
Respirators	**2 vs. 0**	**20**	**7, 33**	**<0.01**
Reg time	**2 vs. 0 hr**	**7.8**	**3.2, 12.3**	**<0.01**
Xray time	0.5 vs. 0 hr	0.48	−2.8, 3.7	0.77
Lab time	0.5 vs. 0 hr	−0.20	−6.4, 6.0	0.95
ED time	15 vs. 10 hr	−0.8	−3.6, 2.0	0.58
Admit time	**8 vs. 3 hr**	**3.7**	**0.53, 6.7**	**0.02**
Nurses	13 vs. 8	−1.4	−11, 9	0.79
Clerks	4 vs. 2	2.1	−4.8, 9.1	0.54
24hrdiversion	1 vs. 0	5.1	−6.5, 16.8	0.38

NOTES. Significant variables and data are in bold. The full model explains 49% of the variability in outcomes ($R^2 = 0.49$). $N = 336$ (42 samplings at eight institutions).
24hrdiversion = 24-hour diversion; CI = confidence interval; CT = computed tomography; ED = emergency department; lab = laboratory; patindex = patient index; reg = registration; wtroom = waiting room.
*The effects are for the specified comparisons. For example, the 6.2 effect for wtroom index is the effect on outcome when comparing a wtroom index of 10% to 30%, or a 20% increase.
†Effects are considered statistically significant if the 95% CI does not include 0.

nonlinear terms were assessed simultaneously by a multiple degree of freedom test. If significant, nonlinear terms were retained. Because there was no a priori information concerning interactive effects, none were specified in the full models. However, interactions are considered with the reduced model below, because the reduced model that approximates the full models was more workable (see below). R^2 statistics are used to compare the predictions from the full-model fit to the observed outcomes.

Simplifying the Full Model. The full model, with all variables considered, was the most accurate at predicting the results of a new data set. The full-model fit established validity of the screening tool and had confidence limits and p-values that are more accurate than other methods. However, the full model was not very practical in an ED setting, because it requires too much data collection and relies on center as a variable. One common procedure is to use a stepwise variable selection process, but when many predictor variables are analyzed, it often yields unreliable models.[29] Instead, we chose the approach outlined by Harrell.[30] The full model with all variables was considered the criterion standard and was the model in which p-values were correct. Excluding center as a variable, a backward step-down procedure was used to reduce the model to the fewest possible questions that approximated the full-model predictions. Variables were deleted until deleted variables would make the approximation to the predictions of all questions inadequate. For the reduced model, all pair-wise interactions among the variables were checked, because the reduced model was more workable. A nomogram for obtaining predictions from the reduced model was developed. Although center effects are not included at the beginning of the step-down procedure, the accuracy of the reduced model is checked against the full model that includes center as a variable. R^2 statistics are used to compare the predictions from the reduced model to the predictions from the full model. Additionally, bias-corrected estimates of the reduced model's predicted values versus observed outcomes are compared. A nomogram for obtaining predictions from the reduced model was developed.

Predictive Model Evaluation. Model fit assessment used R^2 for the final NEDOCS score. A calibration plot was used to illustrate the model's fit across the range of predicted overcrowding from the screening tool and compared with the observed final NEDOCS score. Bootstrapping was used to estimate the model's predictive performance with the original data. For each of 200 equally sized bootstrap samples, the full model was refitted using all 18 predictors and evaluated on the original sample. R^2 statistics for each of the bootstrapped models were averaged to yield an overall estimate of model performance. This method corrects for overfitting and is a realistic estimate of the model's performance. The bias-corrected R^2 statistic is reported and derived by subtracting the expected value of the optimism from the original R^2 index. A conservative estimate of the reduced model's optimism was the full model's optimism. Therefore, the bias-corrected R^2 for the reduced model was derived by subtracting the expected value of the full model's optimism from the reduced model's R^2.

All statistical computations were performed using the S-Plus Statistical Language Version 2000 (Insightful Corp.) in conjunction with the Windows S-Plus Design library (Insightful Corp.) of modeling functions.[24,30]

RESULTS

Site Information. Table 1 shows how the study hospitals compared with all U.S. academic EDs. The eight study sites were statistically equivalent on these measures to all academic EDs except for having a slightly larger average number of ED beds, although confidence interval analysis includes zero, indicating that there may be no difference between the two groups. Based on this information, we considered the "average" academic ED to be 50 beds in the ED and 500 beds in the hospital.

Site-information Form Results. Table 3 is a compilation of the data available from the site sampling survey for the eight institutions included in the study. These data show the wide range of demographic information from our institutions. As can be seen, most of these hospitals were tertiary-care centers with active programs to prevent or decrease the impact of overcrowding. Of interest, all of these hospitals' inpatient beds were closed to new traffic at some time interval in the three months preceding our study because of staffing or overcrowding.

TABLE 3. Demographic Information on the NEDOCS Study Group

Characteristic	Number of Sampling Sites Answering in the Affirmative*	%
Tertiary care facilities	6	75
Amount of overcrowding		
Always	6	75
Common	2	25
Have compared periods of overcrowding and nonovercrowding	0	0
At least some patients spend their entire admission in the ED	7	87
Patients are held in hallways on the floors	2	25
Inpatients beds were closed down in the preceding three months because of inadequate staffing	8	100
Has diversion policy	6	75
Is this policy effective? (opinion)	5	63
Diversion is based on		
ED census	4	50
Hospital census	3	37
Have overcrowding policy	3	37
Are these present policies effective? (opinion)	1	13
Recently expanded the ED	2	25
Have fast tracks	8	100
Have observation units	2	25
Triage patients to other areas of the hospital	7	87
Triages patients out of the hospital	1	13

NOTE. Information completed by site investigator before data collection. All opinion questions left to the investigator to define.
*Out of eight total sites.
ED = emergency department; NEDOCS = National ED Overcrowding Study.

Site-sampling Form. Table 4 shows the compiled data for each question from the 336 site-sampling times. The last column indicates the variable name and whether or not it was indexed to site-specific demographics. Patient numbers varied widely, with up to 88 patients in the departments and 162 waiting to be seen at any given time. In some cases, times were up to 100 hours to complete tests and admit patients to the hospital.

Six questions were omitted from the model development at this point. The numbers of full rooms, hallways, and doubled were combined into one overall variable (patindex). Total number of registered patients (total reg) was interpreted differently at different institutions and had to be dropped. Time from triage until called from waiting room (triage time) was omitted because of the almost complete correlation with time from registration until called from the waiting room, indicating that it did not add any information. Similarly, longest time in the ED for admitted patient since admission (bed time) was omitted because it was too similar to longest time in the ED for admitted patient since registration, a more practical and more easily documented variable. Finally, variation in number of physicians (MDs) was too small to add any information to the model. The 19 predictor variables listed in Table 2 were entered in the final model.

Model Development. Datasets for nine of 336 (3%) of the times and days were missing and were replaced by site-samplings at equivalent days and times at the same institution. Data were imputed for missing values by an imputation scheme. The overall nonlinear effect was nonsignificant ($p > 0.20$), indicating an insignificant departure from linearity.

The results for the three outcome questions and the composite outcome variable are shown in Table 5.

TABLE 4. Results of Objective Questions on the Study Form

	Overall Range	Overall Median (IQ Range)	Variable Name (if indexed, to what)
ED numbers			
In waiting room	0–70	15 (5–29)	Wtroom index (index—Number of ED beds)
At triage	0–40	2 (0–4)	Triage
At registration	0–22	2 (0–3)	Registration
Full rooms	0–86	25 (18–32)	Full rooms†
Hallway patients	0–33	2 (0–4)	Hallways†
Doubled-up patients	0–13	0 (0–0)	Doubled†
Total ED patients	0–88	27 (21–38)	Patindex (index—Number of ED beds)
Awaiting CT scans	0–15	1 (0–3)	CT
Awaiting x-rays	0–26	2 (0–4)	X-ray
Awaiting consults	0–19	2 (1–4)	Consults
Awaiting inpatient beds	0–42	5 (3–9)	Admitindex (index—Number of hospital beds)
Awaiting transfer out	0–7	0 (0–0)	Transfer out
Total registered	1–162	37 (24–57)	Total reg†
On respirators	0–9	0 (0–0)	Respirators
ED times (hours)			
Time from registration until called from waiting room	0.0–40.0	0.6 (0.1–2.0)	Reg time
Time from triage time until called from waiting room	0.0–40.3	0.6 (0.1–2.0)	Triage time†
Time from x-ray order until called for x-ray	0.0–18.0	0.3 (0–0.6)	X-ray time
Time from lab order until completed	0.0–4.0	0.5 (0.4–1.0)	Lab time
Longest time in ED for ED patient since registration	1.1–53.5	12.5 (8.6–20.0)	ED time
Longest time in ED for admitted patient since registration	0.5–100.2	12.1 (8.3–20.0)	Admit time
Longest time in ED for admitted patient since admission	0–89.8	5.5 (2.5–11.0)	Bed time†
ED staffing			
Number of nurses	(4–18)	11 (8–14)	Nurses
Number of clerical personnel	(0–10)	3 (2–4)	Clerks
Number of attending physicians	(1–4)	2 (1–3)	MDs†
Diversion status	Overall	95% CI	
Number of hours out of the last 24 on diversion	0–24	5.3 (0,9)	24 hr diversion
% of time on diversion* (6/8 hospitals)	26%	22%, 31%	Diversion
% of time on community diversion plan? (2/8 hospitals)	61%	50%, 71%	Community†

CT = computed tomography; ED = emergency department; lab = laboratory; reg = registration.
*Percent of times over the sampling period at which the ED was on diversion.
†Variable not used in the model because of either incomplete data or overlap with other variables (See text).

Pearson correlation coefficients were between 0.86 and 0.93, indicating that the variables were all very highly correlated with the combined outcome variable. The Spearman coefficients, which do not assume linearity, deviated at most 0.011 units from the Pearson coefficients. Cronbach alpha for the three outcome questions was 0.87 (95% confidence interval (CI) = 0.84 to 0.89), indicating good internal validity. The simple composite score was almost perfectly correlated (Spearman $\rho = 0.998$) with the first principal component score, which explained 80% of the variance of the three questions; therefore, the composite score was an appropriate value to represent outcome. Based on the outcome variable, our hospital samples were overcrowded an average of 35% of the time. For individual EDs, this ranged from 12% at the least overcrowded to 73% at the most overcrowded ED. Two of the six hospitals were overcrowded 50% or more of the time.

Results of the mixed-effects linear regression model for the full fit of object data are shown in Table 2. Nonlinear terms were not included because overall, nonlinear tests were not significant in preliminary analyses (p-values > 0.20). For the full-model fit, the

TABLE 5. The Composite Outcome Variable and the Characteristics of its Components

	Overall Range	Overall Mean (± SD)	Correlation with the Composite Outcome Variable
ED physician opinion on overcrowding (average of all areas)	0–6	2.7 (±1.4)	0.93
ED physician feeling of being rushed	0–6	2.6 (±1.3)	0.89
Charge nurse opinion on overcrowding	0–6	2.9 (±1.4)	0.86
Composite outcome variable	0–6	2.7 (±1.4)	1.00

NOTE: The composite outcome variable was the mean of these three questions.
ED = emergency department; SD = standard deviation.

wtroom index, patindex, and admit index were all significant ($p < 0.01$). Also, respirators, reg time, and admit time were significant predictors as well ($p < 0.02$). The regression coefficients are given for specified comparisons in Table 2. For example, a 0.20 increase in the patindex (comparing a sampling with a patindex of 0.50 with 0.70) will correspond to a 14-unit increase in the overcrowding effect (95% CI = 11 to 17; $p < 0.0001$). The full-model fit explained 49% of the observed outcome ($p < 0.001$). The intraclass correlation coefficient, a measure of the degree of dependence within clusters, was 0.18 ($p = 0.13$).

Because full-model fits contain many predictors that do not appreciably affect the predictions, simpler solutions were sought. A reduced model was developed to approximate the full-model fit. The reduced model approximated the full-model fit with the five predictors patindex, respirators, admit time, admit index, and reg time. Allowing for all pairwise interactions between those five variables produced one significant interaction effect between patindex and reg time (p-value = 0.01). As reg time increases, then patindex has a larger effect. Figure 1 shows the results for the individual variables in the reduced model. The final reduced model using five variables and no interaction term for presentation purposes predicted the full-model fit with 88% accuracy. Figure 2 is the validation curve for the NEDOCS scale for the reduced model that compares the predicted and observed outcomes.

Figure 3 shows the nomogram developed from the objective questions used to determine the NEDOCS score for the reduced model with an example of its use. Because the interaction term of patindex and reg time represented two continuous variables and only added an additional 1% explanation of the variability over the no-interaction model, the nomogram for the reduced model does not include the interaction term. The nomogram allows the reader to determine the NEDOCS score for any academic institution on a given day and time. Points from the top line are assigned based on the variable sets. These point values are added and used on the line marked "total points"(line E), and a perpendicular line is dropped to the NEDOCS score (line F). An example of the determination of the NEDOCS score is clearly spelled out in Figure 3. Figure 4 shows a simple Web-based calculator that was developed to determine the reduced-model score based on input of the predictor variables.

DISCUSSION

A single, universally acceptable definition of ED overcrowding does not exist. In contrast to overcrowding at public places such as supermarkets, international airports, and national parks, the precise scientific definition and threshold for overcrowding in EDs is the subject of debate. Similarly, when the demand for ED services exceeds the ability to provide service, overcrowding occurs. No simple definition that succinctly describes overcrowding exists, perhaps because of the multiplicity of causes. Based on the opinion of ED directors, the definition of overcrowding in the ED should include the following[1]: 1) all available beds in the ED are full, 2) patients are placed in ED hallways because there are no inpatient beds available, 3) the ED is at some point closed because of saturation or on diversion to ambulance traffic, 4) the waiting room is full, 6) emergency physicians feel rushed, and 7) waits to see a physician are greater than one hour. In addition, there are numerous causes of overcrowding in the ED that are related to services not controlled by the ED, including the following: 1) hospital bed shortage,[23] 2) delays in radiograph and laboratory results,[31] 3) consultation delay, 4) delay of laboratory and radiology services, 5) ED nursing shortages, and 6) insufficient ED space. We incorporated all of these possibilities into a sampling form with the intent of determining which of these and others best define overcrowding. We were able to find six questions that would determine overcrowding with a high degree of reliability and discrimination. This is applicable to most academic EDs and probably other EDs as well; however, we have not yet studied the differences in problems in private EDs that might affect the calculations in this study.

Overcrowding was separated into four categories for development of a survey form. These were the following: 1) number of patients at various steps in ED management, 2) the times needed for various

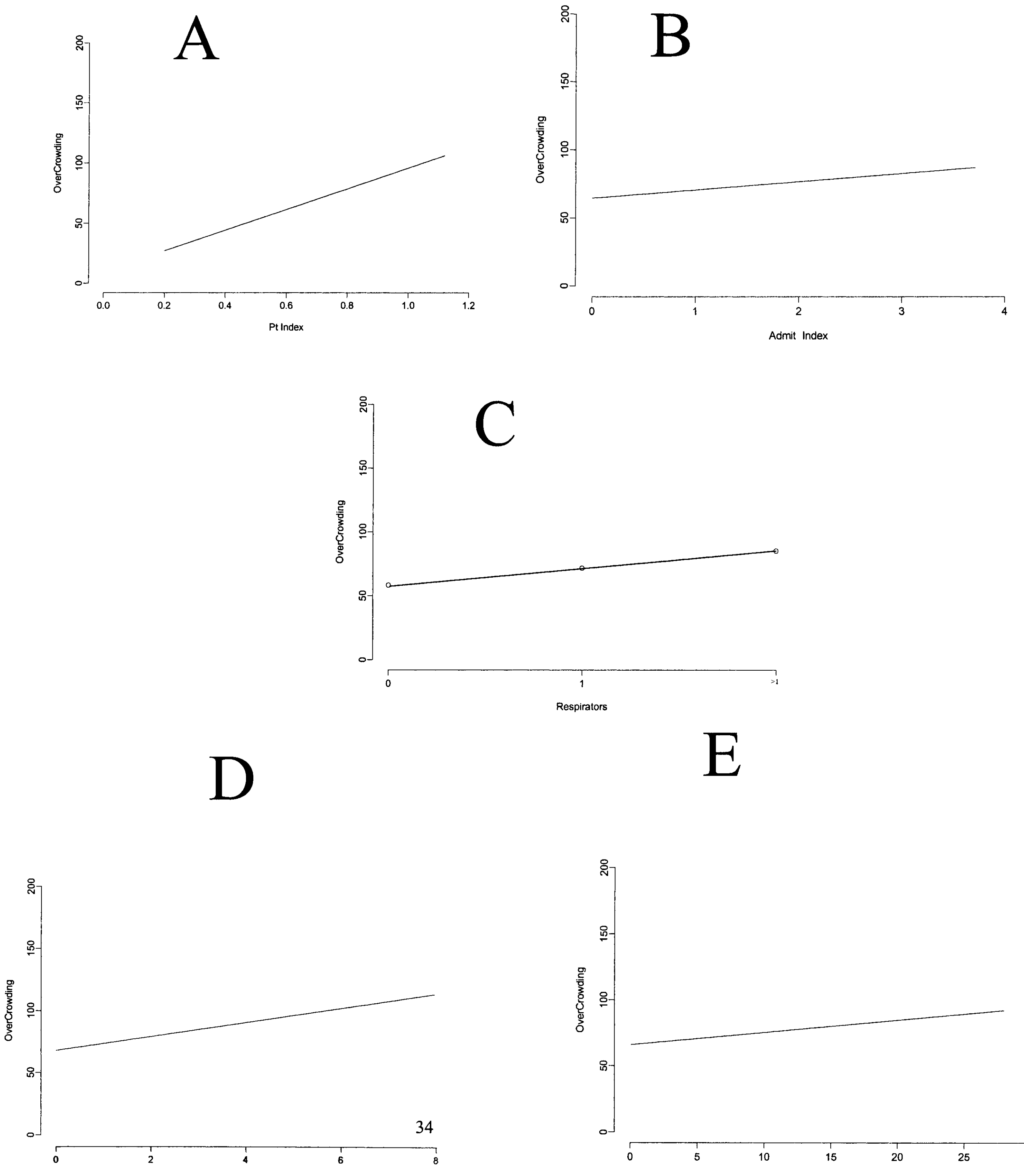


Figure 1. The relationship of the reduced NEDOCS variables patindex (**A**), admit index (**B**), respirators (**C**), reg time (**D**), and admit time (**E**) to the outcome variable representing overcrowding.

steps in ED management, 3) staffing in the ED, and 4) diversion status. Opinions on level of overcrowding among staff members were collected and used to represent the actual sense of overcrowding according to the ED staff. These opinions were recognized to be subjective and not appropriate in the development of an objective screen. However, we used this set of variables to evaluate how well the final objective screen correlated with staff opinions. As can be seen, there was an excellent correlation between these three variables and between a composite variable made up of these three and the final screening questionnaire. This

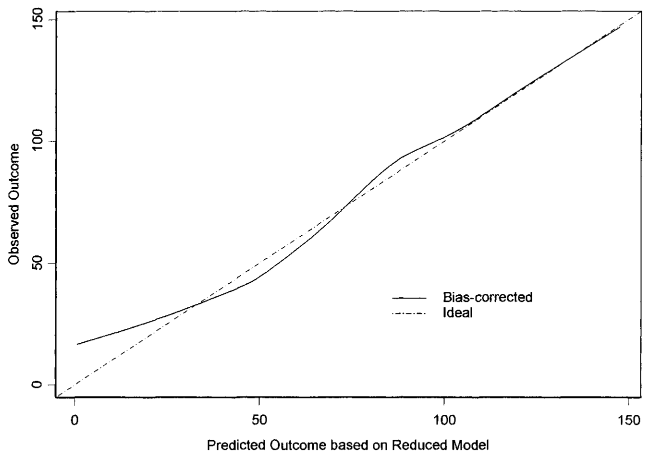


Figure 2. Bootstrap validation of the NEDOCS scale for the reduced model compared with the outcome variable.

is the best evidence we can mount to indicate that these are a standard for overcrowding and correlate well.

No other studies have directly compared academic centers using a standard objective definition of overcrowding. Our 336 samplings at eight sites indicate that we observed the full spectrum of the hospitals' census from very quiet to severely overcrowded. The data are reflective of the problem as it exists in these eight medical centers and would therefore be generalizable to other large academic EDs as shown by data from the SAEM website (www.saem.org).

The NEDOCS score is a useful indicator of the degree of overcrowding and can be used to determine the status of an ED at any given time. Because a value of over 100 represents overcrowding, it might be used to stimulate certain changes in the approach to ED administration when that level is reached. Similarly, higher numbers might trigger the implementation of safeguards to limit further overcrowding.

In addition, based on the effect size in Table 2, we can calculate the change in NEDOCS score that would be seen for changes in each of the significant variables in the model. In our example, we base our assumptions on a hospital with 50 waiting room chairs, 50 ED beds, and 500 inpatient beds (very close to our median results). An increase in ten patients in the waiting room would increase the NEDOCS score by 5, an increase in ten patients in the ED would increase the NEDOCS score by 15, and an increase in ten patients waiting to be admitted would increase the score by 7.5. An increase in number of respirators by one or an increase of five radiographs pending both increase the NEDOCS score by 8 points. Finally, we can say that an increase in waiting time to be seen of two hours increases the NEDOCS score by 6 points.

Using our definition of overcrowding based on subjective criteria, we found overcrowding to be a universal problem ranging as high as 73% of the time at the busiest center in our study. Major factors related to overcrowding were number of patients in the ED, number of people in the waiting room, number of admitted patients awaiting beds, number of respirators in use in the ED, time from registration to being called from the waiting room, and time from admission to floor transfer. Diversion status, although not used in this predictive model, was highly significantly correlated with overcrowding ($p < 0.01$).

We found that the shortage of nursing staff that is clearly a nationwide problem was reflected in our staffing subscale. There was little flexibility in the staffing based on overcrowding in the ED. At the staffing levels reflected in this study, number of

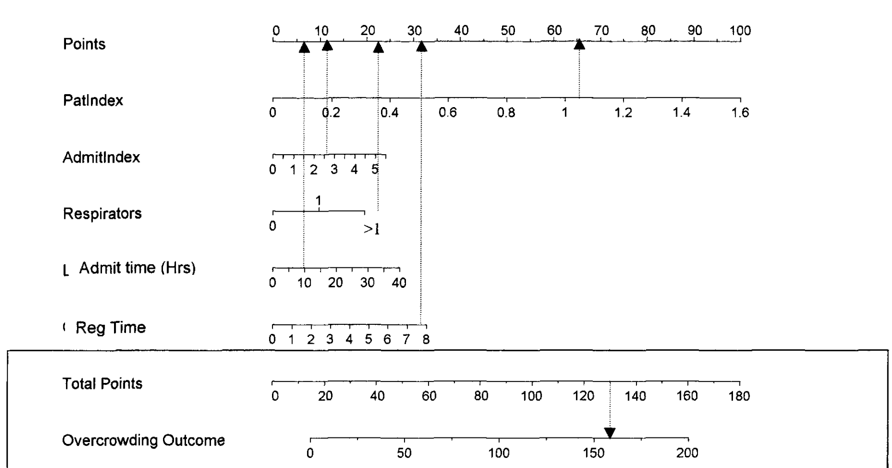
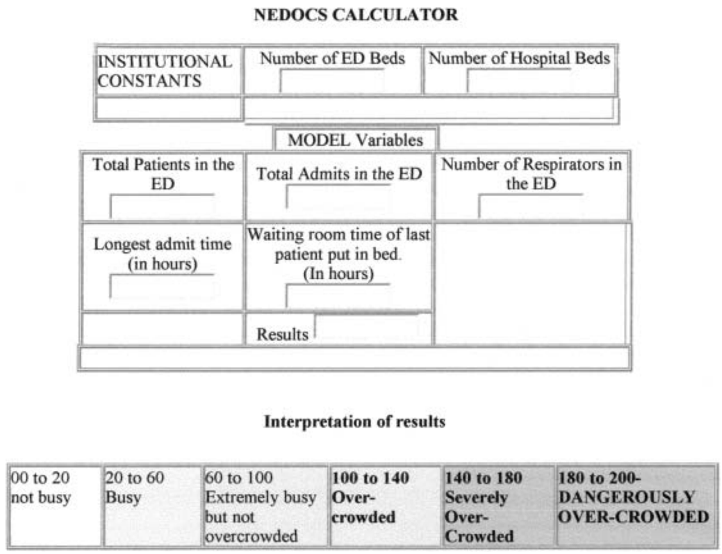


Figure 3. Nomogram of reduced model. The variables included in the nomogram are the following: patindex— the total number of patients in the ED divided by the number of ED beds; admit index—the total number of admitted patients divided by the number of hospital beds; respirators—the number of patients on respirators; admit time—the number of hours the patient waiting the longest has waited for a bed, and reg time—the number of hours the last patient called from the waiting room waited from registration until called to a bed. The example shown is of a 100% full ED (63 points), with 10 patients awaiting admission (500-bed hospital, admit index of 2 = 9 points), two respirators (19 points), a five-hour admit time (3 points), and a seven-hour wait in the waiting room (29 pts) for a total of 123 points and a NEDOCS score of 150.


NEDOCS CALCULATOR

INSTITUTIONAL CONSTANTS | Number of ED Beds | Number of Hospital Beds

MODEL Variables

Total Patients in the ED | Total Admits in the ED | Number of Respirators in the ED

Longest admit time (in hours) | Waiting room time of last patient put in bed. (In hours)

Results

Interpretation of results

00 to 20 not busy	20 to 60 Busy	60 to 100 Extremely busy but not overcrowded	**100 to 140 Over-crowded**	**140 to 180 Severely Over-Crowded**	**180 to 200- DANGEROUSLY OVER-CROWDED**

Figure 4. A Web-based calculator using the NEDOCS algorithm to determine the degree of overcrowding. (The address of this form may be requested from the author by e-mail.)

nurses did not affect the degree of perceived overcrowding. With the incipient problem of a national nursing shortage, we may find that in the future, EDs may be overcrowded without too many patients simply based on an inadequate nurse–patient ratio.

Other possible subscales exist that may not overlap with the data collected. For example, the acuity of the patients was not considered. This was because it did not fit the criteria of being readily available and easy to quantify. We believe that it is reflected in the other more quantifiable variables such as times of care, numbers of patients, and number of patients intubated.

Numerous points emerged during the data collection phase of this study that helped to elucidate the problems in studying overcrowding. Most of these represented the qualitative differences between sites that could not be quantified but are illustrated by a few examples. The diversion policies ranged from having community diversion plans to a policy of never diverting patients. Closing the hospitals occurred when the ED was full in some institutions and when the inpatient beds were full in others. Handling of pediatrics ranged from intermixed with adult patients to children's EDs at different sites. Types of admissions to intensive care units versus step-down units had too many variations to quantify. These and other differences surfaced as we tried to compare a group of widely diverse EDs in widely diverse communities.

No single question addressed the numerous factors regarding ED overcrowding. Some questions thought to be reflective of overcrowding were less so than expected. A good example is the question about how rushed the emergency physician feels. To a certain point this reflected the condition of the ED; however, when the ED became severely overcrowded, the ED beds were full, and the emergency physician felt less rushed. At this point, the opinion of the charge nurse may have been markedly different from that of the physician.

Diversion was another area of controversy. The reasons for ED diversion were numerous, often not related to the degree of overcrowding in the ED. Many hospitals went on diversion based on in-hospital requirements even when the EDs were quiet. Two systems stated that they never go on diversion. Communities, in turn, had different manners of coping with these problems. When too many hospitals were on diversion, some communities devised a rotational strategy or considered all hospitals open.

We believe that the simple questionnaire can be used in any ED to successfully separate those days that are overcrowded from those that are not, because the subscales would be similar in all environments. This would allow ED administration to present a simple number to their hospital administration that indicates the degree of overcrowding and to do so in a manner that is comparable to other EDs.

Our goal was to create an overcrowding information form, which can be easily completed by answering five to ten simple questions and can be standardized to ED size. This allows for classification of each ED time as either "overcrowded" or "not overcrowded" in a manner that has consistency across the country. Based on this classification scheme, further studies directed at overcrowded times will be possible. In addition, ED administration can use this tool to present overcrowding information to their hospital administrators with confidence that the numbers obtained are uniform in other academic EDs.

LIMITATIONS

The limitation in the study of overcrowding is the lack of a criterion standard definition. We sought to begin the process of developing such a definition. This study is for the purpose of using a previously developed data sampling form and to develop an overcrowding model that could be used to determine the degree of ED overcrowding. As such, it is a successful introductory step in the process.

Another limitation is the variable definitions. For instance, time for radiographs was based on order-to-test completion. This variable is actually more complex because the time to completion may not include the time necessary for an official reading, which may be a determining factor in the utilization of the test for decision making. In addition, ancillary services, hospital staffing, and rate of ambulance arrival were not assessed because of difficulty obtaining these data.

Although we used the best possible choice for an outcome variable, it has its problems. We were not able to evaluate interrater reliability of the scale by people of equal qualifications at equal times because we never have two attending-level staff members in the same portion of the ED at the same time. All we could do was evaluate whether multiple site samplers would have good interrater reliability in determining values from the ED at the same time. Although correlations between site samplers were good, this does not clearly define the reliability of the outcome measure.

This study is generalizable only to academic EDs at the present time. This is based only on the similarities between our sample and the rest of the academic centers. We were able to measure comparisons on only four measures (Table 1), but the differences between EDs are far more complex than this simple approach can uncover. Also, it represents a February sampling only and therefore cannot necessarily be generalized to the remainder of the year. In the future, we need to quantify the problem in private EDs, which are the great majority of EDs in the United States. We also need to prospectively validate the models developed in this study.

We did not specifically address pediatric EDs. Although some of our sites collected this information, we found that the inclusion of children changed the study considerably, because EDs manage children in different ways. Our sites ranged from ones in which children were in the patient mix to others, that did not see children at all or saw them at a different site. We therefore included pediatrics only when then were mixed inseparably with all other patients. We did collect some data from sites with specific pediatric EDs and found that the times and conditions defining overcrowding were different from adult EDs. These data will be reported separately at a future time.

Future overcrowding research needs to evaluate some of the more difficult objective variables that did not fit the constraints of the model. This includes important variables such as the overall ED severity of illness and hospital staffing. It would be valuable to determine if they add information about the condition of the ED that was not captured by the present variables. Variables such as illness severity, however, were so discrepant between sites that only the development of a standardized severity score could allow for comparisons. Community diversion, as another example, was so divergent in meaning from one institution to the next that it was simply not amenable to a universal interpretation.

Finally, overcrowding research needs to address certain outcome variables related to overcrowding such as number of medical errors and number of patients who leave without being seen. Both of these serious problems may be consequences of overcrowding. We are very interested in determining if they correlate to overcrowding as reflected by the NEDOCS score. Conceivably, by showing that there is a direct relationship between NEDOCS score and these two variables, it should be possible to convince administrators that it is unacceptable to operate an ED at a high NEDOCS score.

CONCLUSIONS

Overcrowding varied widely between academic centers during the study period. Results of a five-question reduced model are valid and accurate in predicting the degree of overcrowding in academic centers.

References

1. Richards JR, Navarro ML, Derlet RW. Survey of directors of emergency departments in California on overcrowding. West J Med. 2000; 172:385–8.
2. Derlet RW, Richards JR. Overcrowding in the nation's emergency departments: complex causes and disturbing effects. Ann Emerg Med. 2000; 35:63–8.
3. Derlet RW. Primary care and emergency department overcrowding. 2. Successful referrals to primary care clinics. Am J Public Health. 1994; 84:123–4.
4. Derlet RW, Nishio D, Cole LM, Silva J Jr. Triage of patients out of the emergency department: three-year experience. Am J Emerg Med. 1992; 10:195–9.
5. Lynn SG, Kellermann AL. Critical decision making: managing the emergency department in an overcrowded hospital. Ann Emerg Med. 1991; 20:287–92.
6. Andrulis DP, Kellermann A, Hintz EA, Hackman BB, Weslowski VB. Emergency departments and crowding in United States teaching hospitals. Ann Emerg Med. 1991; 20:980–6.
7. Derlet RW. ED overcrowding in New York City. J Emerg Med. 1992; 10:93–4.
8. Eastaugh SR, Eastaugh JA. Putting the squeeze on emergency medicine. The many pressures on today's ED. Hosp Top. 1990; 68:21–4.
9. Fromm RE Jr, Gibbs LR, McCallum WG, et al. Critical care in the emergency department: a time-based study. Crit Care Med. 1993; 21:970–6.
10. Ornato JP. Problems faced by the urban emergency department in providing rapid triage and intervention for the patient with suspected acute myocardial infarction. Heart Lung. 1991; 20:584–8.
11. Baker DW, Stevens CD, Brook RH. Patients who leave a public hospital emergency department without being seen by a

physician. Causes and consequences. JAMA. 1991; 266: 1085–90.
12. Hagland MM. ED overcrowding spurs interest in quality and credentialing issues. Hospitals. 1991; 65:33–6.
13. Gallagher EJ, Lynn SG. The etiology of medical gridlock: causes of emergency department overcrowding in New York City. J Emerg Med. 1990; 8:785–90.
14. Gibbs N. Do you want to die? The crisis in emergency care is taking its toll on doctors, nurses, and patients. Time. 1990; May 28:58-65.
15. Eisenberg D. Critical condition. Time. 2000; Jan 31:52-4.
16. Derlet R, Richards J, Kravitz R. Frequent overcrowding in U.S. emergency departments. Acad Emerg Med. 2001; 8:151–5.
17. Feferman I, Cornell C. How we solved the overcrowding problem in our emergency department. Can Med Assoc J. 1989; 140:273–6.
18. Purnell L. Reducing waiting time in emergency department triage. Nurs Manage. 1995; 26:64Q, 64T, 64V.
19. Nordberg M. Closed for business: overcrowding in the ED. Emerg Med Serv. 2001; 30:31–44.
20. Thompson DA, Yarnold PR, Williams DR, Adams SL. Effects of actual waiting time, perceived waiting time, information delivery, and expressive quality on patient satisfaction in the emergency department. Ann Emerg Med. 1996; 28:657–65.
21. Nerney MP, Chin MH, Jin L, et al. Factors associated with older patients' satisfaction with care in an inner-city emergency department. Ann Emerg Med. 2001; 38:140–5.
22. Crawford SL, Tennstedt SL, McKinlay JB. A comparison of analytical methods for non-random missingness of outcome data. J Clin Epidemiol. 1995; 48:209–19.
23. Krochmal P, Riley TA. Increased health care costs associated with ED overcrowding. Am J Emerg Med. 1994; 12:265–6.
24. Harrell FE. Design: S functions for biostatistical/epidemiologic modeling, testing, estimation, validation, graphics, and prediction. 2003. Available at: hesweb.med.virginia.edu/biostat/s/Design.html. Accessed in 2001.
25. Weiss SJ, Arndahl J, Ernst AA, Derlet R, Richards J, Nick TG. Development of a site sampling form for evaluation of ED overcrowding. Med Sci Monit. 2002; 8: CR549–53.
26. Wears RL. Advanced statistics: statistical methods for analyzing cluster and cluster-randomized data. Acad Emerg Med. 2002; 9:330–41.
27. Donner A, Klar N. Design and Analysis of Cluster Randomization Trials in Health Research. New York, NY: Oxford University Press, 2000.
28. Harrell FE Jr, Lee KL, Pollock BG. Regression models in clinical studies: determining relationships between predictors and response. J Natl Cancer Inst. 1988; 80:1198–202.
29. Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med. 1996; 15:361–87.
30. Harrell FE. Regression modeling strategies: with applications to linear models, logistic regression and survival analysis. New York, NY; Springer-Verlag, 2001.
31. Davis B, Sullivan S, Levine A, Dallara J. Factors affecting ED length-of-stay in surgical critical care patients. Am J Emerg Med. 1995; 13:495–500.