Critical Review Form

Prognosis

Development and Validation of a New Index to Measure Emergency Department Crowding, *Acad Emerg Med* 2003;10:938-942

Objectives: **"To develop a simple, quantitative, reproducible, valid measure of ED busyness, which could be integrated into clinical information systems and used as a real-time tool". In addition, "a secondary goal of this study was to examine whether adverse events are more common during periods of peak ED crowding." (p. 938)**

Methods: **To develop and validate the Emergency Department Work Index (EDWIN), the investigators conducted a prospective observational survey of on-duty attending physicians and charge nurses at Beth Israel Medical Center in Newark, NJ, a Level 2 trauma center. From March 28, 2002 to May 2, 2002 two-hour intervals during 60 eight-hour shifts were selected as a convenience sampling. A trained research assistant approached the on-duty shift physician and charge nurse with a single question: "*How busy would you say the ED is right now?* Please take into account your workload, the workload of all the other doctors and nurses, the number of patients in the ED and waiting room, and numbers of holds (admitted patients waiting for beds):**

1) **not busy at all, not crowded**
2) **steady, easily keeping up**
3) **average: working hard, but keeping up**
4) **more crowded and busy than desirable**
5) **extremely busy, very crowded"**

Nurses and physicians were blinded to each other's responses. These responses were then collapsed into three categories (1or 2 = not busy; 3 = average busyness; 4-5 = busy) and correlated with the EDWIN score via the Kruskal-Wallis test. The EDWIN score was defined as

$$\sum n_i t_i / N_a (B_T - B_A)$$

where n_i **= the number of patients in the ED with triage category *i;***
t_i **= inverted ESI triage category (1-5, 5 = most acute);**

N_a= number of attending physicians on duty;
B_T = the total number of beds in the ED (filled & empty);
B_A = the number of admitted patients in the ED.

Therefore, the EDWIN unit can be described as patient triage units per attending physician per available bed. (p. 939)

To assign each patient individual EDWIN scores the average department EDWIN scores for the time period they were in the ED was used. Two composite secondary outcomes were also evaluated relative to the individual EDWIN scores. First, the composite QI referrals, radiology overreads and 72-hour returns. The second composite outcomes: walkout and AMA patients. Demographic, clinical parameters, and return rates were obtained by a review of the ED's electronic medical record.

Guide		Comments
I.	**Are the results valid?**	
A.	**Was the sample of patients representative?** *In other words, how were subjects selected and did they pass through some sort of "filtering" system which could bias your results based on a non-representative sample. Also, were objective criteria used to diagnose the patients with the disorder?*	Although a convenience sampling of subjects over 35-days, 60/105 eight-hour shifts were sampled with 23.5% midnight – 8am; 45.8% 8am-4pm; and 30.7% 4pm-midnight. Unfortunately only 2647/4816 patients (55%) had EDWIN scores calculated. As per Table 1 (p.940) those with unavailable EDWIN scores were older (median 38 vs. 35, $p<0.001$) and more likely to be admitted (22% vs. 15%, $p<0.001$). These patients probably represent a reasonable sampling of this New Jersey population but the single-center design limits external validity.
B.	**Were the patients sufficiently homogeneous with respect to prognostic risk?** *In other words, did all patients share a similar risk from during the study period or was one group expected to begin with a higher morbidity or mortality risk?*	As per Table 1, those without EDWIN scores appear less ill and younger.
C.	**Was follow-up sufficiently complete?** *In other words, were the investigators able to follow-up on subjects as planned or were a significant number lost to follow-up?*	No lost to follow-up is reported.

D.	**Were objective and unbiased outcome criteria used?** Investigators should clearly specify and define their target outcomes before the study and whenever possible they should base their criteria on objective measures.	For the secondary outcomes the investigators do not report their electronic medical record review techniques, radiology overread communication interval or QI methods or interval of follow-up.
II.	**What are the results?**	
A.	**How likely are the outcomes over time?**	• 2647/4816 (55%) of subjects had EDWIN score calculated. • 225/420 (53.6%) of potential EDWIN two-hour scores were calculated. • Nurse-physician agreement on dept busyness was moderate (weighted κ=0.61; 95% CI 0.53-0.69) and there was excellent correlation of EDWIN score with their busyness assessment ($p<0.001$). • Median EDWIN with IQ range for not busy 1.07 (.80-1.55) average 1.55 (1.16-1.93) and busy 1.83 (1.42-2.45). • EDWIN score was associated with diversion with median EDWIN on diversion 2.77 compared with EDWIN not on diversion 1.45. (p. 940) • For the composite outcome including QI referrals 34 patients were included with median EDWIN 1.90 (compared with 1.66 for all others) which was not significant. • When patients were stratified into one of three groups by EDWIN score (not busy EDWIN <1.31, average busy 1.32-1.69, busy >1.70), then the 67 subjects with the QI composite had significantly higher EDWIN scores (p.=0.03)
B.	**How precise are the estimates of likelihood?** *In other words, what are the confidence intervals for the given outcome likelihoods?*	The median interquartile ranges widely overlap so not very precise.

III.	How can I apply the results to patient care?	
A.	**Were the study patients and their management similar to those in my practice?**	Yes, ED patients in one busy academic urban ED.
B.	**Was the follow-up sufficiently long?**	Yes, 72-hour ED recidivism is sufficiently long to judge over-crowding related adverse medical events.
C.	**Can I use the results in the management of patients in my practice?**	Yes, the EDWIN score offers a simple algebraic computation to judge departmental patient volume with overall institution-specific resources.

Limitations

1) **Single center study limits external validity**

2) **Convenience-sampling may represent a selection bias and not be representative of ED at all time periods (time-of-day, time-of-year) or represent patient flow during extraneous circumstances (flu pandemics, epidemics, or disasters).**

3) **Failure to include confounding variables in EDWIN including nursing staffing, resident staffing, ED square footage, timeliness of ancillary services, procedural load, and overall patient illness severity. Although inclusion of ESI is an attempt to gauge the latter, alternative patient-specific computation might yield more accurate measurements for inclusion in EDWIN (APACHE, MEDS, TIMI, etc).**

4) **Is the definition of overcrowding (clinician perception) adequate?**

Bottom Line

Single-center one-month review suggests that the EDWIN score is a valid marker of ED overcrowding with score <1.5 not busy and >2 overcrowded. If validated, EDWIN can serve as a standardized marker of overcrowding for researchers and policy makers across varied ED settings. In addition, EDWIN might be used locally to gauge impending overcrowding, and optimize activation of institutional crisis plans including diversion.