Critical Review Form
Clinical Prediction or Decision Rule

Validation of the Mortality in Emergency Department Sepsis (MEDS) score in patients with the systemic inflammatory response syndrome (SIRS), *Crit Care Med* 2008; 36:421-426

Objectives: **"To prospectively validate the MEDS score in a heterogeneous, multi-center group of patients with SIRS. The secondary objectives were to recalibrate the MEDS score, to compare the original MEDS score and the recalibrated MEDS score to serum lactate concentration alone, and to assess the reliability of the individual MEDS score components and the score as a whole when used in the ED". (p. 422)**

Methods: Prospective multicenter convenience sampling at the University of Colorado Hospital (Level II trauma center), Denver Health Medical Center (Level I), Albert Einstein Medical Center (Philadelphia, Level I), and the Hospital of the University of Pennsylvania (Level I) from August 1, 2005 through January 31, 2006. Inclusion criteria included adults > 18 years old with SIRS criteria who were admitted to the hospital. Note that this is a different population than Shapiro's derivation trial that required a blood culture be sent (suspected infectious etiology). Exclusion criteria included trauma, previous study enrollment, direct admission or transfer from another institution, or failure to enroll within two hours of ED presentation.

Enrollment was the responsibility of triage nurses or treating physicians/nurses at each site. Data collection at the University of Colorado was by treating resident or attending physicians, whereas every other site used trained research assistants "supervised by an investigator from each site" (p. 422) using a "standardized closed-response data collection instrument". The MEDS variables were defined as follows:

terminal illness – physicians judgment of > 50% likelihood of 30-day mortality after the index ED visit;

septic shock – sepsis with systolic BP < 90 mmHg that persists despite a crystalloid bolus;

respiratory difficulty – respiratory rate > 20, oxygen saturation < 90%, or need for supplemental oxygen by either face mask or 100% non-rebreather;

lower respiratory tract infection – infiltrate on chest x-ray or the presence of clinical findings suggestive of this diagnosis;

altered mental status – any difference from the patient's baseline in any of the three spheres of orientation or in their level of alertness.

The primary outcome was 28-day mortality as determined by "systematically reviewing patient hospital records resulting from the index hospitalization". (p. 422) Unfortunately the investigators provide no methods for this chart review (Gilbert 1996, Worster 2004). If discharged before Day 28 they were assumed to have survived. At one institution, one investigator assessed the reliability of individual MEDS variables by independently abstracting them from the medical record retrospectively.

The investigators assessed overall discriminatory ability of the MEDS score for the 28 day mortality by analyzing ROC AUC's with 95% CI's. They reconstructed the five risk stratification categories described by the derivation authors and reported mortality rates for each risk group. They also sought to derive a better risk model by recalibrating the MEDS score via a *post hoc* logistic regression analysis and comparing the ROC AUC of the recalibrate score with the original MEDS score.

Guide		Comments
I.	***Is this a newly derived instrument* (Level IV)?**	
A.	Was validation restricted to the retrospective use of statistical techniques on the original database? (If so, this is a Level IV rule & is not ready for clinical application).	No, this multi-center study was prospectively validated so not a Level IV CDR.
II.	**Has the instrument been validated? (Level II or III). If so, consider the following:**	
1a	Were all important predictors included in the derivation process?	As described in Table 2 all of the MEDS predictor variables were present to a varying degree across the four heterogeneous institutions. The AEMC population was primarily geriatric (58% > 65 years) whereas the HUP population was younger (only 21% > 65 years and 3% in nursing homes).

1b	Were all important predictors present in significant proportion of the study population?	Yes, the varying degrees across institutions. The heterogeneity increases the external validity of the MEDS score.
1c	Does the rule make clinical sense?	Yes, each variable is logically related to progressively more severe SIRS-related pathology and intuitively ought to be associated with increasing mortality. More importantly unlike the ICV-derived and ICU-validated APACHE-II and SAPS scores the MEDS score includes variables available in the ED.
2	Did validation include prospective studies on several different populations from that used to derive it (II) or was it restricted to a single population (III)?	Multicenter prospective data collection so a Level II CDR.
3	*How well did the validation study meet the following criteria?*	
3a	Did the patients represent a wide spectrum of severity of disease?	Yes, the validation cohort represented a wide variety of terminally ill (range 3 – 16%), respiratory distress (29 – 77%), septic shock (4 – 20%), thrombocytopenia (4 – 23%), lower respiratory tract infections (7- 45%), nursing home residents (3 – 40%), and AMS (7 – 38%). Mortality ranged from 7 – 11%, however, well below the 18% mortality commonly reported from sepsis (Angus 2001) or the 28% mortality expected in septic shock with lactate > 4 mmol/L (Nguyen 2004)so not testing the highest risk septic shock population with this validation trial. It is important to note that 27% of this cohort did not have an infectious source for their SIRS and included DKA, GI bleeding and thromboembolic disease. (p. 423)
3b	Was there a blinded assessment of the gold standard?	Yes. "All MEDS score data also were recorded in a fashion in which the physician was blinded to the patient's outcome". (p. 422). It is not already stated that outcome assessors were blinded to the MEDS score, but presumably they were to avoid ascertainment bias.

<table>
<tr>
<td>3c</td>
<td>Was there an explicit and accurate interpretation of the predictor variables & the actual rule without knowledge of the outcome?</td>
<td>As described in the methods, the MEDS predictor variables were explicitly described which ought to enhance accuracy. Furthermore, the reliability assessment at one institution provided the following Kappa values:

Reliability of the individual Mortality in Emergency Department Sepsis (MEDS) score variables (n = 75)

<table>
<tr><th>Variables*</th><th>κ</th><th>95% CI</th></tr>
<tr><td>Terminal illness</td><td>0.64</td><td>0.40-0.89</td></tr>
<tr><td>Tachypnea or hypoxemia</td><td>0.96</td><td>0.89-1.00</td></tr>
<tr><td>Septic shock</td><td>0.96</td><td>0.87-1.00</td></tr>
<tr><td>Platelets <150,000/mm^3</td><td>0.95</td><td>0.87-1.00</td></tr>
<tr><td>Bands >5%</td><td>0.82</td><td>0.57-1.00</td></tr>
<tr><td>Age >65 yrs</td><td>1.00</td><td>0.92-1.00</td></tr>
<tr><td>Lower respiratory Infection</td><td>1.00</td><td>0.92-1.00</td></tr>
<tr><td>Nursing home resident</td><td>0.95</td><td>0.86-1.00</td></tr>
<tr><td>Altered mental status</td><td>1.00</td><td>0.92-1.00</td></tr>
</table>
CI, confidence interval.
*See text for definition of each variable.

The Kappa for terminal illness is low and may represent one weakness of the MEDS score since different physician may score the same patient differently on this variable. Since the variables and the summed MEDS score were recorded during the ED enrollment while the outcome was not known for 28 days, it is unlikely that nurses, physicians, or research assistants were aware of the outcome at the time values for MEDS predictor variables were recorded.</td>
</tr>
<tr>
<td>3d</td>
<td>Did the results of the assessment of the variables or of the rule influence the decision to perform the gold standard?</td>
<td>No “Information recorded on study participants was not used to direct therapy or disposition”. (p. 422) All subjects had 28-day, mortality recorded regardless of their MEDS score.</td>
</tr>
</table>

4	How powerful is the rule (in terms of sensitivity & specificity; likelihood ratios; proportions with alternative outcomes; or relative risks or absolute outcome rates)?	• 385 patients were enrolled (39%, HUP, 21% UCH, 24% ACMC, 16% DHMC) with 9% (95% CI 6 – 12%) 28-day mortality and median age 56 years. The median hospital LOS was five days and 11% required mechanical ventilation. • 46% of patients had a lactate level available • ED diagnosis was available for 225 and 165 (73%) had a source of infection. • The majority of subjects were classified into very low (48%) or low (21%) or moderate (21%) risk groups. • The MEDS score predicted mortality rates similar to the original derivation trial.

Mortality in Emergency Department Sepsis (MEDS) score mortality categories

MEDS Score	Original Mortality %*b*	95%CI	Validated Mortality, %	95% CI
0-4	0.9	0-2	0.6	0-3
5-7	2	1-3	5	1-13
8-12	8	6-10	19	11-29
13-15	20	13-27	32	15-54
> 15	50	36-64	40	12-74

		• The ROC AUC was 0.88 (95% CI 0.83 – 0.92) c/w 0.78 (95% CI 0.66 – 0.90) for lactate level alone or 0.84 (95% CI 0.78 – 089) for the recalibrated MEDS score.
III.	**Has an impact analysis demonstrated change in clinical behavior or patient outcomes as a result of using the instrument? (Level I). If so, consider the following:**	

1	How well did the study guard against bias in terms of differences at the start (concealed randomization, adjustment in analysis) or as the study proceeded (blinding, co-intervention, loss to follow-up)?	No impact analysis was performed. This is simply a non-interventional descriptive cohort to assess the prognostic accuracy of the MEDS score on heterogeneous populations distinct from that upon which it was originally derived. Despite the well-designed protocol, the investigators note two potential sources of bias: selection bias (non-consecutive sampling) and misclassification bias (variables not consistently recorded by physicians).
2	What was the impact on clinician behavior and patient-important outcomes?	Impact on clinician decision-making and patient outcomes via the incorporation of MEDS into ED operations was not assessed in this non-interventional trial. Future trials will need to assess clinician acceptance of MEDS scores into patient care protocols and the impact of such use upon patient-oriented outcomes (disease-free survival). In the meantime, this trial in conjunction with Chen 2006, Howell 2007, Jones 2008 and Lee 2008 mean that the MEDS score has been externally validated and can confidently be incorporated into goal-directed pathways for SIRS and sepsis, as well as sepsis bundles and research protocols to prognostically stratify subjects into low- and high-risk groups pre-intervention.

Limitations

1) **Failure to incorporate lactate into the MEDS score.**

2) **Failure to conduct sensitivity analysis for the subset where bands were not assessed and by default listed as "not present".**

3) **Failure to report AUC for various institutions to assess diagnostic accuracy in heterogeneous populations.**

4) **No chart review method reported or referenced.**

5) **Reliance upon a well-accepted but probably grossly insufficient metric for altered mental status since EM nurses and physician under-recognize > 60% of geriatric patients with cognitive dysfunction (Carpenter 2010). Although this represents practice reality, in a study setting it would have been helpful to report percentage with caregiver reporting mental status in increments. Future research may want to assess the MEDS score using more formal assessments of mental status (Carpenter 2008).**

6) **No assessment of clinician's gestalt estimate of 28-day mortality to compare with MEDS AUC.**

Bottom Line

The MEDS score is a reliable and accurate predictor of 28-day mortality in ED patients admitted to the hospital with SIRS (infectious and non-infectious etiologies). Future impact analyses will need to assess whether EP's and admitting physicians accept the MEDS score as a SIRS prognostic instrument AND incorporate it into management protocols. In addition, the impact of the MEDS score on overall resource utilizations, costs and patient outcomes will need to be assessed.

MEDS Score

Risk Factor	Points
Rapidly terminal co-morbid illness	6
Age > 65	3
Bands > 5%	3
Tachypnea or hypoxia	3
Shock	3
Platelet < 150,000 mm^3	3

Altered mental status	2
Nursing home resident	2
Lower respiratory infection	2

Interpretation of the MEDS score from the accumulated validation trials through late 2007 (Carpenter 2008):

Score	**Label**	**Range in 28-day Mortality**
0-4	Very low risk	0.4%-11.0%
5-7	Low risk	3.3%-5.0%
8-12	Moderate risk	6.6%-19.0%
12-15	High risk	16.1%-32.0%
>15	Very high risk	39.1%-40.0%