Critical Review Form
Clinical Prediction or Decision Rule

Mortality in Emergency Department Sepsis (MEDS) score: A prospectively derived and validated clinical prediction rule, *Crit Care Med* 2003; 31:670-675

Objectives: **"To identify univariate correlates of death in emergency department patients at risk for infection; b) to perform multivariate analyses and identify independent predictors of death; and c) to develop and internally validate a prediction rule that may be used in the emergency department to risk stratify patients into different risk groups to predict their mortality". (p. 671).**

Methods: **All adult patients (age ≥ 18 years) presenting to Beth Israel Deaconess Medical Center (Boston, MA) ED from Feb 2000 – February 2001 were eligible if the treating emergency physician obtained a blood culture and were admitted to the hospital. Patients were also included if a blood culture was obtained on the floor within three hours of admission. Data was abstracted by blinded trained chart reviewers who reviewed ED charts for the following confounding variables: chills, cough, fever, altered mental status (by history or exam), cerebrovascular disease, chemotherapy, congestive heart failure, chronic obstructive pulmonary disease (COPD), dementia, diabetes, HIV or AIDS, Hodgkin's disease, intravenous drug (IVDA) use, leukemia, liver disease, myocardial infarction (MI), non-Hodgkin's lymphoma, peripheral vascular disease, renal disease, splenectomy, steroid use, ulcer disease, transplant, presence of any malignancy, residence in a nursing home, subjective shortness of breath (SOB), and an acute abdomen by exam. Any covariate in history or physical exam not noted as present was considered to be absent. The authors also computed the Charlson co-morbidity index and used the following definitions:**

immunocompromised state - presence of HIV or AIDS, leukemia, any malignancy, any history of chemotherapy, neutropenic fever, non-Hodgkin's lymphoma (NHL), transplanted organ, or current use of steroid.

major co-morbidity - coma or brain death, bowel perforation, multiple trauma or burns, cardiopulmonary arrest within 24 hrs, or chronic hepatic failure.

The authors used a method previously described by McCabe and Jackson (*Arch Int Med* 1962; 110: 856-864) to define severe underling disease including metastatic cancer or disease with > 50% one-month mortality

Abnormal vital signs were defined as follows: 35.5°C > temperature > 38°C, heart rate > 90 bpm, systolic BP< 90 mmHg. Respiratory difficulty was defined by respiratory rate > 20, pulse oximetry < 90%, or the need for supplemental oxygen to maintain adequate oxygenation. Laboratory data assessed included CBC (total WBC, percent bands, platelets count), chemistries and anion gap. Lactate was not routinely measured.

Systemic inflammatory response syndrome (SIRS) was defined by the presence of two or more of the following: heart rate > 90, respiratory rate > 20, oxygenation < 90% or need for supplemental oxygen, 35.5°C > temp > 38°C , WBC > 15 or bands > 10%. Severe sepsis was SIRS plus organ dysfunction (AG > 16, AMS, pulse oximetry < 90%). Septic shock was severe sepsis plus BP < 90 mmHg after a 30 cc/kg bolus.

Two-thirds of patients were randomly assigned to a derivation set and one-third to a validation set. The primary outcome was 28-day mortality. First, univariate analysis identified relationships between confounding variables and 28-day mortality. If $p < 0.1$ for an individual variable on univariate analysis then that variable was entered in a forward selection logistic regression model to identify independent predictors of 28-day mortality with $p < 0.05$. Next, the stability of the model was assessed on the derivation set by bootstrap method and covariates that remained significant were retained in the final model that was tested by the Hosmer-Lemeshow goodness-of-fit test.

Investigators then took the LR models β-coefficients, divided by 0.3 and rounded to the nearest integer to derive the MEDS score (below). The MEDS score was then tested on the validation set with c-statistic and AUC ROC were computed. The MEDS score was also tested on the subset of patients more likely to receive aggressive care by excluding patients with a fatal disease, dementia, or nursing home residence.

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).	Yes, so a Level IV CDR that will require prospective validation.
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?	Yes except for lactate. However, by assuming lack of documentation equates to not present, the authors are undoubtedly missing many cases. For example, at BJH (and almost every other ED worldwide) dementia is present in 30 – 40% of adults > age 65 but only noted in ~ 5% (Carpenter 2010). The MEDS score, therefore, will require validation when these confounding variables are more proactively sought.
1b	Were all important predictors present in significant proportion of the study population?	Uncertain. Table 1 lists some, but not all of the predictor variables' prevalences. (see below)

Table 1. Patient characteristics in the derivation and validation sets

	Derivation (n = 2,070)	Validation (n = 11.090)
Age, mean (SD)	61.8 (19.6)	61.0 (19.9)
Male sex, %	45.3	47.6
Charlson score, mean (SD)	2.4 (2.5)	2.4 (2.5)
Nursing home resident, %	14.0	12.8
Terminal illness (<30 days), %	6.5	7.4
Cerebrovascular disease, %	8.5	9.7
Congestive heart failure, %	15.1	14.8
Diabetes, %	23.9	22.5
HIV, %	6.5	7.0
Malignancy, %	18.9	19.1
WBC, mean (SD)	12.2 (10.0)	11.6 (6.7)
Neutrophil %, mean (SD)	76.2 (14.2)	76.5 (14.4)
HR, mean (SD)	96.2 (21.9)	95.5 (21.9)
SBP, mean (SD)	121 (27.1)	120.5 (26.6)
Mortality rate, %	5.3	5.7

WBC, white blood cell count; HR, heart rate; SBP, systolic blood pressure.

Most importantly, the derivation process did not include lactate.

1c	Does the rule make clinical sense?	Yes, the 9-item MEDS score has content-validity in that each component would be expected to contribute to sepsis-related mortality. Although the complicated scoring system might be onerous to compute at the bedside, it is an improvement over previous scores (APACHE II, APACHE III, SAPS, MPM II, SOFA) that contain elements not available in the ED and were derived/validated in ICU settings. Incorporation of the MEDS score into EMRs or PEPID-like PDA platforms could facilitate their use 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)?	This rule was not prospectively derived on validated so it remains a Level IV CDR in this paper not yet ready for widespread use since it might not validate or be used appropriately by clinicians in other patient settings.
3	*How well did the validation study meet the following criteria?*	
3a	Did the patients represent a wide spectrum of severity of disease?	Probably. The derivation and validation cohort represent all patients with SIRS presenting to a tertiary academic medical center in Boston if blood cultures were drawn. Although the authors do not offer any other metric of acute illness severity (ESI or Injury Severity Score), the demographics they do provide in Table 1 seem representative of a busy ED's patient population.
3b	Was there a blinded assessment of the gold standard?	Uncertain. "The ED charts of all patients were prospectively reviewed either the same day or on a subsequent day by trained reviewers blinded to knowledge of the patient's hospital course or outcomes" (p. 671). However, the authors do not clearly state who identified the outcomes (28-days mortality) or whether those person(s) were blinded to the hypothesis, predictor variables, or MEDS score.
3c	Was there an explicit and accurate interpretation of the predictor variables & the actual rule without knowledge of the outcome?	The authors clearly describe their definitions for the predictor variables, but there was no *a priori* standardization prior

		to study onset by the clinicians recording the data used in building the MEDS score. This is alright in deriving the score, but future validation trials will need to assess how reliably each element of the MEDS score performs in real-world settings. For example, if the Kappa for "rapidly terminal co-morbid illness" at BJH is 0.4 than two physicians are likely to rate the same patient's risk differently which may negate the prognostic accuracy of the instrument and concurrently reduce physician acceptance.
3d	Did the results of the assessment of the variables or of the rule influence the decision to perform the gold standard?	Unlikely. The MEDS score had not yet been derived (and therefore could not have been computed) at the time that the data were being accumulated. Although elements of the MEDS score may have influenced clinical decision-making (aggressiveness of care, ICU admission), there would not have been an ethical mechanism to avoid this co-intervention bias in the rule's derivation. Future research might be able to control for such confounding variables.
4	How powerful is the rule (in terms of sensitivity & specificity; likelihood ratios; proportions with alternative outcomes; or relative risks or absolute outcome rates)?	• From 3301 encounters (2649 patients), 3179 (96%) were enrolled in the study and 2070 randomly assigned to the derivation set (5.3% mortality) an 1109 to the validation set (5.7% mortality). • The 9-item MEDS score (below) was used on the derivation set to yield 5 mortality risk groups: very low 0.9% (95 CI 0.2 – 1.5%), low 2.0% (0.8 – 3.2%), moderate 7.8% (5.6 – 10%), high 20% (13 – 27%), and very high 50% (36 – 64%) with a ROC AUC 0.82. • In the validation set the five mortality risk groups stratified as follows: very low 1.1% (95% CI 0.1 – 2.1%), low 4.4% (2.0 -6.8%), moderate 9.3% (6 – 13%), high 16% (6.5 – 26%), and very high 39% (19 – 51%) with ROC AUC 0.76.

		• After excluding patients unlikely to receive aggressive care (fatal disease, nursing home resident or dementia) the MEDS score ROC AUC was 0.80.
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 has been performed so no randomization or blinding pertinent. Consecutive patients were included and randomly assigned to derivation and validation subsets but by all available prognostic variables the derivation and validation subsets appear similar.
2	What was the impact on clinician behavior and patient-important outcomes?	No impact analysis has been performed so we remain uncertain how clinicians would accept, interpret or use the MEDS score or actual patients. Would they accurately and reliably compute the MEDS score on individual patients? Would knowledge of the MEDS score impact diagnostic, therapeutic, or disposition decisions? Will knowledge of MEDS score translate into improved patient outcomes?

Limitations

1) **Single-center retrospective derivation and validation trial so uncertain if valid in other populations.**

2) **No impact analysis. Future research will be essential to demonstrate clinician's accurately and reliably calculate MEDS score and ultimately that it improves patient-important outcomes.**

3) **Selection bias in that blood-stream infection had to be suspected by ordering blood cultures so less clinically apparent populations not represented in this research.**

4) **Uncertain reliability of risk-factor "severe underlying disease". The original methods by McCabe and Jackson did not describe reliability (reproducibility) of this term so we are uncertain if two EP's would label the same patient the same way with this risk factors.**

5) **No inclusion of lactate or lactime in the derivation set.**

Bottom Line

The MEDS score has the potential to risk-stratify septic ED patients for 28-day mortality. Currently, it is the only instrument validated on ED patients (as opposed to APACHE, SAPS, etc. Future prospective research should assess the external validity, accuracy, and impact on patient-oriented outcomes before widespread use of this clinical decision rule is recommended.

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 this validation set:

Score	Label	28-day Mortality (95% CI)
0-4	Very low risk	1.1% (0.1%-2.1%)
5-7	Low risk	4.4% (2.0%-6.8%)
8-12	Moderate risk	9.3% (6.0%-13.0%)
12-15	High risk	16.0% (6.5%-26.0%)
>15	Very high risk	39% (19.0%-51.0%)