

## METHODOLOGY

## Understanding receiver operating characteristic (ROC) curves

Jerome Fan, MD; Suneel Upadhye, MD, MSc; Andrew Worster, MD, MSc

SEE ALSO PAGE 13.

In this issue of the Journal, Auer and colleagues<sup>1</sup> conclude that serum levels of neuron-specific enolase (NSE), a biochemical marker of ischemic brain injury, may have clinical utility for the prediction of survival to hospital discharge in patients experiencing the return of spontaneous circulation following at least 5 minutes of cardiopulmonary resuscitation. The authors used a receiver operating characteristic (ROC) curve to illustrate and evaluate the diagnostic (prognostic) performance of NSE. We explain ROC curve analysis in the following paragraphs.

The term “receiver operating characteristic” came from tests of the ability of World War II radar operators to determine whether a blip on the radar screen represented an object (signal) or noise. The science of “signal detection theory” was later applied to diagnostic medicine.<sup>2</sup> The determination of an “ideal” cut-off value is almost always a trade-off between sensitivity (true positives) and specificity (true negatives). As both change with each “cut-off” value it becomes difficult for the reader to imagine which cut-off is ideal. The ROC curve offers a graphical illustration of these trade-offs at each “cut-off” for any diagnostic test that uses a continuous variable.<sup>3</sup> Ideally, the best “cut-off” value provides both the highest sensitivity and the highest specificity, easily located on the ROC curve by finding the highest point on the vertical axis and the furthest to the left on the horizontal axis (upper left corner) (Fig. 1). However, it is rare that this ideal can be achieved, so that, for example, one may opt to choose a higher sensitivity at the cost of lower specificity. In the NSE study,<sup>1</sup> the authors chose a cut-off point of  $>30 \mu\text{g/L}$  with a specificity of 100% and sensitivity of 79% (Fig. 2). A cut-off point

with high specificity allows the authors to “rule-in” the outcome for all patients with a NSE value above the selected cutoff.<sup>4</sup> The study indicates that patients with a NSE level  $>30 \mu\text{g/L}$  will die before hospital discharge and those with a NSE level  $<29 \mu\text{g/L}$  will possibly survive to hospital discharge.

The area under the ROC curve (AUC) is widely recognized as the measure of a diagnostic test’s discriminatory power.<sup>5</sup> The maximum value for the AUC is 1.0, thereby indicating a (theoretically) perfect test (i.e., 100% sensitive

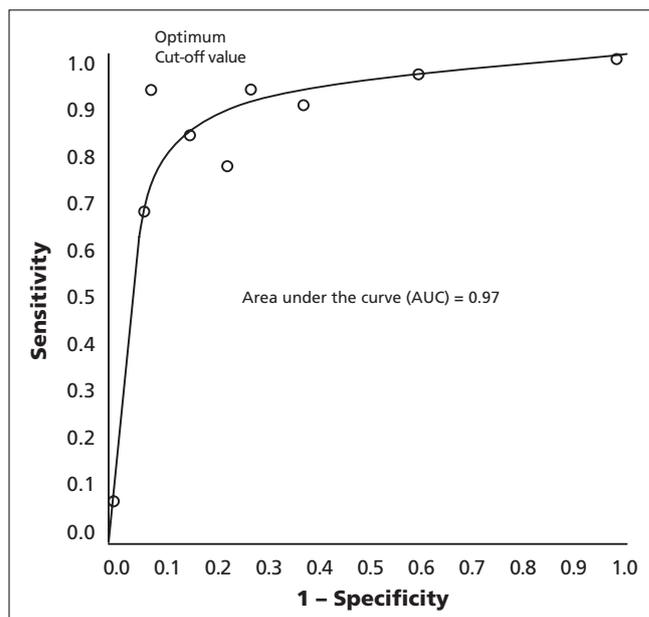


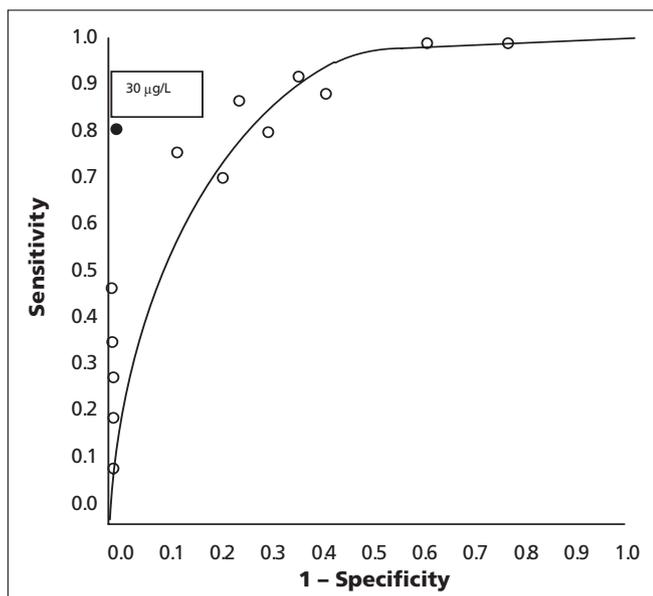
Fig. 1. Receiver operating characteristic curve illustrating high discriminatory power.

Division of Emergency Medicine, McMaster University, Hamilton, Ont.

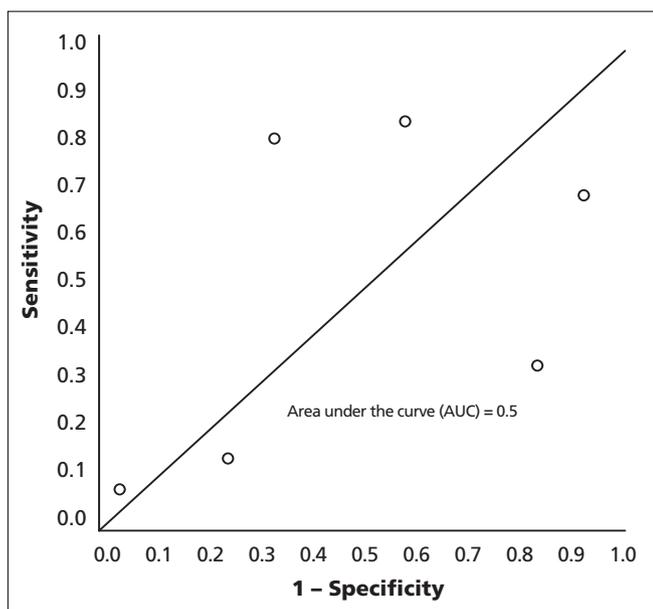
Received: Nov. 15, 2005; accepted: Nov. 22, 2005

*Can J Emerg Med* 2006;8(1):19-20

and 100% specific). An AUC value of 0.5 indicates no discriminatory value (i.e., 50% sensitive and 50% specific) and is represented by a straight, diagonal line extending from the lower left corner to the upper right (Fig. 3). There are several scales for AUC value interpretation but, in general, ROC curves with an AUC  $\leq 0.75$  are not clinically useful and an AUC of 0.97 has a very high clinical value, correlating with likelihood ratios of approximately 10 and 0.1. The AUC for NSE was 0.87, demonstrating moderate



**Fig. 2.** Receiver operating characteristic curve for the overall performance of neuron-specific enolase to predict survival at 48 hours after return of spontaneous circulation.



**Fig. 3.** Receiver operating characteristic curve illustrating no discriminatory power.

discriminatory power and, therefore potential utility as a diagnostic test in determining the non-survivors of a cardiac arrest with return of spontaneous circulation.

It is important to note that ROC performance may change when the diagnostic test is applied to different clinical situations (e.g., patient populations) or under different phases of test development (derivation, validation). The most useful information from a diagnostic test likely originates by pooling the results of several studies examining the same test in different situations, generating averaged specificity, sensitivity and ROC, so as to be able to get a true understanding of the diagnostic test's utility.<sup>6</sup>

In summary, ROC analysis provides important information about diagnostic test performance: the closer the apex of the curve toward the upper left corner, the greater the discriminatory ability of the test (i.e., the true-positive rate is high and the false-positive [1 - Specificity] rate is low). This is measured quantitatively by the AUC such that a value of  $>0.96$  indicates excellent discriminatory ability. Like all summary measures, however, there are confidence intervals around this value that must be taken into consideration. In the end, it will be rare for a diagnostic test to have both 100% specificity and sensitivity. The clinician will have to decide which cut-off value will provide the likelihood ratios and sensitivity and specificity values that have the greatest clinical value in the diagnosis of any disorder.

**Competing interests:** None declared.

**Key words:** receiver operating characteristic curve; diagnosis; sensitivity; specificity

## References

1. Auer J, Berent R, Weber T, et al. Ability of neuron-specific enolase to predict survival to hospital discharge after successful cardiopulmonary resuscitation. *Can J Emerg Med* 2006;8(1):13-8.
2. Lusted LB. Decision-making studies in patient management. *N Engl J Med* 1971; 284:416-24.
3. Grzybowski M, Younger JG. Statistical methodology: III. Receiver operating characteristic (ROC) curves. *Acad Emerg Med* 1997;4:818-26.
4. Worster A, Innes G, Abu-Laban RB. Diagnostic testing: an emergency medicine perspective. *Can J Emerg Med* 2002;4(5): 348-54.
5. Faraggi D, Reiser B. Estimation of the area under the ROC curve. *Stat Med* 2002;21:3093-106.
6. Deeks J. Systematic reviews of evaluations of diagnostic and screening tests. In: Egger M, Smith GD, Altman DG, editors. *Systematic reviews in health care: meta-analysis in context*. BMJ Publishing Group; 2001.

**Correspondence to:** Dr. Andrew Worster, Emergency Department, Hamilton Health Sciences, McMaster University Medical Centre, 1200 Main St. W, Hamilton ON L8N 3Z5