A brief note about likelihood ratios

Andrew Worster, MD, MSc; Christopher Carpenter, MD, MSc

In this issue of the journal, Azzam and colleagues review the validation study of a clinical prediction rule for termination of resuscitation in patients suffering an out-of-hospital cardiac arrest. The clinical prediction rule is similar to a diagnostic test in that it predicts whether or not the patient will suffer the disorder of interest, in this case death, before hospital discharge. Given this, the results can be expressed using summary measures that are typically used for diagnostic efficacy. The most familiar of these are sensitivity and specificity, terms that have been adopted into clinical medicine from laboratory research on assays for the detection of substances. Application of the sensitivity to the clinical prediction rule refers to the percentage of patients that die before discharge who met the clinical prediction rule criteria. Clinically, this is counterintuitive since what we really want to know is what percentage of those who met the criteria will die before being discharged. This is the positive predictive value. The greatest problem with predictive values is that they are dependent on the prevalence of the disorder in the population in which the measurements are performed; therefore, they have limited generalizability, that is, limited clinical applicability outside of the study population.

The solution to both of these problems, and perhaps the best single measure of a diagnostic test’s strength, is the likelihood ratio (LR). As its name implies, the LR is a ratio of probabilities: the probability of a given test result (all clinical prediction rule criteria, met or not) among people who died divided by the probability of that test result among people who survived. The LR employs a mathematical relationship known as Bayes Theorem to indicate how much a given diagnostic test result will raise or lower the pretest probability of the patient having the disorder in question by multiplying the pretest odds by the LR. This is typically calculated using a nomogram or Web-based program (http://araw.med.uic.edu/cgi-bin/testcalc.pl) but can easily be done by simply understanding the diagnostic strength of different LR values. Because the LR is a ratio, a value of 1 indicates that the numerator is the same as the denominator and multiplying the pretest probability by 1 yields a posttest probability of the same value, hence, an LR of 1 has no diagnostic value. As a ratio, the diagnostic test (the clinical prediction rule) has greater predictive value as the LR moves away from 1 toward ∞ or toward 0 such that an LR of less than 2 and an LR of greater than 0.5 are each of little clinical value but an LR of greater than 10 and an LR less than 0.1 are both considered conclusive. It should be noted that positive test results will often be expressed as a positive LR (LR+) and a negative test result as a negative LR (LR–). This is based on the assumption that the test result is dichotomous and overlooks one of the advantages of the LR: it can be used for test results with multiple categories. On the other hand, sensitivity and specificity assume that a diagnostic result is dichotomous. In fact, the interval LR is the slope between 2 points on a receiver operating characteristic curve.

Azzam and colleagues have used the LR to express the results of the termination of resuscitation rule with different...
pretest probabilities of death and, as we saw, there is little impact on the posttest probability in part due to the inconclusive LR. However, the LR is of greatest diagnostic value when the pretest probability is in the range of 30%–70%. In the face of a diagnostic dilemma with a pretest probability of 50%, the same LR of 6.6 reported by Azzam and coauthors will significantly increase the posttest probability to more than 80% and might be considered conclusive. A perceived disadvantage of the LR is the need to determine a pretest probability from which to launch one’s Bayesian analysis. Although pretest probability estimates may be evidence-based, when confronted with the life-or-death decision of not applying cardiopulmonary resuscitation, emergency physicians must ensure that the baseline cardiac arrest survival rates within their populations approximate those in the derivation study before they can confidently apply the yet-to-be-validated termination of resuscitation clinical prediction rule. If baseline survival rates are higher within an individual community, emergency medicine clinicians and prehospital health care providers would fail to attain equally low posttest probabilities within their settings.

Perhaps, as Azzam and colleagues have indicated, the measure that emergency medical services directors and other stakeholders are likely most interested in is not the sensitivity, specificity, predictive value or even the LR of the clinical prediction rule, but simply the percentage or proportion of incorrect termination of resuscitation.

Competing interests: None declared.

References


Correspondence to: Dr. Christopher R. Carpenter, Division of Emergency Medicine, Evidence Based Medicine, Washington University School of Medicine in St. Louis, Box 8072, 660 S. Euclid Ave., St. Louis MO 63110; carpenterc@wusm.wustl.edu