Signal Detection Theory Optimized Risk Cut-Points as a Function of Grade and Substance

Authors: Jason Machan, Ph.D., P. Allison Minugh, Ph.D., Nicoletta Lomuto, M.A., Joseph Drop, M.S.

Presenter: Kelly R. Breeden, M.S.

Introduction

Current convention in prevention needs assessment is to report prevalence rates of individuals defined to be "at risk" by their scores on risk scales. These scales are generally continuous measures, or aggregate scores amounting to continuous measures. This requires the researcher to select a cut point at which to separate individuals who are "at risk" from those who are "not at risk". The methods by which cut points are determined vary by study. Among the methods reported are the mid-point of the scale, algorithms based on the distribution of the scale scores, and algorithms based on both scales scores and outcomes. The event for which an individual is considered to be at risk also varies. Often, the event is either unstated or a broad aggregate of several outcomes (e.g. substance use and anti-social behaviors). This paper presents the method used by DATACORP as part of survey of students conducted in Alabama and compares the results with those obtained using other methods. The principles and methods of Signal Detection Theory were used to filter uninformative risk scales and to select cut-points optimized for equal sensitivity and specificity for predicting lifetime use of different substances.

Methods

Sample: Approximately 93,000 Alabama 6th thru 12th grade students.

Midpoint Method: constant across all scales, substances and grades.

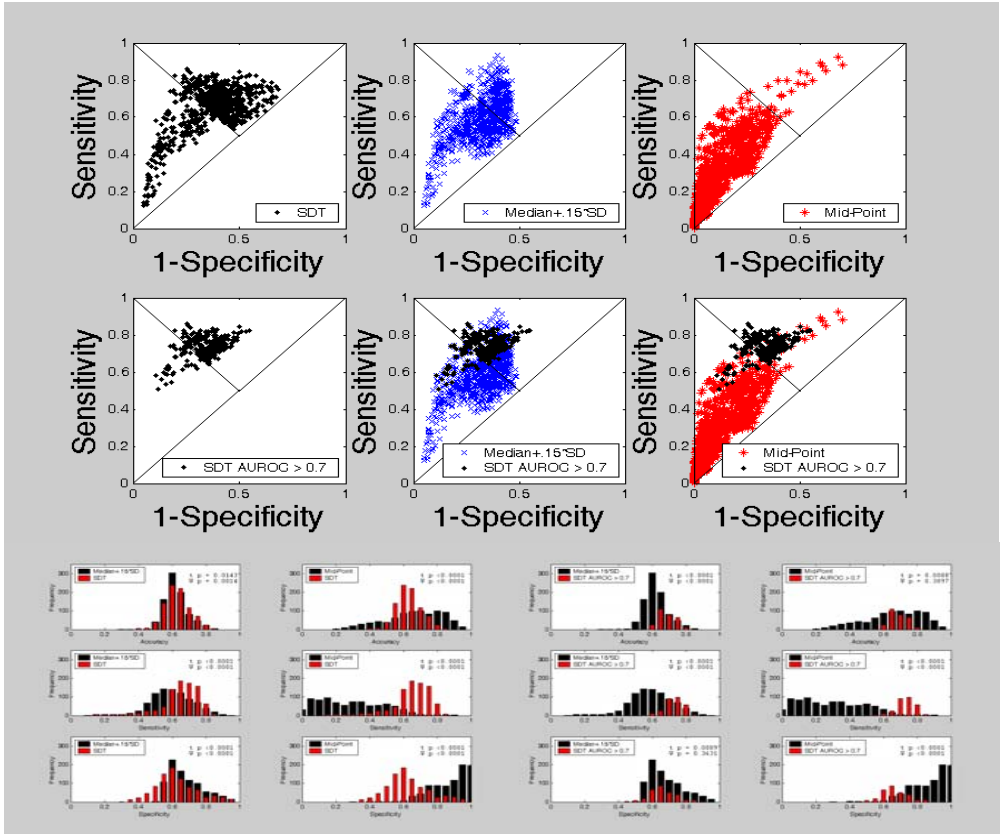
Median-based Method: calculated for each scale x grade; constant across substances.

SDT Method: calculated for each scale x substance x grade. Optimized cut-points were based on the point in the risk scale where sensitivity and specificity were (most) equal.

Area Under the ROC curve (AUROC): strength of the relationship between the risk scale and probability of lifetime use of the substance within a grade. Scale x outcome x grade combinations were considered informative if the AUROC exceeded .70.

Figures: The top six figures plot the sensitivity as a function of 1-specificity for each cut point method. The further points lie from the longer diagonal line, the higher the sensitivity and/or specificity. The closer points lie to the shorter diagonal line, the more balanced the sensitivity and specificity.

The bottom twelve figures show the distributions and significance tests comparing filtered and unfiltered SDT to median and mid-point based cut points on accuracy, sensitivity and specificity. The t-test and Wilcoxon rank-sum statistics are both given.



Results

Mid-point based cut-points were overly specific relative to their sensitivity. Median-based cut-points were more balanced, but individual cut-points tended to be either more sensitive or specific. SDT cut-points, particularly those filtered, exhibited better balanced sensitivity and specificity.

Both filtered and unfiltered SDT optimized cut-points had higher sensitivities than either alternative methods. Specificities were lower for unfiltered SDT than median-based cut-points and comparable when filtered. SDT specificities were lower than mid-point based, regardless of filtering. Accuracies were higher for SDT cut-points than for median based. They were lower but more consistent for SDT than for mid-point based.

There were a total of 1092 combinations of grade, risk scale, and substance (7 grades x 6 substances x 26 risk scales). Of these, 798 (73%) had AUROCs of less than or equal to .70 indicating that they did not discriminate well between at risk and not at risk individuals.

Discussion

SDT offers a method for creating cut points in risk scales that are directly related to and optimized for distinguishing between the risk of specific events. It assures that planners can control the balance between sensitivity and specificity. It also offers a method for selecting scales with the strongest relationship to the chosen outcomes (AUROC). The importance of this filter was demonstrated by the finding that for 73% of the scale x outcome x grade combinations, the risk scale was not informative in discriminating between individuals at risk and those not at risk.

Another strength of this method is that it permits for differential relationships between different scales and different outcomes (i.e., different cut-points for different substances). While the median-based cut points were calculated by grade, they were not able account for differences in the relationship between risk scales and outcomes.