On August 9, 2011, SCOR SE, a global reinsurer with offices in more than 31 countries, acquired substantially all of the life reinsurance business, operations and staff of Transamerica Reinsurance, the life reinsurance division of the AEGON companies. The business of Transamerica Reinsurance will now be conducted through the SCOR Global Life companies, and Transamerica Reinsurance is no longer affiliated with the AEGON companies.

While articles, treaties and some historic materials may continue to bear the name Transamerica, AEGON is no longer producing new reinsurance business.

Archive Materials

Receiver Operating Characteristic Curves

Originally published online in January 2010

By David Wesley, Vice President, Medical Research and Development

Receiver operating characteristic curves (ROCs) are a popular way to summarize the resolution power of a diagnostic test. ROCs plot the trade-off between a test’s sensitivity and its false positive rate (1- specificity) as the cutoff point that defines a positive test is varied.

An ideal test would have 100-percent sensitivity and 100-percent specificity while the ROC for a test that provides no information at all would follow a 45-degree line. ROCs and a related measure, area under curve (AUC), are ubiquitous in reports about tests and are often the focus when discussing the relative merits of different tests. However, ROCs are fairly useless.

ROCs are literally all about sensitivity and specificity - nothing else. So they tell us nothing about positive predictive value (need prior probability or prevalence) and nothing about protective value or cost/benefit (ditto plus need outcome costs as well). Even for comparing diagnostic tests, their usefulness is limited.

Figure 1 displays the ROCs for two tumor markers, CA 19-9 and CA 125, when used to detect pancreatic cancer. These are empirical results from actual data.
Focusing on CA 125, if the cutoff for the definition of a positive test is set at zero, the sensitivity is 100 percent but 1 - specificity is also 100 percent, and the corresponding point on the ROC curve is at the upper right corner. As the cutoff is increased through the range of possible values for CA 125, both the sensitivity and 1 - specificity decrease (specificity increases).

From Figure 1 it is apparent that CA 19-9 performs better than CA 125 – its ROC is further from the 45-degree line and closer to an ideal plot. The way to quantify this fact is with the AUC.

An ideal test would have an AUC of 1.0 while the non-informative test has an AUC of 0.5. Many researchers use AUCs as the key criteria for gauging test performance; i.e., the test with the higher AUC is the better test. This is not necessarily the case, because “the AUC is not clinically relevant.”1 What does this mean?

Figure 2 compares the ROC and AUC for a tumor marker I’ve named logit (actually, it’s a logistic combination of both CA 19-9 and CA 125). Logit has a higher AUC than CA 19-9 alone, but is it truly superior?

Figure 2

The clinician or Medical Director is likely to focus on only the portion of an ROC where test performance meets a certain requirement -- for example, a high specificity. In the region of high specificity (see red line for Sp = 90 percent) for these two ROCs, there is no reason to prefer logit over CA 19-9. Incidentally, these tumor markers actually do not have a role in screening for pancreatic CA primarily because of the low prevalence of the disease.
AUCs are computationally equivalent to the “c-statistic,” a measure of resolution in risk prediction models (e.g., the Framingham equation). This fact has contributed to unfortunate misunderstandings and controversies.\(^2\,^3\)

ROCs and AUCs are appropriate where one is detecting the actual presence of a condition or predicting the occurrence of an event. They are not appropriate for use in risk models that predict the probability of an event.

Even with diagnostic test literature where ROC publication seems obligatory, the ROC provides less information than you would think.

Say I want to use CA 19-9 for the detection of pancreatic CA, and I want to have a specificity of 10 percent. The ROC in Figure 1 tells me the corresponding sensitivity is 75 percent, but I cannot determine the corresponding cutpoint from the ROC. From the data, I know it is about 40 U/mL.

Someone recently asked me what to do with ROCs, and I replied, “take them with a grain of salt.” Perhaps better advice would be to ignore them completely, because ROCs are often more misleading than informative.

Sources: