5.2 Cutoff Point and Its Effects on Sensitivity and Specificity

We have been discussing sensitivity and specificity as characteristic of a diagnostic test; however, they can be modified by the choice of the *cutoff point between normal and abnormal*. For example, we may want to diagnose patients as hypertensive or normotensive by their diastolic blood pressure. Let us say that anyone with a diastolic pressure of 90 mmHg or more will be classified as "hypertensive." Since blood pressure is a continuous and variable characteristic, on any one measurement, a usually nonhypertensive individual may have a diastolic blood pressure of 90 mmHg or more, and similarly a truly hypertensive individual may have a single measure less than 90 mmHg. With a cutoff point of 90 mmHg, we will classify some nonhypertensive individuals as hypertensive, and these will be false positives. We will also label some hypertensive individuals as normotensive and these will be false negatives. If we had a more stringent cutoff point, say, 105 mmHg, we would classify fewer nonhypertensives as hypertensive since fewer normotensive individuals would have such a high reading (and have fewer false positives).

However, we would have more false negatives (i.e., more of our truly hypertensive people might register as having diastolic blood pressure less than 105 mmHg on any single occasion). These concepts are illustrated in Figure 5.3.

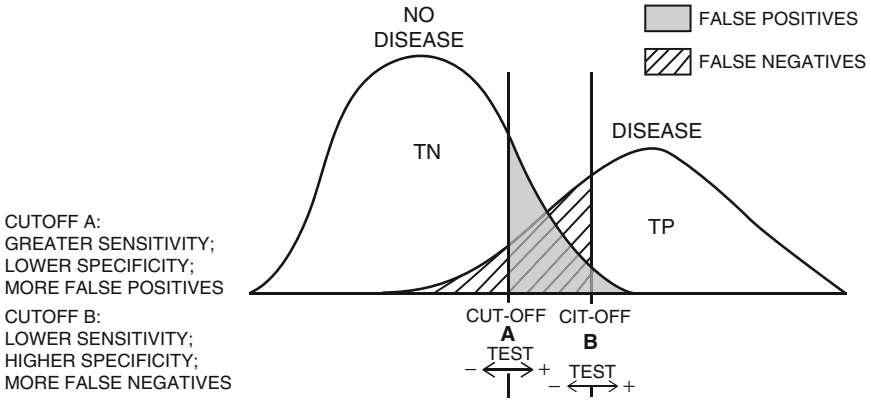

Figure 5.3 Different test cutoff points and false positives and false negatives

There are two population distributions, the diseased and nondiseased, and they overlap on the measure of interest, whether it is blood pressure, blood glucose, or other laboratory values. There are very few screening tests that have no overlap between normal and diseased individuals.

One objective in deciding on a cutoff point is to strike the proper balance between false positives and false negatives. As you can see in Figure 5.3, when the cutoff point is at A, all values to the right of A are called positive (patient is considered to have the disease). In fact, however, the patient with a value at the right of cutoff A could come from the population of non-diseased people, since a proportion of people who are perfectly normal may still have values higher than those above A, as seen in the normal curve. The area to the right of A under the no-disease curve represents the false positive.

If an individual has a test value to the left of cutoff A, he may be a true negative or he may be a false negative because a proportion of individuals with the disease can still have values lower than cutoff A. The area under the "disease" curve to the left of cutoff A represents the proportion of false negatives.

If we move the cutoff point from A to B, we see that we decrease the area to the right of the cutoff, thereby decreasing the number of false positives but increasing the number of false negatives. Correspondingly, with cutoff A, we have a greater probability of identifying the truly diseased correctly, that is, pick up more true positives, thereby giving the test with cutoff A greater sensitivity. With cutoff B, we are less likely to pick up the true positives (lower sensitivity) but more likely to correctly identify the true negatives (higher specificity).

Thus, by shifting the cutoff point beyond what we call a test positive, we can change the sensitivity and specificity characteristics of the test. The choice of cutoff, unless there is some special physiological reason, may be based on consideration of the relative consequences of having too many false positives or too many false negatives. In a screening test for cancer, for example, it would be desirable to have a test of high sensitivity (and few false negatives), since failure to detect this condition early is often fatal. In a mass screening test for a less serious condition or for one where early detection is not critical, it may be more desirable to have a high specificity in order not to overburden the health-care delivery system

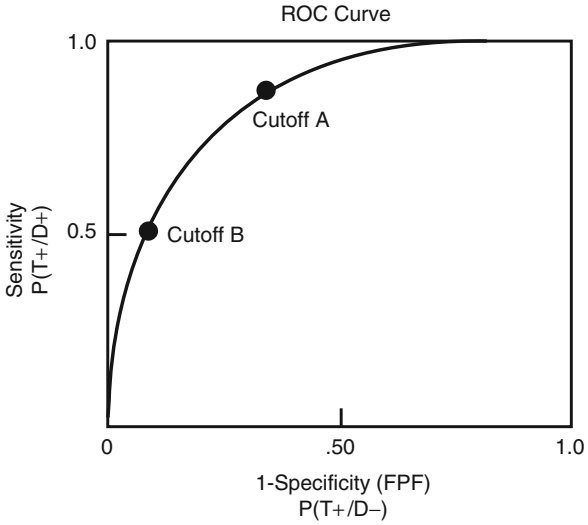

Figure 5.4 Receiver operating characteristic (ROC) curve, sensitivity versus false positive fraction

with too many false positives. Cost consideration may also enter into the choice of cutoff point.

The relationship between sensitivity (the ability to correctly identify the diseased individuals) and the false-positive fractions is shown in Figure 5.4.

This is called the receiver operating characteristic (ROC) curve of the test. Often we can select the cutoff point between normal and abnormal, depending on the trade-off we are willing to make between sensitivity and the proportion of false positives.

We can see that with cutoff A, while we can detect a greater percentage of truly diseased individuals, we will also have a greater proportion of false-positive results, while with cutoff B we will have fewer false positives but will be less likely to detect the truly diseased. Screening tests should have corresponding ROC curves drawn.