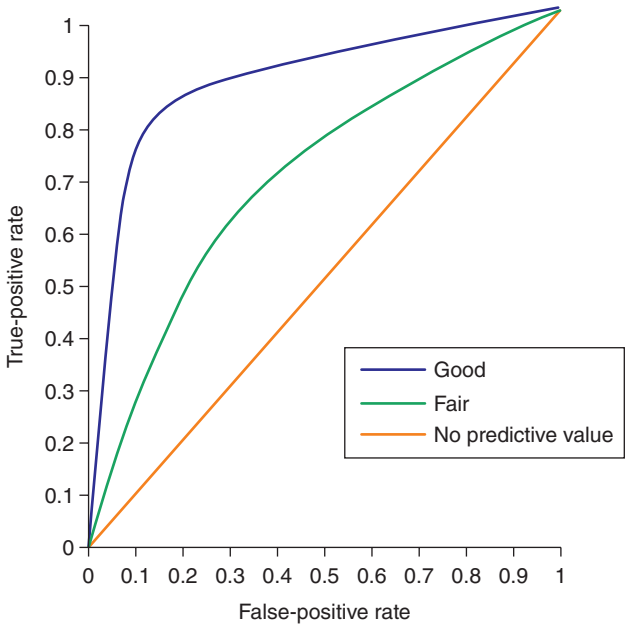

FIGURE 3-1 Each receiver operating characteristic (ROC) curve illustrates a trade-off that occurs between improved test sensitivity (accurate detection of patients with disease) and improved test specificity (accurate detection of patients without disease), because the test value defining when the test turns from "negative" to "positive" is varied. A 45° line would indicate a test with no predictive value (sensitivity = specificity at every test value). The area under each ROC curve is a measure of the information content of the test. Thus, a larger ROC area signifies increased diagnostic accuracy.

incidence. For some common conditions, such as CAD, nomograms and statistical models generate estimates of pretest probability that account for history, physical examination, and test findings. The posttest probability (also called the predictive value of the test) is a revised statement of the likelihood of the diagnosis, accounting for both pretest probability and test results. For the likelihood of disease following a positive test (i.e., positive predictive value), Bayes' rule is calculated as:

$$\text{Posttest probability} = \frac{\text{Pretest probability} \times \text{Sensitivity}}{\text{Pretest probability} \times \text{Sensitivity} + (1 - \text{Pretest probability}) \times \text{False-positive Rate}}$$

For example, with a pretest probability of 0.50 and a "positive" diagnostic test result (test sensitivity = 0.90 and specificity = 0.90):

$$\text{Posttest probability} = \frac{0.50 \times 0.90}{0.50 \times 0.90 + (1 - 0.50) \times 0.10} = 0.90$$

The term *predictive value* often is used as a synonym for the posttest probability. Unfortunately, clinicians commonly misinterpret reported predictive values as intrinsic measures of test accuracy. Studies of diagnostic tests compound the confusion by calculating predictive values on the same sample used to measure sensitivity and specificity. Since all posttest probabilities are a function of the prevalence of disease in the tested population, such calculations may be misleading unless the test is applied subsequently to populations with the same disease prevalence. For these reasons, the term *predictive value* is best avoided in favor of the more informative *posttest probability* following a positive or a negative test result.

The nomogram version of Bayes' rule (Fig. 3-2) helps us to conceptually understand how it estimates the posttest probability of disease. In this nomogram, the impact of the diagnostic test result is summarized by the *likelihood ratio*, which is defined as the ratio of the probability of a given test result (e.g., "positive" or "negative") in a patient with disease to the probability of that result in a patient without disease, thereby providing a measure of how well the test distinguishes those with from those without disease.

For a positive test, the likelihood ratio positive is calculated as the ratio of the true-positive rate to the false-positive rate (or sensitivity/[1 – specificity]). For example, a test with a sensitivity of 0.90 and a specificity of 0.90 has a likelihood ratio of 0.90/(1 – 0.90), or 9. Thus, for this hypothetical test, a "positive" result is nine times more likely in a patient with the disease than in a patient without it. Most tests in medicine have likelihood ratios for a positive result between 1.5 and 20. Higher values are associated with tests that more substantially increase the posttest likelihood of disease. A very high likelihood ratio positive (exceeding 10) usually implies high specificity, so a positive high-specificity test helps "rule in" disease. If sensitivity is excellent but specificity is less so, the likelihood ratio will be reduced substantially (e.g., with a 90% sensitivity but a 55% specificity, the likelihood ratio is 2.0).

For a negative test, the corresponding likelihood ratio negative is the ratio of the false-negative rate to the true-negative rate (or [1 – sensitivity]/specificity). Lower likelihood ratio values more substantially lower the posttest likelihood of disease. A very low likelihood ratio negative (falling below 0.10) usually implies high sensitivity, so a negative high-sensitivity test helps "rule out" disease. The hypothetical test considered above with a sensitivity of 0.9 and a specificity of 0.9 would have a likelihood ratio for a negative test result of (1 – 0.9)/0.9, or 0.11, meaning that a negative result is about one-tenth as likely in patients with disease than in those without disease (or 10 times more likely in those without disease than in those with disease).

APPLICATIONS TO DIAGNOSTIC TESTING IN CAD

Consider two tests commonly used in the diagnosis of CAD: an exercise treadmill test and an exercise single-photon emission CT (SPECT) myocardial perfusion imaging test (Chap. 270e). Meta-analysis has shown that a positive treadmill ST-segment response has an average sensitivity of 66% and an average specificity of 84%, yielding a likelihood ratio of 4.1 (0.66/[1 – 0.84]) (consistent with small discriminatory ability because it falls between 2 and 5). For a patient with a 10% pretest probability of CAD, the posttest probability of disease after a positive result rises to only about 30%. If a patient with a pretest probability of CAD of 80% has a positive test result, the posttest probability of disease is about 95%.

In contrast, exercise SPECT myocardial perfusion test is more accurate for CAD. For simplicity, assume that the finding of a reversible exercise-induced perfusion defect has both a sensitivity and a specificity of 90%, yielding a likelihood ratio for a positive test of 9.0 (0.90/[1 – 0.90]) (consistent with moderate discriminatory ability because it falls between 5 and 10). For the same 10% pretest probability patient, a positive test raises the probability of CAD to 50% (Fig. 3-2). However, despite the differences in posttest probabilities between these two tests (30% versus 50%), the more accurate test may not improve diagnostic likelihood enough to change patient management (e.g., decision to refer to cardiac catheterization) because the more accurate test has only moved the physician from being fairly certain that the patient did not have CAD to a 50:50 chance of disease. In a patient with a pretest probability of 80%, exercise SPECT test raises the posttest probability to 97% (compared with 95% for the exercise treadmill). Again, the more accurate test does not provide enough improvement in posttest confidence to alter management, and neither test has improved much on what was known from clinical data alone.

In general, positive results with an accurate test (e.g., likelihood ratio positive 10) when the pretest probability is low (e.g., 20%) do not move the posttest probability to a range high enough to rule in disease (e.g., 80%). In screening situations, pretest probabilities are often particularly low because patients are asymptomatic. In such cases, specificity becomes particularly important. For example, in screening first-time female blood donors without risk factors for HIV, a positive test raised the likelihood of HIV to only 67% despite a specificity of 99.995% because the prevalence was 0.01%. Conversely, with a high pretest