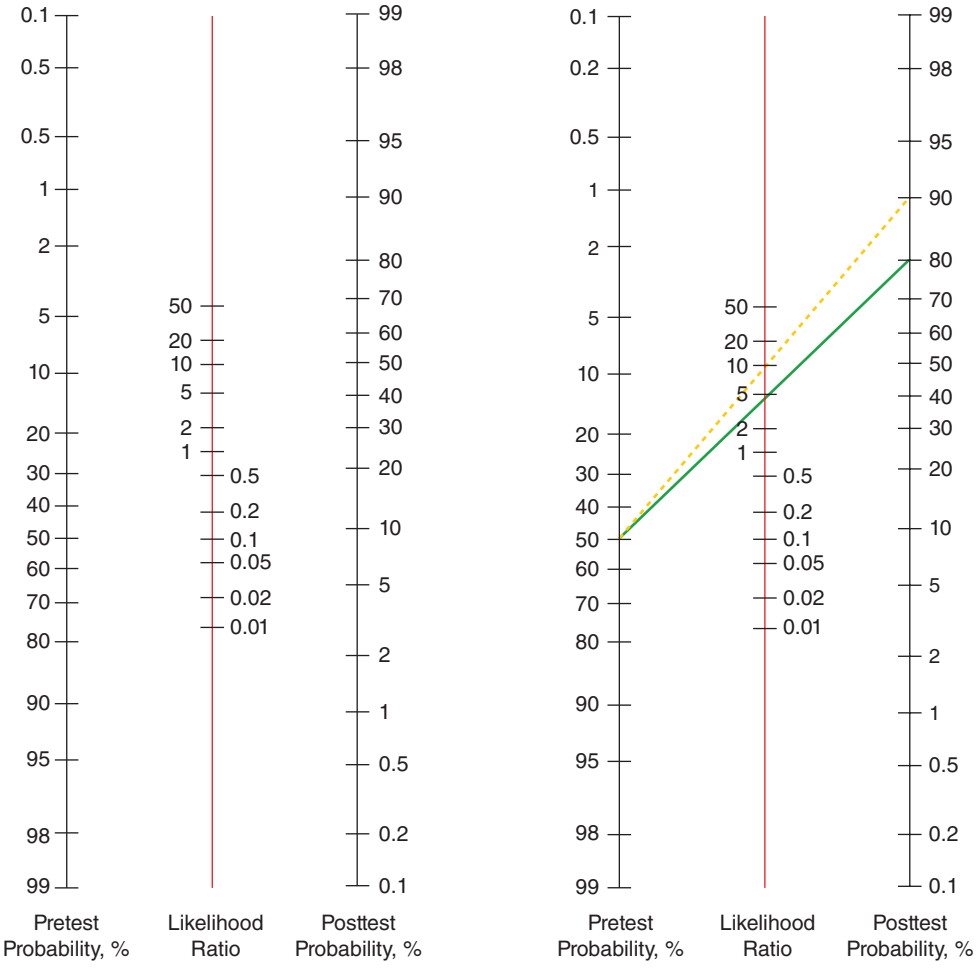

FIGURE 3-2 **Nomogram version of Bayes' rule** used to predict the posttest probability of disease (right-hand scale) using the pretest probability of disease (left-hand scale) and the likelihood ratio for a positive test (middle scale). See text for information on calculation of likelihood ratios. To use, place a straight edge connecting the pretest probability and the likelihood ratio and read off the posttest probability. The right-hand part of the figure illustrates the value of a positive exercise treadmill test (likelihood ratio 4, green line) and a positive exercise thallium single-photon emission computed tomography perfusion study (likelihood ratio 9, broken yellow line) in a patient with a pretest probability of coronary artery disease of 50%. *(Adapted from Centre for Evidence-Based Medicine: Likelihood ratios. Available at http://www.cebm.net/index.aspx?o=1043.)*

probability, a negative test may not rule out disease adequately if it is not sufficiently sensitive. Thus, the largest change in diagnostic likelihood following a test result occurs when the clinician is most uncertain (i.e., pretest probability between 30% and 70%). For example, if a patient has a pretest probability for CAD of 50%, a positive exercise treadmill test will move the posttest probability to 80% and a positive exercise SPECT perfusion test will move it to 90% (Fig. 3-2).

As presented above, Bayes' rule employs a number of important simplifications that should be considered. First, few tests have only positive or negative results, and many tests provide multiple outcomes (e.g., ST-segment depression and exercise duration with exercise testing). Although Bayes' rule can be adapted to this more detailed test result format, it is computationally more complex to do so. Similarly, when multiple tests are performed, the posttest probability may be used as the pretest probability to interpret the second test. However, this simplification assumes conditional independence—that is, that the results of the first test do not affect the likelihood of the second test result—and this is often not true.

Finally, it has long been asserted that sensitivity and specificity are prevalence-independent parameters of test accuracy, and many texts still make this statement. This statistically useful assumption, however, is clinically simplistic. A treadmill exercise test, for example, has a sensitivity in a population of patients with one-vessel CAD of around 30%, whereas its sensitivity in patients with severe three-vessel CAD approaches 80%. Thus, the best estimate of sensitivity to use in a particular decision may vary, depending on the severity of disease in the local population. A hospitalized, symptomatic, or referral population typically has a higher prevalence of disease and, in particular, a higher prevalence of more advanced disease than does an outpatient population. Consequently, test sensitivity will likely be higher in hospitalized patients, and test specificity higher in outpatients.

STATISTICAL PREDICTION MODELS

Bayes' rule, while illustrative as presented above, provides an unrealistically simple solution to most problems a clinician faces. Predictions based on multivariable statistical models, however, can more accurately address these more complex problems by accounting for specific patient characteristics. In particular, these models explicitly account for multiple possibly overlapping pieces of patient-specific information and assign a relative weight to each on the basis of its unique contribution to the prediction in question. For example, a logistic regression model to predict the probability of CAD considers all the relevant independent factors from the clinical examination and diagnostic testing and their significance instead of the limited data that clinicians can manage in their heads or with Bayes' rule. However, despite this strength, prediction models are usually too complex computationally to use without a calculator or computer (although this limitation may be overcome once medicine is practiced from a fully computerized platform).

To date, only a handful of prediction models have been validated properly (for example, Wells criteria for pulmonary embolism) (Table 3-2). The importance of independent validation in a population