thereby perpetuating their overuse, e.g., annual cardiac exercise testing in asymptomatic patients.

Practice Setting Factors Factors in this category relate to the physical resources available to the physician's practice and the practice environment. *Physician-induced demand* is a term that refers to the repeated observation that once medical facilities and technologies are made available to physicians, they will use them. Other environmental factors that can influence decision-making include the local availability of specialists for consultations and procedures; "high-tech" advanced imaging or procedure facilities such as MRI machines and proton beam therapy centers; and fragmentation of care.

Economic Incentives Economic incentives are closely related to the other two categories of practice-modifying factors. Financial issues can exert both stimulatory and inhibitory influences on clinical practice. In general, physicians are paid on a fee-for-service, capitation, or salary basis. In fee-for-service, physicians who do more get paid more, thereby encouraging overuse, consciously or unconsciously. When fees are reduced (discounted reimbursement), doctors tend to increase the number of services provided to maintain revenue. Capitation, in contrast, provides a fixed payment per patient per year to encourage physicians to consider a global population budget in managing individual patients and ideally reducing the use of interventions with small marginal benefit. In contrast to inexpensive preventive services, however, this type of incentive is more likely to affect expensive interventions. To discourage volume-based excessive utilization, fixed salary compensation plans pay physicians the same regardless of the clinical effort expended, but may provide an incentive to see fewer patients.

INTERPRETATION OF DIAGNOSTIC TESTS IN THE CONTEXT OF DECISION-MAKING

Despite the great technological advances in medicine over the last century, uncertainty remains a key challenge in all aspects of medical decision-making. Compounding this challenge is the massive information overload that characterizes modern medicine. Today's clinician needs access to close to 2 million pieces of information to practice medicine. According to one estimate, doctors subscribe to an average of seven journals, representing over 2500 new articles each year. Of course, to be useful, this information must be sifted for applicability to and then integrated with patient-specific data. Although computers appear to offer an obvious solution both for information management and for quantification of medical care uncertainties, many practical problems must be solved before computerized decision support can be routinely incorporated into the clinical reasoning process in a way that demonstrably improves the quality of care. For the present, understanding the nature of diagnostic test information can help clinicians become more efficient users of such data. The next section reviews important concepts related to diagnostic testing.

DIAGNOSTIC TESTING: MEASURES OF TEST ACCURACY

The purpose of performing a test on a patient is to reduce uncertainty about the patient's diagnosis or prognosis in order to facilitate optimal management. Although diagnostic tests commonly are thought of as laboratory tests (e.g., blood count) or procedures (e.g., colonoscopy or bronchoscopy), any technology that changes a physician's understanding of the patient's problem qualifies as a diagnostic test. Thus, even the history and physical examination can be considered a form of diagnostic test. In clinical medicine, it is common to reduce the results of a test to a dichotomous outcome, such as positive or negative, normal or abnormal. Although this simplification ignores useful information (such as the degree of abnormality), such simplification does make it easier to demonstrate the fundamental principles of test interpretation discussed below.

The accuracy of diagnostic tests is defined in relation to an accepted "gold standard," which defines the presumably true state of the patient (Table 3-1). Characterizing the diagnostic performance of a new test requires identifying an appropriate population (ideally, patients in whom the new test would be used) and applying both the new and the gold standard tests to all subjects. Biased estimates of test performance may occur from using an inappropriate population or from incompletely applying the gold standard test. By comparing the two tests, the characteristics of the new test are determined. The *sensitivity* or *true-positive rate* of the new test is the proportion of patients with disease (defined by the gold standard) who have a positive (new) test. This measure reflects how well the new test identifies patients with disease. The proportion of patients with disease who have a negative test is the *false-negative rate* and is calculated as 1 – sensitivity. Among patients without disease, the proportion who have a negative test is the *specificity,* or *true-negative rate*. This measure reflects how well the new test correctly identifies patients without disease. Among patients without disease, the proportion who have a positive test is the *false-positive rate*, calculated as 1 – specificity. A perfect test would have a sensitivity of 100% and a specificity of 100% and would completely distinguish patients with disease from those without it.

TABLE 3-1 MEASURES OF DIAGNOSTIC TEST ACCURACY

	Disease Status	
Test Result	**Present**	**Absent**
Positive	True positive (TP)	False positive (FP)
Negative	False negative (FN)	True negative (TN)
Test Characteristics in Patients with Disease		
True-positive rate (sensitivity) = TP/(TP + FN)		
False-negative rate = FN/(TP + FN)		
True-positive rate = 1 – false-negative rate		
Test Characteristics in Patients without Disease		
True-negative rate (specificity) = TN/(TN + FP)		
False-positive rate = FP/(TN + FP)		
True-negative rate = 1 – false-positive rate		

Calculating sensitivity and specificity requires selection of a threshold value or cut point above which the test is considered "positive." Making the cut point "stricter" (e.g., raising it) lowers sensitivity but improves specificity, whereas making it "laxer" (e.g., lowering it) raises sensitivity but lowers specificity. This dynamic trade-off between more accurate identification of subjects with disease versus those without disease is often displayed graphically as a receiver operating characteristic (ROC) curve (Fig. 3-1) by plotting sensitivity (*y* axis) versus 1 – specificity (*x* axis). Each point on the curve represents a potential cut point with an associated sensitivity and specificity value. The area under the ROC curve often is used as a quantitative measure of the information content of a test. Values range from 0.5 (no diagnostic information from testing at all; the test is equivalent to flipping a coin) to 1.0 (perfect test). The choice of cut point should depend on the relative harms and benefits of treatment for those without versus those with disease. For example, if treatment was safe with substantial benefit, then choosing a high-sensitivity cut point (upper right of the ROC curve) for a low-risk test may be appropriate (e.g., phenylketonuria in newborns), but if treatment had substantial risk for harm, then choosing a high-specificity cut point (lower left of the ROC curve) may be appropriate (e.g., amniocentesis that may lead to therapeutic abortion of a normal fetus). The choice of cut point may also depend on the likelihood of disease, with low likelihoods placing a greater emphasis on the harms of treating false-positive tests and higher likelihoods placing a greater emphasis on missed benefit by not treating false-negative tests.

MEASURES OF DISEASE PROBABILITY AND BAYES' RULE

Unfortunately, there are no perfect tests. After every test is completed, the true disease state of the patient remains uncertain. Quantifying this residual uncertainty can be done with Bayes' rule, which provides a simple way to calculate the likelihood of disease after a test result or posttest probability from three parameters: the pretest probability of disease, the test sensitivity, and the test specificity. The pretest probability is a quantitative estimate of the likelihood of the diagnosis before the test is performed and is usually the prevalence of the disease in the underlying population although occasionally it can be the disease