The Impact of Disease Prevalence on the Predictive Value of Laboratory Tests in Primary Care

H. Thomas Wiegert, MD, and Olga Wiegert, MPA
Lexington, Kentucky

The *predictive value* of any diagnostic process is related to the *sensitivity* and *specificity* of the test and the *prevalence* of the underlying pathology for which the diagnostic testing is accomplished. Prevalence is the most important, but least understood, factor affecting the usefulness of a test result. For each accomplished test which is not helpful in either supporting a diagnosis or assisting in a differential diagnosis, the health care cost is increased without a corresponding benefit in the value of the information obtained.

The physician's interest in the usefulness of a diagnostic test depends upon how accurately the test predicts the presence or absence of disease. That is, if given a normal result, what are the chances that the patient is free of the disease and if given an abnormal result, what are the chances that disease is actually present? An additional consideration is whether the cost of doing a test yields an equivalent value in knowledge on which to base the management of the condition under consideration. If several tests can be accomplished to increase the probability of correctly determining the existence of a disease state, the physician must determine if the marginal increase in predictive value gained from doing each additional test is likely to result in a change in the management of the disease. The cost of doing the test must accrue equivalent expected benefit to the patient or society.

Probability calculations permit physicians to make better decisions about the efficacy of a laboratory test in confirming or excluding a diagnosis and provide a predictive value of a positive or negative test result.[1] The predictive value of test results is important to determine the probability of the presence or absence of the disease and to assess the extent of possible benefit, psychological trauma, or economic burden that could ensue from embarking on a particular mode of management.

McNeil et al,[2] in a cooperative study on the relative cost of finding patients with renovascular disease using the renogram, the intravenous pyelogram, or both, used Bayes theorem to calculate the a posteriori probabilities of disease at several cutoff points. The gain in the number of patients found with the additional use of the intravenous pyelogram, over the use of the renogram alone, was associated with a marginal cost of $18,708 per patient for surgically correctable

From the Department of Family Practice, University of Kentucky, College of Medicine, Lexington, Kentucky. Requests for reprints should be addressed to Dr. H. Thomas Wiegert, Department of Family Practice, MN-140D, University of Kentucky, College of Medicine, Lexington, KY 40536.

0094-3509/79/061199-05$01.25

hypertension. This is a high cost for the additional information gained.

The predictive value of a test is determined by the complex interaction of three variables: the incidence of true positive results in patients with the disease, the incidence of true negative results in subjects without disease, and the number of subjects with the disease in the group examined. In this interaction of variables, the prevalence of disease is far more important than common sense would suggest.[1] Tests established in tertiary care settings with a high prevalence of disease must be reevaluated in terms of the actual prevalence that exists in the setting in which the test is being used for diagnostic purposes. Indiscriminate use of laboratory tests with moderate or high false positive rates is not helpful in establishing the predictive value of a positive result, if the prevalence of the disease being tested for is low. Making certain that critical signs and symptoms are present for the diagnosis in question can effectively increase the prevalence of disease in the population selected for testing, and thereby increase the predictive value of a positive result (ie, the proportion of test positives that are true positives) of any test procedure applied to the selected population.

In 1763, the Reverend Thomas Bayes devised the conditional probability formula that permits the use of new information in the calculation of the predicted value of a positive test result. His formula is

$$p(\Theta_1/R) = \frac{p(\Theta_1)\, p(R/\Theta_1)}{p(\Theta_1)\, p(R/\Theta_1) + p(\Theta_2)\, p(R/\Theta_2)}$$

where $p(\Theta_1)$ is the a priori probability of a condition existing in a population; $p(\Theta_2)$ is the a priori probability of non-condition existing in a population; $p(R/\Theta_1)$ is the probability of a positive test result R, given that the subject does have the condition; $p(R/\Theta_2)$ is the probability of a positive test result R, given that the subject does not have the condition; and $p(\Theta_1/R)$ is the a posteriori probability of the condition Θ_1, given a positive test result R.[3]

Bayes formula for conditional probability can be used to determine the predictive value of a positive or negative test result under any set of prevalence conditions and hence can be used for comparing the predictive value of a test under different conditions. If a test is positive, the probability that the person actually has the condition under question can be predicted; and if a test is negative, the probability that the person does not have the disease can be predicted.

In the laboratory test context the predictive value of a positive test can be determined by the insertion of the appropriate values for the following characteristics of the test and setting:

$$\text{Predictive Value of a Positive Test} = \frac{(\text{Prevalence})\,(\text{Sensitivity})}{(\text{Prevalence})\,(\text{Sensitivity}) + (1\text{-Prevalence})\,(1\text{-Specificity})}$$

Sensitivity is the term used to characterize the *true positive (TP) results obtained when a test is applied to patients known to have the disease.* It is the percentage of all people with the disease who test positive. The formula is

$$\frac{(TP)}{(TP+FN)} \times 100 = \text{\_\_\_\_\_percent}$$

(FN equals false negative.) A good test gives a high percentage of positive results in diseased subjects.

Specificity is the term used to characterize the *true negative (TN) results obtained when a test is applied to persons known to be free of the disease.* It is the percentage of all disease-free people who test negative. The formula is

$$\frac{(TN)}{(TN+FP)} \times 100 = \text{\_\_\_\_\_percent}$$

(FP equals false positive.)

The *point prevalence rate* for a disease equals the number of patients who have the disease per 100,000 at the *time of the study. Prevalence* is a function of both the *incidence and the years of disease duration.* The incidence of chronic disease may be low, but its prevalence high. The incidence is low because only a small percentage of the population may contract the disease per year, but the prevalence is high because the disease may last for a lifetime in each subject.

When quantitative tests are used, it is possible to vary the number of true positives and true negatives by changing the level at which the test is considered positive. The cutoff level used can increase or decrease the sensitivity or the specificity. For some diagnostic circumstances, such as anemia, it is better to have a high specificity rather than a high sensitivity in order to avoid the misclassification of healthy subjects. For conditions that should not be missed, it is important to have a high sensitivity in order to avoid missing a treatable case. In qualitative tests, like pregnancy tests, sensitivity and specificity are inherent in the test system and cannot be varied by the laboratory.

In practice, tests do not have 100 percent sensitivity and 100 percent specificity at the same time, therefore the physician needs to decide what factors are important in the management of the patient. The physician's decision to use the test must be based on the prevalence of the disease, the severity of the disease, the cost of the test, the amenability of the disease to treatment, and the consequences of treatment.

Physicians must become knowledgeable about the circumstances in which it would be preferable to use tests that have certain characteristics. There are times when a test is of little value unless it has one or a combination of the following characteristics: high sensitivity, high specificity, high predictive value, or a high efficiency value.

Tests that have a high sensitivity, preferably 100 percent, should be used when the disease is serious and should not be missed, or when the disease is treatable. Disease conditions that fall into this classification are pheochromocytoma, phenylketonuria, venereal disease, and other treatable infectious diseases.

Tests that have a high specificity, preferably 100 percent, should be used when the disease is serious but is not treatable or when the knowledge that the disease is absent has psychologic or public health value. Disease conditions that fall into this category are multiple sclerosis and occult cancers. If the disease is in an early phase of its natural history, the patient will generally return due to continuing difficulty.

Tests that have a high predictive value for a positive result are essential in a circumstance in which the treatment of a false negative could have dire consequences. Lung cancer is a diagnosis for which the predictive value for the presence of the disease should be 100 percent because the treatment by lobectomy or radiation involves considerable risk for the patient.

Tests that have high efficiency are desired for disease that is treatable and for which false positive or false negative results are equally serious or damaging. Myocardial infarction may be fatal, but it is treatable. If it is missed, irreparable harm may be done; but nearly equal harm can result if a false positive diagnosis is made. Other examples of diseases that require the highest diagnostic efficiency include lupus erythematosus, diabetes mellitus, lymphoma, and some forms of leukemia.

Application

Illustrative examples of the effect of prevalence, sensitivity, and specificity on the predictive value of a positive test result are readily apparent from two commonly used tests in the family physician's office; the Monospot test and the throat culture.

One of the Monospot tests commonly used has a sensitivity of 94 percent and a specificity of 95 percent according to the manufacturer.* The test was developed where the prevalence rate was 45 percent, ie, 64 positive specimens and 79 controls from patients without evidence of the disease. Under these circumstances, the predictive value of a positive result is calculated as follows:

$$\text{Predictive Value of a Positive Test} = \frac{(0.45)\,(0.94)}{(0.45)\,(0.94) + (1 - 0.45)\,(1 - 0.95)} \times 100 = 94 \text{ percent}$$

However, when one applies the same Monospot test to a population such as that described by English and Geyman[4] where the incidence rate in a series of patients with symptoms compatible with infectious mononucleosis was 6.1 percent, the predictive value of a positive test becomes:

*Monosticon by Organson, West Orange, NJ

Predictive Value of a Positive Test =

$$\frac{(0.061)\,(0.94)}{(0.061)\,(0.94) + (1 - 0.061)\,(1 - 0.95)} \times 100 = 55 \text{ percent}$$

A positive result under these circumstances has the probability of being a false positive 45 percent of the time.

Next, consider the throat culture for B hemolytic streptococcal infection. The sensitivity of this test is 86 percent in the circumstances of aerobic incubation using the "stab" technique.[5] This assumes that two swabs are used.[6] The specificity of the throat culture is 80 percent in children[7,8] and 92 percent in adults.[9] If patients who present with sore throat, anterior adenopathy, exudate, and fever have throat cultures accomplished, the incidence for positive recovery of B hemolytic streptococci is 30.4 percent in children and 15.1 percent in adults.[10] This is the same order of magnitude reported by others.[7-9] However, the true infection as evidenced by rising ASO and anti-DNAse B titers reduces the prevalence rates to 13 percent in children[7,9] and to 5 percent in adults.[9,10]

The predictive value of a positive test using these data for children is:

Predictive Value of a Positive Result =

$$\frac{(0.13)\,(0.86)}{(0.13)\,(0.86) + (1 - 0.13)\,(1 - 0.80)} \times 100 = 39 \text{ percent}$$

The predictive value of a positive test using these data for adults is:

Predictive Value of a Positive Result =

$$\frac{(0.05)\,(0.86)}{(0.05)\,(0.86) + (1 - 0.95)\,(1 - 0.92)} \times 100 = 36 \text{ percent}$$

False positive results are probable 61 percent of the time in children and 64 percent of the time in adults.

The problem with interpretation of throat cultures relates to the relatively low prevalence of *true* group A beta hemolytic streptococcal *infections* and the relatively poor sensitivity if accomplished under aerobic conditions. The sensitivity is further reduced if only one swab is used, if the stab technique is not utilized, or if the "A" disc is not utilized. The low predictive value of a positive culture for beta hemolytic streptococci to indicate the presence of infection makes it a poor criterion for decisions regarding treatment. Komaroff,[6] Kaplan,[7] Forsyth,[10] and Breese[11] have outlined critical clinical signs and symptoms which, if present, are correlated with positive cultures for group A beta hemolytic streptococci. There is no perfect set of criteria on which to make a firm diagnosis of true group A beta hemolytic streptococcal infection.

As with all tests applied to conditions of low prevalence where the specificity is reasonably high, the predictive value of a negative result is good.

Discussion

Postgraduate medical education including primary care training is concentrated in tertiary care medical settings. Since many of the technological advances originate in tertiary care settings and the process of medical care taught in these settings is felt to constitute "good medicine," graduates of such programs regardless of practice setting emulate this process of medical care in their practices.

The "good medicine process" in the tertiary care setting is not ideal in a primary care setting for all decisions. Primary care physicians provide health care to a random population with unselected conditions which have prevalence and incidence rates lower than those found in tertiary care settings but greater than in the population at large. The different prevalence rates of disease in the different settings and the physician's estimate of disease frequency are important in decision making.[12] For example, in a tertiary care setting where the prevalence rate is 50 percent for the diagnosis in question, and the theoretical test has a 95 percent sensitivity and 95 percent specificity (a

good test), the predictive value of a positive test is 95 percent. However, in a primary care setting where the prevalence of the disease is one percent, the predictive value of a positive test is only 16 percent.

The medical student or resident who learns to use the "good test" in the tertiary care setting, but subsequently attempts to apply it uncritically in a primary care setting, is both perplexed and annoyed that the previously learned "good medicine process" is of little value to differential diagnosis or management of the disease under consideration. For each test ordered and accomplished which is not efficient in either supporting a diagnosis or assisting in a differential diagnosis, or which is ambiguous in assisting management, the health care cost is disproportionately increased to its marginal value.

One value of laboratory tests for the physician lies in his/her subsequent ability to sort patients into disease categories and to restore health or alleviate suffering for a particular level of cost.[13] It is a fallacy to believe that health care resources are not scarce and that people place health as a goal above all other goals; people are interested in adjusting their choices to the satisfaction they receive compared to the cost.[14] To reach equality at the margin, the last dollar spent on health care should give exactly the same increase in human satisfaction as if it were devoted to other goals. It is imperative that physicians have ready access to information on the comparative outcomes of health care decisions and the relative costs of those decisions.

Since most health care is given in primary and secondary care settings, the appropriate utilization of technology developed in the tertiary care setting can have a marked impact on cost of health care. There is an urgent need for educators and health care cost analysts to provide physicians with predictive values of laboratory tests under varying prevalence conditions, and methods by which to delineate the critical factors on which to base the decisions to use laboratory testing to aid in the diagnosis and management of disease.

The containment of the increasing costs of medical care has taken on a high priority. In a search to find ways to hold cost in line yet not reduce the quality of care, Bayesian decision theory is promulgated as a decision tool in cost effective or cost-benefit approaches to managing health care.[13-15] The application of Bayesian conditional decision analysis to the area of laboratory testing will assist in selecting those alternatives with the highest probability for desirable outcomes.[1,2,16] If physicians will estimate accurate probabilities of positive results before ordering diagnostic tests, the quality of medical care will improve and the cost of medical care for laboratory testing will decrease.

References

1. Galen S, Gambino SR: Beyond Normality: The Predictive Value and Efficiency of Medical Diagnosis. New York, John Wiley, 1975
2. McNeil J, Varady PD, Burrows BB, et al: Measures of clinical efficacy: Cost-effectiveness calculations in the diagnosis and treatment of hypertensive renovascular disease. N Engl J Med 293:216, 1975
3. Bayes T: An essay toward solving a problem in the doctrine of chance. Philo Trans R Soc 53:370, 1763
4. English EC, Geyman JP: The efficiency and cost effectiveness of diagnostic tests for infectious mononucleosis. J Fam Pract 6:977, 1978
5. McGonagle LA: Evaluation of a screening procedure for the isolation of beta hemolytic streptococci. Health Lab Sci 11:61, 1974
6. Komaroff L: A management strategy for sore throat. JAMA 239:1429, 1978
7. Kaplan EL, Top FH, Dudding BA, et al: Diagnosis of streptococcal pharyngitis: Differentiation of active infection from the carrier state in the symptomatic child. J Infect Dis 132:490, 1971
8. Quinn W, Denny W, Riley D: Natural occurrence of hemolytic streptococci in normal school children. Am J Public Health 47: 995, 1957
9. Commission on Acute Respiratory Diseases, Fort Bragg, NC: Role of B hemolytic streptococci in common respiratory disease. Am J Public Health 35:675, 1945
10. Forsyth RA: Selective utilization of clinical diagnosis in treatment of pharyngitis. J Fam Pract 2:173, 1975
11. Breese B: A simple scorecard for the tentative diagnosis of streptococcal pharyngitis. Am J Dis Child 131: 514, 1977
12. Barr D: Primary and specialized care: Implications for learning situations. J Med Educ 53: 176, 1978
13. McNeil J, Adelstein SJ: Measures of clinical efficacy: The value of case finding in hypertensive renovascular disease. N Engl J Med 293: 221, 1975
14. Fuchs VR: Who Shall Live? New York, Basic Books, 1974
15. Bunker JP, Barnes BA, Mosteller F: Costs, Risks and Benefits of Surgery. New York, Oxford University Press, 1977
16. Weinstein D, Statson WB: Foundations of cost effectiveness analysis for health and medical practices. N Engl J Med 296: 716, 1977