



A Comparison of C/B Ratios from Studies Using Receiver Operating Characteristic Curve Analysis

Scott B. Cantor,^{1,*} Charlotte C. Sun,² Guillermo Tortolero-Luna,²
Rebecca Richards-Kortum,³ and Michele Follen²

¹DEPARTMENT OF INTERNAL MEDICINE SPECIALTIES, SECTION OF GENERAL INTERNAL MEDICINE, AND ²DEPARTMENT OF GYNECOLOGIC ONCOLOGY, THE UNIVERSITY OF TEXAS M. D. ANDERSON CANCER CENTER, HOUSTON, TEXAS; AND ³DEPARTMENT OF COMPUTER AND ELECTRICAL ENGINEERING, THE UNIVERSITY OF TEXAS AT AUSTIN, AUSTIN, TEXAS

ABSTRACT. In receiver operating characteristic (ROC) curve analysis, the optimal cutoff value for a diagnostic test can be found on the ROC curve where the slope of the curve is equal to $(C/B) \times (1-p[D])/p[D]$, where $p[D]$ is the disease prevalence and C/B is the ratio of net costs of treating nondiseased individuals to net benefits of treating diseased individuals. We conducted a structured review of the medical literature to examine C/B ratios found in ROC curve analysis. Only two studies were found in which a C/B ratio was explicitly calculated; in another 11 studies, a C/B ratio was based on a so-called holistic estimate, an all-encompassing educated estimate of the relative costs and benefits relevant to the clinical situation. The C/B ratios ranged from 0.0025 (tuberculosis screening) to 2.7 (teeth restoration for carious lesions). Clinical scenarios that are directly life threatening but curable had C/B ratios of less than 0.05. This analysis led us to construct a table of ordered C/B ratios that may be used by investigators to approximate C/B ratios for other clinical situations in order to establish cutpoints for new diagnostic tests. *J CLIN EPIDEMIOL* 52;9:885–892, 1999. © 1999 Elsevier Science Inc.

KEY WORDS. ROC curve, sensitivity and specificity, decision support techniques, routine diagnostic tests, laboratory diagnosis, cost-benefit analysis

INTRODUCTION

A receiver operating characteristic (ROC) curve provides a visual comparison of the trade-offs between the true-positive rate (or sensitivity) and the false-positive rate (or one minus specificity) of a diagnostic test (Figure 1). ROC curve analysis may be used for three purposes: (1) to determine the discriminative ability of a diagnostic test, (2) to compare the discriminative abilities among several different diagnostic tests in order to identify the preferred one, and (3) to determine the optimal cutpoint of a diagnostic test. This article focuses on the third purpose.

When a single diagnostic test is available, a cutoff point or threshold distinguishes cases from noncases, i.e., those who will be treated from those who will not be treated. The optimal cutpoint of a diagnostic test is defined in the medical literature as the point at which the expected utility of a diagnostic test is maximized, where utility is a measure of

the strength of preference for an outcome. Metz [1] has shown that this optimal point on the ROC curve is the spot at which the slope R satisfies the following equation:

$$R = C/B \times (1 - p[D])/p[D]$$

where

C = net costs of treating nondiseased individuals,
 B = net benefits of treating diseased individuals,
and

$p[D]$ = prevalence of disease.

The first term of the Metz equation, the C/B ratio, represents the ratio of the net costs of treating nondiseased individuals to the net benefits of treating diseased individuals. The C/B ratio is expressed as a numerical ratio, without units, because it represents a relative weighting of advantages and disadvantages of treatment. For example, if the net costs of treating nondiseased individuals were to be three times as great as the net benefits of treating diseased individuals, then the C/B ratio would be equal to 3.

The C/B ratio can be dissected into its component parts regarding the eventual outcomes of the test result. Two perspectives for evaluating net costs to net benefits are available, depending on whether one wants to view outcomes as either “costs” (in a negative frame) or as “utilities” (in a positive

*Address correspondence to: Scott B. Cantor, Ph.D., The University of Texas M. D. Anderson Cancer Center, Department of Internal Medicine Specialties, Section of General Internal Medicine, 1515 Holcombe Blvd., Box 40, Houston, TX 77030-4095.

This paper was presented, in part, at the 18th Annual Meeting of the Society for Medical Decision Making, Toronto, Ontario, Canada, October 1996. Accepted for publication on 2 April 1999.

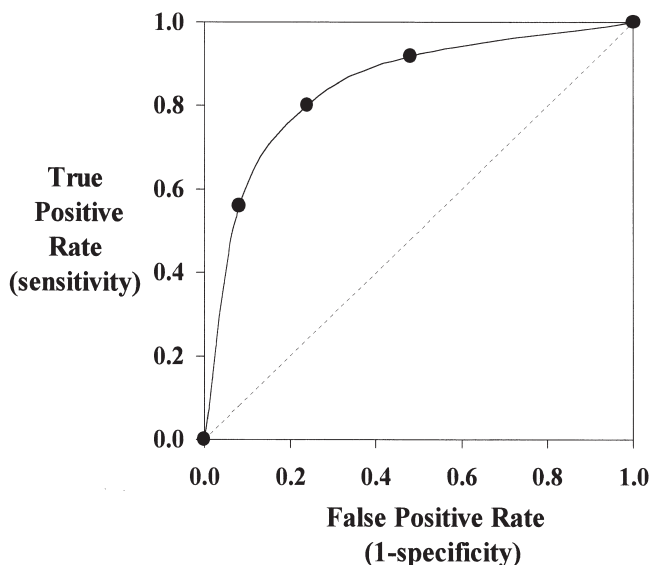


FIGURE 1. A receiver operating characteristic (ROC) curve, demonstrating the tradeoff between the true-positive rate (sensitivity) and the false-positive rate (1-specificity) for a diagnostic test. The 45-degree diagonal line represents a test that does not provide any useful discriminative information between diseased and nondiseased patients.

frame). In the first perspective, “cost” is a negative outcome measure, referring to the negative effects in terms of monetary cost, adverse health risks, or a combination of the two. The C/B ratio based on costs can be determined by formal methods (e.g., clinical decision analysis or cost-benefit analysis) or by a subjective estimate. The following is the framework developed by Metz [1] and Weinstein and Fineberg [2]. In general:

$$C/B = (C_{TN} - C_{FP}) / (C_{TP} - C_{FN})$$

where:

C_{TN} = cost of a true-negative result; typically best outcome of not having disease

C_{FP} = cost of a false-positive result; inappropriate (over)treatment

C_{TP} = cost of a true-positive result; appropriate treatment

C_{FN} = cost of a false-negative result; typically worst outcome of missed case

In the second perspective, outcomes can be viewed in terms of utilities, “utility” being a positively framed outcome measure that refers to monetary savings, health benefits, or a combination of the two. It should be noted that “utility” refers to a generally positive outcome measure and not necessarily a utility defined in the von Neumann-Morgenstern [3] framework as an outcome evaluated using the standard gamble method. The paradigm of using utilities in the ratio of net costs to net benefits was developed by Sox *et al.* [4]:

$$C/B = (U_{TN} - U_{FP}) / (U_{TP} - U_{FN})$$

where:

U_{TN} = utility of a true-negative result; typically best outcome of not having disease

U_{FP} = utility of a false-positive result; inappropriate treatment

U_{TP} = utility of a true-positive result; appropriate treatment

U_{FN} = utility of a false-negative result; typically worst outcome of missed case

The second term in the Metz equation involves the prevalence of disease. For diseases of low prevalence (which typically result in more false positives than true positives) and for conditions in which a false positive would result in painful or dangerous testing or treatment, the analyst should select a cutoff point that yields fewer false positives. Such a cutoff point should be chosen from the segment of the ROC curve in the lower-left quadrant of the ROC plot in Figure 1. A line tangent to a point of the ROC curve selected from the lower-left quadrant of the ROC plot has a relatively steep slope. While this cutoff point yields fewer false positives, it also results in a trade-off of more false negatives.

Conversely, for highly prevalent diseases and for conditions in which the number of false-negative test results should be minimized, the analyst should select a cutoff point that limits the number of false negatives. Such a point should be chosen from the segment of the ROC curve in the upper-right quadrant of the ROC curve plot in Figure 1. A line tangent to a point of the ROC curve selected from the upper-right quadrant of the ROC plot has a relatively flatter slope than tangents drawn for other points on the ROC curve. However, such a cutoff point would yield more false-positive results [5].

For diseases in which treatment or testing is toxic to nondiseased patients but has little chance of curing diseased patients, the C/B ratio is moderate to large. Again, this results in a steep slope, and both true positives and false positives are minimized. Likewise, the C/B ratio is small to moderate for conditions in which treatment or testing is relatively benign for nondiseased patients and presents an excellent chance of cure for diseased patients. This results in a flatter slope in which the number of true positives is as large as the number of false positives [1].

We undertook this study as a part of a larger project to evaluate the cost-effectiveness of strategies for the diagnosis and management of patients with cervical precancerous conditions, in which the results of the newer technology of fluorescence spectroscopy were compared with those of the usual care of colposcopic examination. In the present study, we wished to examine how researchers assigned or determined a C/B ratio in the management of clinical problems in order to determine the optimal cutpoint along a ROC curve for a diagnostic test. Our goal was to find and compare studies that formally justified an optimal cutpoint using a C/B ratio within a ROC curve analysis.

METHODS

Published reports of clinical and diagnostic applications of ROC curve analysis were reviewed for this inventory. One of the investigators (CCS) conducted a literature search using MEDLINE for all reports published between 1976 and 1997, inclusive. (We made sure to include 1978 because that was the year of Metz's landmark publication on ROC curve analysis.) Articles that met the following selection criteria were chosen for initial review: (1) keyword = "ROC curve analysis"; and (2) textword = "cutoff" OR "criterion" OR "threshold." The articles' references were also examined for additional sources.

We noted the papers in which the C/B ratio was implicitly established by the authors' interpretation of the relative net costs of treating nondiseased individuals compared with the net benefits of treating diseased individuals. This assignment of a C/B ratio based on encompassing all attributes deemed appropriate, including clinical outcomes, economic impact, and psychological concerns, will be referred to as a "holistic" estimate, similar to the sense when Fischer [6] described an approach for simultaneously incorporating multiple attributes in the outcomes evaluation process. Alternatively, we noted when a C/B ratio was explicitly calculated, as in a cost-benefit analysis, which requires health outcomes to be valued in monetary units [7], or as part of a clinical decision analysis, which uses the maximization of a clinical outcome (e.g., survival probability, life expectancy, or quality-adjusted life expectancy) as the decision criterion.

We initially identified 377 abstracts that satisfied the MEDLINE search criteria. The number of published abstracts increased over the periods selected: 8 abstracts were found for the years 1976 to 1980; 41 abstracts were found for the years 1981 to 1986; 96 abstracts were found for the years 1987 to 1992; and 232 abstracts were found for the years 1993 to 1997. Based on the content of the abstracts, only 48 of the 377 identified abstracts were selected to be reviewed in detail; the remaining ones were not reviewed in detail because their abstracts did not indicate that a C/B ratio or an explicit method for determining a threshold was described.

To validate that the employed methodology for selecting articles was accurate, two investigators (SBC, GTL), different from the investigator who conducted the initial literature search, reviewed a random sample of 73 articles identified from the 329 abstracts not initially reviewed. This validation sample was undertaken to determine whether or not relevant articles had been inadvertently omitted in the original review process. An error rate of 5% was considered acceptable.

RESULTS

Of the 48 articles selected for review, including others found by checking secondary references, only 13 articles in-

cluded a C/B ratio as part of the ROC curve analysis for determining an optimal cutpoint. The validation sample yielded only one article that employed the Metz [1] methodology; it was a study of determining the optimal cutpoint levels of plasma des-gamma-carboxy prothrombin and serum alpha-fetoprotein for the diagnosis of hepatocellular carcinoma [8]. The prevalence of appropriate studies missed by the original search was therefore found to be 1.4% (1/73), leading us to conclude that our strategy was appropriate.

In the studies that did not include a C/B ratio as part of the ROC curve analysis, (including those from the validation sample as well as the ones from the initial review), the investigators used one of several different methods to determine a cutpoint:

- An arbitrary point was selected, with no justification or explanation given.
- The point on the ROC curve that was closest to the upper-left corner of the ROC plot was selected. (The point in the upper-left corner of the ROC plot represents a perfect test with 100% sensitivity and specificity.)
- A desired level of sensitivity or specificity was predetermined (or deemed acceptable), and the corresponding operating characteristic was found from the curve. For example, for a desired level of sensitivity equal to 0.95, the corresponding level of specificity could be determined from the curve
- The sum of sensitivity and specificity was maximized (see, for example, Guechot *et al.* [9]). This is equivalent to finding a point on the ROC curve whose tangent slope is equal to 1.
- The point at which sensitivity was equal to specificity was chosen (e.g., Arnese *et al.* [10]).

None of these methods of cutpoint selection has a theoretical foundation or scientific justification. None of these methods explicitly considers the risks and benefits of overtreatment and undertreatment nor the prevalence of disease in the clinical situation for which the diagnostic test is applicable.

Table 1 enumerates the papers that included a C/B ratio as part of their ROC curve analysis for determining an optimal cutpoint [2,4,5,8,11–21]. They are presented in order from smallest to largest C/B ratio, similar to the way clinical interventions in league tables of incremental cost-effectiveness ratios are presented from smallest to largest dollars per quality-adjusted life year gained [22]. The studies that provided a single C/B ratio are depicted in Figure 2.

The smallest C/B ratio found was 1/400 or 0.0025 for tuberculosis screening [11]. In this clinical scenario, the consequences of a missed case are serious and may potentially result in death. Consequently, the costs of overtreatment are small relative to the benefits of treating true cases. Although the overall sample size of 13 studies is small, we found the ratios to have consistency. Clinical scenarios that have direct life-threatening relevance, including tuberculo-

TABLE 1. Summary of C/B ratios in receiver operating characteristic curve analysis

Reference	Clinical scenario	Country/ perspective	C/B ratio	C/B method	C/B basis
Lusted [11]	Tuberculosis testing	USA/physician	1/400 = 0.0025	Holistic	Physician value judgments and dollar costs
Gohagen <i>et al.</i> [12,13] ^a	Mammography and palpation for breast screening (primary care)	USA/patient	1/1000 = 0.001 1/100 = 0.01	Holistic	External, psychosocial, and dollar costs
Sox <i>et al.</i> [4]	Screening for curable infectious disease of serious consequence (primary care setting)	USA/patient	1/50 = 0.02	Holistic	No details provided
Hagen [5]	Stress test and three-vessel coronary disease (primary care)	USA/patient	0.2/5.4 = 0.037	Explicit	Costs = percentage of acute mortality; benefits = percentage of improved survival
Weinstein and Fineberg [2]	Schiotz tonometry and glaucoma (in ophthalmology clinic)	USA/patient	1/20 = 0.05	Holistic	No details provided
DeBaun and Sox [14]	Lead poisoning in children (primary care setting)	USA/societal	(168-63)/(6159-4361) = 0.058	Explicit	Cost-benefit analysis including direct and indirect (by human capital approach) dollar costs
Milman and Albeck [15]	Tests for hereditary haemochromatosis	Denmark/patient	1/100 = 0.01 1/10 = 0.10 1/1 = 1.00	Holistic	No details provided
Bergus [16]	Urinary tract infection (primary care setting)	USA/patient	1/1.5 = 0.6667	Holistic	C/B derived from clinical experience
England [17]	CT scan images of patients with neurological problems (based on Hanley and McNeil [18])	USA/patient	(0-3000)/(3000-6000) = 1.0	Holistic	No details provided
Hdez-Armas <i>et al.</i> [19]	Gammagraphic diagnosis of hepatic cirrhosis	Spain/patient	1/1 = 1.0	Holistic	"Simplified cost-benefit analysis" but no details provided
Fombonne [20]	Child behavior checklist for psychiatric research to measure child psychiatric disorders	France/various	1/1 = 1.0	Holistic	Risk/benefit of treatment financial cost, and discomfort
Fujiyama <i>et al.</i> [8]	Plasma DCP and AFP in diagnosis of liver cancer	Japan/patient	1/1	Holistic	No details provided
Sox <i>et al.</i> [4]	Confirmation of incurable infectious disease of serious consequence (primary care setting)	USA/patient	2/1 = 2.0	Holistic	No details provided
Kay and Knill-Jones [21]	Dentists' decisions to restore teeth deemed positive for carious lesions	Scotland/patient	260/96 = 2.7	Holistic ^b	Value judgment left up to dentist

C/B methods: holistic—assignment of C/B ratio based on encompassing all attributes deemed appropriate by investigators (e.g. clinical outcomes, economic impact, etc.); explicit—C/B ratio calculation based on explicit evaluation of health outcomes using cost-benefit or clinical decision analysis. Note that some scenarios and diagnostic tests suggested several possible C/B ratios.

^aFor illustrative purposes, Gohagen *et al.* [13] calculated cutoff points for varying C/B ratios and disease rates, and they used cost estimates cited in their previous work.

^bHolistic estimate based on retrospective examination of operating characteristics.

sis testing [11], breast cancer screening [12,13], testing for curable fatal infectious disease [4], and stress testing for three vessel coronary disease [5], all had C/B ratios lower than 1/20 or 0.05.

The largest C/B ratio found was 260/96 (or 2.7) for teeth restoration necessitated by carious lesions [21]. The clinical situations shown on the bottom portion of Table 1 (child

behavior checklist [20], computed tomography scan for neurological disorders [17], and confirmation of disease without cure [4]) are not immediately life threatening or have little chance of cure when the individual tested actually has the disease or condition of interest. Compared with interventions in some of the other cases, smaller or even no clinical benefit would be gained from detecting disease in

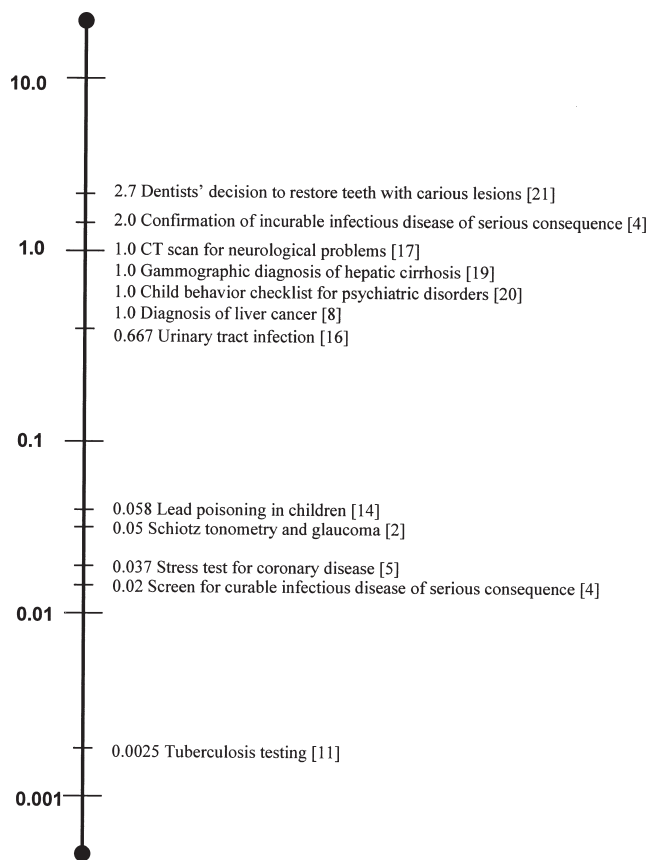


FIGURE 2. Studies that cited a C/B ratio as plotted on a logarithmic scale (references in parentheses). Note: Studies from Table 1 presenting multiple C/B ratios are not included in the figure for purposes of clarity.

these patients. The C/B ratios in these cases are all greater than or equal to 1.

In only two studies were explicit analyses performed that identified the four “costs” of the four diagnostic outcomes: true positives, true negatives, false positives, and false negatives. DeBaun and Sox [14] performed a cost–benefit analysis to determine the optimal threshold for erythrocyte protoporphyrin screening for lead poisoning. In this cost–benefit analysis, the total costs of the possible outcomes of lead poisoning screening were determined. In this context, the total cost referred to the sum of direct and indirect costs. Direct costs referred to medical costs; indirect costs referred to the reduction in future earnings of a child who has either mental retardation or a learning disability secondary to lead poisoning. The Metz equation was used to determine the slope of the tangent of the ROC curve where the optimal cutpoint should be; in this case, the monetary costs of the outcomes formed the basis for the C/B ratio.

In a review article on ROC curve analysis, Hagen [5] performed a simplified clinical decision analysis using 1-year mortality rate as outcome measure for determining the optimal cutpoint for electrocardiographic stress testing for

three-vessel coronary disease. Hagen [5] argued that the net cost of stress testing in normal patients is the 0.2% risk of acute mortality from coronary angiography, used when an abnormal stress test occurs. The net benefits of stress testing consist of the improved survival for true-positive patients: medically treated three vessel coronary disease confers a 9.0% excess annual mortality, with bypass surgery decreasing the coronary-related excess annual mortality by 1.3%. However, bypass surgery has an acute surgical mortality risk of 2.3%. Thus, the net 1-year benefit for stress testing is $(9.0\% - 2.3\%) - 1.3\%$ or 5.4%. The C/B ratio is therefore equal to $0.2\%/5.4\%$, or 0.037. Given the baseline prevalence of three-vessel coronary disease in men, the Metz equation could then be applied to determine the point on the ROC curve with the appropriate slope that would indicate the optimal cutpoint for electrocardiographic stress testing for three-vessel coronary disease.

In the other studies listed in Table 1, few details were provided in reference to how the C/B ratio was determined. Various factors, such as the relative cost of diagnostic errors compared with the value of correct diagnoses . . . [including] dollar costs” [11], “clinical experience” [16], “risk/benefits from treatment, financial costs, discomfort, etc.” [20], or “views about . . . the long-term outcomes of any treatment prescribed . . . [and] beliefs about the sequelae of untreated disease” [21], may have been incorporated in the determination of the C/B ratio. However, no further details were provided; the reader must infer any additional information. For example, in the Bergus study [16], a cost–benefit analysis based on the dollar cost of office visits and antibiotic treatment for urinary tract infection would coincidentally lead to the same C/B ratio as presented in the paper ($2/3$) (personal communication). However, in most of these studies, no details were provided at all.

In the majority of these studies, the authors would write something to the effect that “the costs of treatment of non-diseased patients were x times as great as the net benefits of treatment of diseased patients.” In these cases, the C/B ratios were simple to calculate. For example, in the Bergus [16] study, the author states that the net benefits of treatment were 1.5 times as great as the net costs based on a holistic estimate—therefore, the C/B ratio was calculated as $1/1.5$, or 0.667. In many of the studies, as shown in Figure 2, the authors stated that the net costs of treatment were equivalent to the net benefits; therefore, the C/B ratio would be calculated as $1/1$ or simply, 1. Perhaps this assumption was made for purpose of simplicity; in any case, no detailed rationale was provided to the reader. Thus, the 11 studies other than the DeBaun and Sox [14] and Hagen [5] studies identified in Table 1 recognized the importance of the C/B ratio but made a holistic rather than an explicit estimate of the ratio of net costs to net benefits.

To determine the C/B ratio for fluorescence spectroscopy in a referral setting, we compared the possible outcomes of diagnostic testing for cervical precancerous lesions in a

noncompliant referral population with the other clinical testing scenarios as presented in Table 1. Cervical cancer is a serious disease, but much greater survival benefits are achieved when detected early in its precancerous stages. Thus, diagnostic testing for cervical precancer is more like the clinical scenarios in the top portion of the table. Therefore, we estimated that the appropriate range for the C/B ratio regarding the diagnosis of cervical precancers is between 0.01 and 0.05. This range of estimates can then be used in the Metz equation to determine the optimal cutpoint for fluorescence spectroscopy in a referral setting.

DISCUSSION

The recent emphasis on cost-conserving medical practice [23] and incorporation of patients' values in clinical decision making [24] has emphasized the need to justify the cutoff or threshold values in ROC curve analysis for diagnostic tests. Several articles in the literature have highlighted the formula for the identification of an optimal cutoff point based on the slope of the ROC curve. In the Metz equation, the prevalence $p[D]$ can usually be identified in a relatively straightforward manner. In contrast, the C/B component requires researchers identify explicitly costs or utilities of four possible outcomes: true positives, false positives, true negatives, and false negatives.

The Metz [1] model that is used in this article is a generic model of determining optimal cutpoints on a ROC curve. Although it is not typically explicitly stated in these types of diagnostic studies, the model of choice of a cutpoint is typically applied to an "average" patient. The specific amount of net benefit of treatment for diseased patients or net cost for nondiseased patients depends on a variety of factors. For example, a physician may prefer to perform cardiac catheterization on a 30-year-old patient, compared with a 90-year-old patient, even if both have the same posterior probability of coronary artery disease, because the life expectancy for a successfully treated patient in the former case is much greater. Similarly, the severity of disease may lead to a different estimate of the C/B ratio, again because of the variations in the valuations of the possible outcomes. Thus, we recognize that a "one cutoff does it all" approach may not be appropriate. To accommodate particular or specialized patient populations, an adjustment of the C/B ratio, in either the previously published studies or for use in future studies, may not only be appropriate, but desirable as well.

We have classified the results of our literature search for methodologies for determining optimal diagnostic cutpoints into four categories of analysis. The literature describes the development of many diagnostic tests that seem to pick arbitrary cutoffs for the determination of a "positive" result. The commercial manufacturers of these tests seem to be the least rigorous in their methodology for choosing a cutpoint.

The second category of analysis is used by a minority of researchers who seem to understand the theoretical foundations of ROC curve analysis and even plot the ROC curve but who then claim that the best operating point on the curve is the one that is closest to the point of perfection, one that is on the line perpendicular to the 45-degree line of no diagnostic information. Some investigators have used other criteria for determining the "optimal" cutpoint: acceptance of a preset level of sensitivity or specificity and determining the corresponding operating characteristic, maximizing the sum of sensitivity plus specificity, or finding the point at which sensitivity is equal to specificity. None of these methods takes into consideration the prevalence of disease or the consequences of correctly or incorrectly classifying a test as positive or negative.

The third category of analysis is used by investigators who chose the operating point based on the Metz equation but chose the C/B ratio based on their own clinical experience rather than arbitrary opinion. Although no explicit justification is given for the C/B ratio chosen, the analyses based on these choices have shown some level of consistency.

Finally, the fourth category of analysis for choice of an optimal cutpoint was found only in two studies [5,14]. The investigators using this approach performed a quantitative cost-benefit or clinical decision analysis to determine explicitly costs and benefits of appropriate and inappropriate treatment. This fourth kind of analysis is typically the most rigorous, difficult, and time consuming, but the explicit nature of the analysis makes it the most appealing.

We recognize that many investigators do not encourage the dichotomization of diagnostic test results. Centor [25] recommended the use of likelihood ratios and Bayesian prior-to-posterior analysis to determine the revised probability of disease after the results of a diagnostic test are learned. This was the approach taken in the use of the CAGE questionnaire to diagnose a history of alcohol abuse [26]. Diagnostic tests are still being developed, however, that recommend the use of cutpoints that produce dichotomous "test-positive" and "test-negative" results. Hence, the framework we describe should be highly relevant.

We embarked on this study to help determine what the optimal cutpoint should be for the diagnosis of cervical intraepithelial neoplasia when using fluorescence spectroscopy. Fluorescence spectroscopy uses laser technology to distinguish normal from precancerous tissues [27]. By varying the threshold of classification, pairs of sensitivity and specificity values are established. We determined an appropriate estimate for the ratio of the net costs of treating women without cervical intraepithelial neoplasia to the net benefits of treating women with cervical intraepithelial neoplasia (defined as the C/B ratio) by comparing the clinical situation with the other entries in the C/B summary table. By using this C/B estimate and the prevalence of cervical intraepithelial neoplasia in our referral population, we

can now determine an optimal cutoff point for this innovative clinical technique. We recognize, however, that a detailed, though more complicated, explicit cost-benefit analysis or clinical decision analysis is preferred for determining a C/B ratio. We intend to analyze the diagnostic utility of fluorescence spectroscopy using the most rigorous methodology.

Over the past 20 years, an increasing number of studies involving diagnostic tests and the use of ROC curve analysis have appeared in MEDLINE. This may be attributed to the increasing number of publications catalogued in the literature database. A more likely explanation is that rapid advances in medical science have resulted in increased demand for both development and evaluation of diagnostic tests. With these technological advances, investigators must strive for the appropriate interpretation and classification of such diagnostic tests. Furthermore, the clinical situations for which these tests are applied have become more serious, both from the policy and the individual patient's perspectives. For example, the downstream consequences of diagnostic tests that involve genetic screening for a patient's inherited predisposition for cancer underscore the importance of using the appropriate methodologies for evaluating the classification of a positive test result. These consequences have implications ranging from insurance discrimination to prophylactic surgical interventions to prevent cancer that may develop many years in the future.

An important aspect of this article may be to assist investigators who want to follow the third category of analysis of justification in determining an optimal cutoff point. Our summary table (Table 1) of C/B ratios for many different clinical situations is sorted by smallest to largest C/B ratio. This table may be a guide for determining a C/B ratio for a new diagnostic test to be used in a particular clinical situation. One could then compare the new diagnostic test in the particular clinical situation (e.g., primary care or specialty setting) with the established tests. If similarities between the new test-setting situation and the established test-setting situations can be found, then one could use the same C/B ratios (or range of ratios) to assign a C/B ratio for the new test-setting situation.

For this purpose, however, the table should be used with caution. Only two of the cited articles actually performed explicit analyses of the costs and benefits of the diagnostic outcomes, and these are subject to their own critiques. Some investigators who could not determine a C/B ratio for their particular test clinical scenario suggested a range of ratios of several orders of magnitude. Finally, in the majority of studies in which all-encompassing estimates of the C/B ratio were made, little or no information was provided on how the estimates were determined. Nevertheless, it was interesting to note that there was a general trend regarding the relative order of many of these testing situations; the generalizability of the data remains to be determined, as does the theoretical and scientific basis of selecting the cutpoint.

Meanwhile, however, researchers who want to be explicit in their methodology can use our table of C/B ratios as a guide to assess the net costs and benefits of overtreatment and undertreatment compared to previously evaluated clinical situations. As shown in the Metz equation, the selected C/B ratio, multiplied by the reciprocal of the odds of disease in the designated setting, yields a number that should be equal to the slope of the tangent to the ROC curve at its most optimal operating point. Given the uncertainty that usually surrounds the estimation of costs and benefits, a sensitivity analysis that incorporates variations in the C/B ratio can be performed to see how the computed slope and subsequent choice of optimal operating point will be affected. Use of this analytical methodology should lead to better clinical management decisions and enhance the overall health of the community.

This study was funded, in part, by support from the National Science Foundation/The Whitaker Foundation "Integration of Optical Engineering, Efficacy and Cost Analyses to Improve Diagnosis of Cervical Pre-Cancer." This manuscript was improved by suggestions made by J. Robert Beck, Robert Centor, Dennis Fryback, Michael Kattan, Lore Feldman, Walter Pagel, and colleagues in the Section of General Internal Medicine of The University of Texas M. D. Anderson Cancer Center.

References

1. Metz CE. Basic principles of ROC analysis. *Semin Nucl Med* 1978; 8: 283-298.
2. Weinstein MC, Fineberg HV. **Clinical Decision Analysis**. Philadelphia: W. B. Saunders; 1980.
3. von Neumann J, Morgenstern O. **Theory of Games and Economic Behavior**. Princeton: Princeton University Press; 1953.
4. Sox HC, Blatt MA, Higgins MC, Marton KI. **Medical Decision Making**. Boston: Butterworths; 1988.
5. Hagen MD. Test characteristics: How good is that test? *Prim Care* 1995; 22: 213-233.
6. Fischer GW. Utility models for multiple objective decisions: Do they accurately represent human preferences? *Decis Sci* 1979; 10: 451-479.
7. Drummond MF, Stoddardt GL, Torrance GW. **Methods for the Economic Evaluation of Health Care Programmes**. Oxford: Oxford University Press, Inc.; 1987.
8. Fujiyama S, Izuno K, Yamasaki K, Sato T, Taketa K. Determination of optimum cutoff levels of plasma des-gamma-carboxy prothrombin and serum alpha-fetoprotein for the diagnosis of hepatocellular carcinoma using receiver operating characteristic curves. *Tumor Biol* 1992; 13: 316-323.
9. Guechot J, Laudat A, Loria A, Serfaty L, Poupon R, Giboudeau J. Diagnostic accuracy of hyaluronan and type III procollagen amino-terminal peptide serum assays as markers of liver fibrosis in chronic viral hepatitis C evaluated by ROC curve analysis. *Clin Chem* 1996; 42: 558-563.
10. Arnese M, Salustri A, Fioretti PM, Cornel JH, Boersma E, Reijs AEM, *et al*. Quantitative angiographic measurements of isolated left anterior descending coronary artery stenosis: Correlation with exercise echocardiography and technetium-99m 2-methoxy isobutyl isonitrile single-photon emission computed tomography. *J Am Coll Cardiol* 1995; 25: 1486-1491.
11. Lusted LB. Decision-making studies in patient management. *N Engl J Med* 1971; 284: 416-424.

12. Gohagen JK, Rodes ND, Ballinger WF, Blackwell CW, Butcher HR, Darby WP, *et al.* **Early Detection of Breast Cancer: Risk, Detection Protocols, and Therapeutic Implications.** New York: Praeger Publishers; 1982.
13. Gohagen JK, Spitznagel EL, McCrate M, Frank TB. ROC analysis of mammography and palpation for breast screening. **Invest Radiol** 1984; 19: 587–592.
14. DeBaun MR, Sox HC. Setting the optimal erythrocyte protoporphyrin screening decision threshold for lead poisoning: A decision analytic approach. **Pediatrics** 1991; 8: 121–131.
15. Milman N, Albeck MJ. Distinction between homozygous and heterozygous subjects with hereditary haemochromatosis using iron status markers and receiver operating characteristics (ROC) analysis. **Eur J Clin Chem Clin Biochem** 1995; 33: 95–98.
16. Bergus GR. When is a test positive? The use of decision analysis to optimize test interpretation. **Fam Med** 1993; 25: 656–660.
17. England WL. An exponential model used for optimal threshold selection on ROC curves. **Med Decis Making** 1988; 8: 120–131.
18. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. **Radiology** 1982; 143: 29–36.
19. Hdez-Armas J, Allende A, Alarco R, Panero J, Oton C. ROC analysis in gammagraphic diagnosis of hepatic cirrhosis. **Eur J Nucl Med** 1982; 7: 417–418.
20. Fombonne E. The use of questionnaires in child psychiatry research: Measuring their performance and choosing an optimal cut-off. **J Child Psychol Psychiatry** 1991; 32: 677–693.
21. Kay EJ, Knill-Jones R. Variation in restorative treatment decisions: Application of receiver operating characteristic curve (ROC) analysis. **Community Dent Oral Epidemiol** 1992; 20: 113–117.
22. Drummond M, Torrance G, Mason J. Cost-effectiveness league tables: More harm than good? **Soc Sci Med** 1993; 37: 33–40.
23. Eddy DM. Health system reform. Will controlling costs require rationing services? **JAMA** 1994; 272: 324–328.
24. Kassirer JP. Incorporating patients' preferences into medical decisions. **N Engl J Med** 1994; 330: 1895–1896.
25. Centor RM. Signal detectability: The use of ROC curves and their analyses. **Med Decis Making** 1991; 11: 102–106.
26. Buchsbaum DG, Buchanan RG, Centor RM, Scholl SH, Lawton MJ. Screening for alcohol abuse using CAGE scores and likelihood ratios. **Ann Intern Med** 1991; 115: 774–777.
27. Ramanujam N, Mitchell MF, Mahadevan A, Thomsen S, Richards-Kortum R. Fluorescence spectroscopy as a diagnostic tool for cervical intraepithelial neoplasia. **Gynecol Oncol** 1994; 52: 31–38.