Relative Operating Characteristic Analysis and Group Modeling for Tumor Markers: Comparison of CA 15.3, Carcinoembryonic Antigen, and Mucin-like Carcinoma-associated Antigen in Breast Carcinoma

Hulbert K. B. Silver, Betty-Lou Archibald, Joseph Ragaz, and Andrew J. Goldman

Department of Advanced Therapeutics [H. K. B. S., B-L. A.], Division of Medical Oncology [J. R.], and Division of Epidemiology, Biometry and Occupational Oncology [A. J. C.], British Columbia Cancer Agency, Vancouver, British Columbia, Canada

ABSTRACT

Relative operating characteristic (ROC) analysis was used to examine the clinical applicability of 3 breast carcinoma tumor markers, CA 15.3, carcinoembryonic antigen, and mucin-like carcinoma-associated antigen. Each tumor marker was quantitated in single serum samples collected from 100 normal blood donors, 60 patients with nonmalignant diseases, 33 women at high risk for breast carcinoma, 30 patients with malignancies other than breast carcinoma, and 158 breast carcinoma patients including 67 with no evidence of disease following surgery, 46 with a tumor burden <5 g, and 45 with a tumor burden >5 g. These were used to construct models for early diagnosis among those at high risk for breast carcinoma, the influence of nonmalignant disease on early diagnosis, discrimination of breast carcinoma from other adenocarcinomas, detection of early recurrence, and assessment of change in tumor burden. For each model ROC data permitted the unbiased selection of the most appropriate critical values based on the interaction of sensitivity and specificity. ROC analysis indicated that in practice the assays were remarkably similar. While CA 15.3 generally performed best, there was significant variation among models. Optimal marker selection can thus depend on specific clinical application. In some cases ROC identified a combination of markers as superior to any single assay, but this was not statistically significant.

INTRODUCTION

Tumor markers have become an important management tool in clinical oncology. While they have proven most useful in monitoring established disease, additional potential applications include diagnosis in the subgroup of patients suspected of harboring a malignancy, identification of a probable primary site for patients with an unknown primary, assessment of prognosis, and detection of early recurrence. Until recently the only well-accepted tumor marker for breast cancer was CEA. Its utility has been well reviewed. Now, with the advent of monoclonal antibody and recombinant DNA technology, there is an increasing number of potentially useful tumor markers. At least a dozen have been proposed for breast carcinoma in the past decade. Yet, it is often difficult or impossible to assess the relative clinical merit of a given marker from published reports. Part of the problem is that, while a number of statistical evaluation methods are available, results from one method may not be strictly comparable to another. In some cases the selected patient groups within a study do not reflect the intended clinical application or results for a tumor marker in one group of patients are used as a comparison for another marker used to evaluate different patients. Often unintended bias is introduced in the selection of decision criteria (normal cutoffs or upper limits of normal) required for most comparative analyses.

ROC analysis has evolved from signal detection theory as a general method of analyzing diagnostic systems (4–6). A major advantage of this analysis is that decision criteria need not be identified to compare accuracy of tests. Where the establishment of a decision criterion is indicated, ROC analysis can be used to identify a cutoff providing the best accuracy for a given test or aid in the unbiased selection of a cutoff for comparison among tests. In addition, the method provides for a relatively simple graphic representation of test accuracy.

The objective of this study was to evaluate ROC analysis using selected patient groups as models for three potential breast tumor markers: CEA, CA 15.3, and MCA.

MATERIALS AND METHODS

Serum Sample Collection

Blood was collected in glass tubes and centrifuged at 1400x g for 8 min; then the serum was separated and stored in 0.5-ml aliquots at −70° until assayed. A total of 382 patient samples were collected and divided into groups as described below.

N Group. The Canadian Red Cross kindly provided samples from 100 female donors, 16–91 years of age (mean, 38 years). At the time of sampling patients specifically denied any history of neoplastic, inflammatory, infectious, central nervous system, cardiovascular, or hepatic disease.

NMD Group. The underlying diagnoses for the 60 samples in this category included active rheumatoid arthritis (35 patients), colitis or diverticulitis (6), hypertension (4 patients), peptic ulcer (4 patients), chronic obstructive pulmonary disease (3 patients), pericarditis (3 patients), and one each of renal failure, cirrhosis, idiopathic anemia, lumbar disc disease, and sarcoidosis.

HR Group. This group of 33 women was defined as high risk by the presence of a suspicious breast lesion, as determined by physical examination and mammography, and the identification of cellular atypia after fine needle aspiration. The latter is a recognized risk factor for breast carcinoma.

NBA Group. The histologically confirmed diagnoses in this group of 30 adenocarcinoma patients included ovary (15 patients), lung (8 patients), and pancreas (7 patients). Each of these had advanced regional (15 patients) and/or advanced metastatic disease (15 patients).

Breast Carcinoma Group. Samples from these 158 patients with histologically confirmed breast carcinoma were divided into three groups. Group 1 included 67 premenopausal patients with a history of lymph node involvement. All known disease had been resected, and the patients had completed adjuvant chemotherapy and were on a follow-up protocol with no evidence of recurrence at the time of serum sampling. Group 2 included 46 patients sampled at the time of first recurrence of directly assessable local or regional disease estimated at <5 g. Evaluation to exclude distant metastatic disease included serum liver function studies (aspartate aminotransferase, lactic dehydrogenase, bilirubin), chest radiographs, and 99mTc-diphosphonate bone scintigraphy. Group 3 consisted of 45 patients with advanced disease...
characterized by local, regional, or distant metastases >5 g. This classification is in keeping with our own previous work (8-11) and that of others (12). Although the usual clinical staging is an indirect correlate of tumor burden (13), individuals with advanced disease may have very little tumor burden at some times during their clinical course (for example stage IV patients with no evidence of disease). For the purpose of tumor marker evaluation our system is more refined in that it is more directly quantifiable and clearly denotes tumor burden at the time of serum sampling. Blood samples from group 2 and 3 patients were obtained before institution of treatment.

Tumor Marker Assays

CEA, as initially described by Gold and Freedman (14), is a complex large molecular weight glycoprotein associated with the cellular glyco-calyx. Although best known as a tumor marker for colorectal cancer, the utility of CEA for breast carcinoma has been well documented (1). The method used in this study was an enzyme immunoassay method using heat extraction and polyclonal anti-CEA antibody as described by the manufacturer, Abbott Laboratories, Chicago, IL.

The CA 15.3 test is a double determinant radioimmunoassay utilizing two different monoclonal antibodies. The first, 115D8, was raised against milk fat globule membranes and reacts with a high molecular weight antigen, MAM-6. The second monoclonal antibody, DF3, was raised against a membrane-enriched extract of breast carcinoma cells and reacts with a heterogeneous M, 300,000-450,000 circulating antigen (15, 16). In this assay, 115D8 is immobilized on polystyrene beads to complete the double antibody sandwich with 125I-DF3. The detailed procedure was as described by the manufacturer (Centocor, Malvern, PA) and supplied by Amersham Canada Ltd. (Oakville, Ontario, Canada).

MCA is an M, 350,000 glycoprotein produced by mammary carcinomas and some normal tissues. The monoclonal antibody used in this assay, b-12, recognizes an epitope thought to be closely associated with breast carcinoma (17). The two-step solid phase enzyme immunoassay method was performed as directed by the manufacturer, Hoffman LaRoche Ltd., Etobicoke, Ontario, Canada.

Statistical Methods

The statistical basis for ROC methods have been well described (4-6). Briefly, when a test is used to detect patients having a disease (for example, breast carcinoma) among those free of the disease, a critical test value is usually selected that it is hoped will best distinguish between the two groups. For tumor markers, results greater than the critical value (test positive) generally denote increased probability of disease. This system defines four groups, those who: test positive with the disease (true positive), test positive without the disease (false positive), test negative with the disease (false negative), and test negative without the disease (true negative). ROC analysis takes advantage of a simplification of these familiar categories. The entire population can be described by just two functions: true-positive fraction (the proportion of test positives among those with the disease) and false-positive fraction (the proportion of test positives among those without the disease). These fractions are linked by any given critical value. For tumor markers such as those discussed in this paper, selection of a higher critical value must result in a smaller false-positive fraction as well as a smaller true-positive fraction. Clearly, selection of a critical value can have a profound influence when tumor markers are compared. The simple expedient of plotting true-positive fraction against false-positive fraction for a range of assay values will overcome many of the difficulties inherent in analyses based on critical value, as described in the text. The computer program used by us (June 1989 revision kindly provided by the developer, Charles E. Metz, Department of Radiology, University of Chicago Medical Center, Chicago, IL) analyzes data by a modification of a program by Dorfman and Alf (18). Related statistical tests are based on the bivariate normal model as previously described by Metz et al. (6).

Disease group comparisons were also made using conventional logistic regression analysis (19) in which all three tumor markers were used as variables. Not only did this permit an appraisal of the relative merits of ROC and logistic regression methods but the model developed by logistic regression could then be used in ROC analysis to examine the predictive value of using multiple markers as opposed to any single marker.

RESULTS

The ROC analysis of each tumor marker as a discriminator between normal females and group 3 (advanced disease) breast carcinoma patients is shown in Fig. 1. Each curve delineates the relationship between true- and false-positive fractions for a range of critical marker values. Since overall discrimination is a function of AUC, it is immediately apparent that in this case CA 15.3 is the best of the three markers. The AUC for CA 15.3 is significantly greater than that for MCA (P = 0.03) but not CEA. Critical values providing a 5% false-positive fraction, as interpolated from ROC data, are 25 μg/liter for CA 15.3, 2.5 μg/liter for CEA, and 13 μg/liter for MCA (Fig. 1). These values can, in turn, be used to construct scattergrams (Figs. 2 and 3) in which tumor marker values are rendered comparable by normalizing them in terms of the critical value. This permits a direct visual comparison of assays having different critical values. For example, Fig. 2 shows that a variety of people not harboring breast carcinoma may have marker values greater than perfectly normal controls.

Fig. 1. ROC analysis of 100 normal controls (group N) versus patients with advanced breast carcinoma (group 3, n = 45) for the three individual assays (CA 15.3, CEA, and MCA) and the combination defined by logistic regression analysis (ALL). FPF and TPF are expressed as percentages.

Fig. 2. Scattergram of assay values for CA 15.3, CEA, and MCA in patients who did not have breast carcinoma. Groups included NMD, HR, and NBA. Values are expressed on a log scale as multiples of the cutoff (MOC) defined by a 5% false-positive fraction among normal controls. Numbers below the double-ruled line, numbers of patients with assay values below the cutoff.
Fig. 3. Scattergram of assay values for CA 15.3, CEA, and MCA in breast carcinoma patients. Groups included those being followed up after resection of all known disease (group 1) and patients with limited tumor burden (group 2) and advanced disease (group 3). Values are expressed on a log scale as multiples of the cutoff (MOC) defined by a 3% false-positive fraction among normal controls. Numbers below the double-ruled line, numbers of patients with assay values below the cutoff.

Fig. 4. ROC analysis of high-risk patients (group HR, n = 33) versus those with limited disease (group 2, n = 46) for the three individual assays (CA 15.3, CEA, and MCA) and the combination defined by logistic regression analysis (ALL). FPF and TPF are expressed as percentages.

Fig. 5. ROC analysis of patients with NMD (n = 60) versus breast carcinoma patients with limited disease (group 2, n = 46) for the three individual assays (CA 15.3, CEA, and MCA) and the combination defined by logistic regression analysis (ALL). FPF and TPF are expressed as percentages.

Fig. 6. ROC analysis of patients with advanced carcinoma other than breast (group NBA, n = 30) versus advanced breast carcinoma patients (group 3, n = 45) for the three individual assays (CA 15.3, CEA, and MCA) and the combination defined by logistic regression analysis (ALL). FPF and TPF are expressed as percentages.

Tumor markers are rarely used as in Fig. 1 to discriminate between perfectly normal individuals and cancer patients. More appropriate comparison groups would serve as better models for decision making. One potential application for breast tumor markers is to identify carcinoma patients among women suspected of having underlying breast carcinoma. The model we have chosen to examine is one that tests how well each marker discriminates between the HR and limited disease (group 2) groups, as defined above. ROC analysis is displayed in Fig. 4. The best single discriminator, as determined by AUC, is CA 15.3, and this is significantly better than MCA (P = 0.01). The dip in the CEA curve suggests unreliability for very low CEA values. Logistic regression analysis identified CA 15.3 as having the most independent predictive power (P = 0.014), with MCA also a significant independent predictor (P = 0.037). ROC analysis of the three assays in combination, as defined by logistic regression analysis, showed that the combined function was better than any single marker. However, this approach had statistical significance only for the comparison with CEA (P = 0.092).

For any diagnostic application it is important to evaluate how the identification of small tumor burden might be confounded by other diseases, particularly inflammatory conditions. A possible model for ROC analysis would be one that compares the NMD group with group 2 (Fig. 5). In keeping with Fig. 2, in which marker values tended to be higher in the NMD group than the HR group, the areas under the curves in Fig. 5 tend to be less than in Fig. 4.

Another diagnostic problem is the presentation of a patient with advanced metastatic adenocarcinoma but no obvious primary site. An ROC model for this is one that determines how well a given marker can distinguish between group 3 and patients with advanced adenocarcinomas from sites other than breast (NBA). Fig. 6 shows that in this case CA 15.3 and MCA are better than CEA (for AUC, P = 0.005 and P = 0.01, respectively). A combination of assays was not significantly better than either CA 15.3 or MCA alone.

Our model for the early detection of recurrence compared group 1 (those who had no evidence of disease after surgery) with group 2. As shown in Fig. 7, all three markers were remarkably similar. No single marker AUC was significantly better than another, nor was the AUC significantly better for the combined function defined by logistic regression analysis. The shape of the curves in Fig. 7 is markedly different from those in Fig. 1, in which an FPF of 3% corresponded to a relatively high TPF. However, the curves in this region of Fig. 7 are relatively steep and a modest increase in FPF results in a disproportionate increase in TPF. A critical value dictated by a 10% FPF would yield a TPF of 42% for the best tumor marker.
clinical decision making (20, 21), yet it has not been widely
and TPF are expressed as percentages.


disease (group 2, n = 46) for the three individual assays (CA 15.3, CEA, and
reseetion of all known disease (group I, n = 67) versus patients with limited

DISCUSSION

0.15, 0.33, and 0.88 for CEA, CA 15.3, and MCA, respectively.

was one that used ROC to evaluate group 2 versus group 3 (Fig. 7). The AUC for CEA was less than either CA 15.3 or MCA,

The use of ROC analysis has been suggested as an aid to
cancer and other adenocarcinomas. Our findings are generally

disease (Figs. 2 and 5). However, tumor markers do have
diagnostic potential in the subgroup of people in whom malig
selection of critical values.

having determined critical values, it is possible to construct
scattergrams normalized on this basis (Figs. 2 and 3). This
permits an immediate assessment of the relative merit of the
various assays. In this case the most striking feature is that all
three assays may be confounded by conditions other than breast
carcinoma.

the configuration of the curves in Fig. 1 indicate that all 3
assays have an excellent ability to discriminate between per-
fectly normal individuals and patients with advanced disease.

Unfortunately, these groups do not reflect clinical reality in
which the usual diagnostic problem is to identify relatively
small tumor burden. We know of no data supporting the use of
these assays alone for diagnostic screening. In that application
one would expect the false-positive rate to be excessive, espe-

nancy is already suspected. In the ROC model for this (Fig. 4)
it is immediately apparent that assay discrimination has de-
creased compared with the clinically improbable conditions
represented by Fig. 1. While the markers do have predictive
power in the early diagnosis model, especially for the combined
function, it is unlikely that these tumor markers would replace
mammography. It is possible that a combination of mammog-
ography and a tumor marker would be significantly better than
either alone. This has been examined in a preliminary study of
another tumor marker, MSA (27).

As indicated in Figs. 2 and 6 none of the tumor markers are
truly diagnostic when discriminating between advanced breast
cancer and other adenocarcinomas. Our findings are generally
in keeping with others who found individual markers elevated
in patients harboring a variety of advanced epithelial malignan-
cies (2, 28). Given the above, the discriminating power displayed
in Fig. 6 is surprisingly good for CA 15.3 and MCA. This
suggests that these markers could prove helpful in the clinical
“primary unknown” situation in which one may have to depend
on the best information available to select treatment for trial.
By contrast, in a study using a fixed criterion based on normal
sera, Colomer et al. (28) concluded that CA 15.3 would probably
not be useful in this application. It is possible that ROC
analysis, which is not constrained by a fixed critical value,
would have revealed some discriminating power. This question

Table 1 Critical values for each assay related to comparison groups and FPF

<table>
<thead>
<tr>
<th>Groups compared*</th>
<th>Critical values (µg/liter)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FPF 3%</td>
</tr>
<tr>
<td>CA15.3</td>
<td>CEA</td>
</tr>
<tr>
<td>N vs. 3</td>
<td>25</td>
</tr>
<tr>
<td>I vs. 2</td>
<td>46</td>
</tr>
</tbody>
</table>

* Groups are defined in “Materials and Methods.”

and 48% for the combined function based on logistic regression
analysis. The corresponding critical values, as interpolated from
the ROC data, would be 27 µg/liter for CA 15.3, 2.8 µg/liter
for CEA, and 11 µg/liter for MCA (Table 1).

Tumor markers can be most effectively used to monitor the
progress of established disease. The model chosen in this case
was one that used ROC to evaluate group 2 versus group 3 (Fig.
8). The AUC for CEA was less than either CA 15.3 or MCA,
but this was not statistically significant. By contrast, logistic
regression analysis identified CEA as contributing the most
independent predictive information with significance values of
0.15, 0.33, and 0.88 for CEA, CA 15.3, and MCA, respectively.

DIscussion

The use of ROC analysis has been suggested as an aid to
clinical decision making (20, 21), yet it has not been widely
accepted or fully exploited, especially in the field of tumor
markers. Some authors (22) have included ROC analysis as an
element in their evaluation but have not included crucial ele-
ments in the analysis such as the use of clinically relevant group
models or the importance of ROC for selecting critical values
defined by a specific clinical problem.

The ROC analysis of group N versus group 3 (Fig. 1) has
some practical use as the basis for normalizing simple scatter-
grams. However, the primary reason for including this analysis
is to illustrate the inappropriateness of validating tumor marker
assays with these groups, as has been frequently the case in the
past. Certainly, critical values are most commonly derived from
data of perfectly normal individuals and a cutoff corre-
ponding to the 3–5% FPF. The critical values we obtained
are in keeping with others for CA 15.3 and MCA (17, 23–26).

For CEA our critical value of 2.5 is one half the value often
selected by others from general experience only (1). Inadvertent
selection of an inappropriately high critical value for CEA
would make CEA appear less sensitive in comparative studies.
One advantage of ROC analysis is it lends itself to the unbiased
selection of critical values.

Having determined critical values, it is possible to construct
scattergrams normalized on this basis (Figs. 2 and 3). This
permits an immediate assessment of the relative merit of the
various assays. In this case the most striking feature is that all
three assays may be confounded by conditions other than breast
carcinoma.

The configuration of the curves in Fig. 1 indicate that all 3
assays have an excellent ability to discriminate between per-
fectly normal individuals and patients with advanced disease.

Unfortunately, these groups do not reflect clinical reality in
which the usual diagnostic problem is to identify relatively
small tumor burden. We know of no data supporting the use of
these assays alone for diagnostic screening. In that application
one would expect the false-positive rate to be excessive, espe-

nancy is already suspected. In the ROC model for this (Fig. 4)
it is immediately apparent that assay discrimination has de-
creased compared with the clinically improbable conditions
represented by Fig. 1. While the markers do have predictive
power in the early diagnosis model, especially for the combined
function, it is unlikely that these tumor markers would replace
mammography. It is possible that a combination of mammog-
ography and a tumor marker would be significantly better than
either alone. This has been examined in a preliminary study of
another tumor marker, MSA (27).

As indicated in Figs. 2 and 6 none of the tumor markers are
truly diagnostic when discriminating between advanced breast
cancer and other adenocarcinomas. Our findings are generally
in keeping with others who found individual markers elevated
in patients harboring a variety of advanced epithelial malignan-
cies (2, 28). Given the above, the discriminating power displayed
in Fig. 6 is surprisingly good for CA 15.3 and MCA. This
suggests that these markers could prove helpful in the clinical
“primary unknown” situation in which one may have to depend
on the best information available to select treatment for trial.
By contrast, in a study using a fixed criterion based on normal
sera, Colomer et al. (28) concluded that CA 15.3 would probably
not be useful in this application. It is possible that ROC
analysis, which is not constrained by a fixed critical value,
would have revealed some discriminating power. This question

Fig. 8. ROC analysis of breast carcinoma patients with limited disease (group
2, n = 46) versus patients with advanced disease (group 3, n = 45) for the three
individual assays (CA 15.3, CEA, and MCA) and the combination defined by
logistic regression analysis (ALL). FPF and TPF are expressed as percentages.

Fig. 7. ROC analysis of breast carcinoma patients being followed up after
resection of all known disease (group 1, n = 67) versus patients with limited
disease (group 2, n = 46) for the three individual assays (CA 15.3, CEA, and
MCA) and the combination defined by logistic regression analysis (ALL). FPF
and TPF are expressed as percentages.
would best be settled in a larger study than ours, including a variety of epithelial malignancies individually matched for tumor burden. It is perhaps not surprising that our study showed significantly less specificity for CEA, since that antigen was derived from adenocarcinoma of bowel, while the other markers were developed from breast carcinoma.

The ROC model for detection of early recurrence is more encouraging (Fig. 7). One advantage of ROC analysis is that it permits a ready evaluation of the dynamic interaction of the TPF and FPF over a range of potential critical values. In this case one could take advantage of the relatively steep curves to propose that critical values based on an FPF of 10% would be much more relevant than the more standard 3% used in the discussion of Fig. 1.

Not only is ROC helpful in selecting a critical value, it demonstrates that a critical value selected for one clinical problem may be quite misleading in another application. This is emphasized in Table 1 in which critical values defined by a 3% FPF for group N versus group 3 (25, 2.5, and 13 μg/liter for CA 15.3, CEA, and MCA, respectively) are quite different for a 3% FPF applied to the group 1 versus group 2 early recurrence model (46, 6.5, and 16 μg/liter for CA 15.3, CEA, and MCA, respectively). In fact, the critical values dictated for a 3% FPF in group N versus group 3 are much closer to the 10% FPF in the early recurrence model. Investigators using decision criteria based on the 3–5% FPF for a normal population may be unwittingly dealing with quite different FPF values when the same decision criteria are applied to a different clinical setting. Further examination of Table 1 shows that the shift in critical value of one tumor marker cannot be used to predict the behavior of another. The use of ROC analysis clarifies this problem and encourages the selection of decision criteria tailored to a specific clinical problem.

Identification of early recurrence is a potential clinical application for tumor markers that deserves close scrutiny. Studies of CEA have been inconsistent, perhaps partly because optimum critical values had not been selected for this purpose (1). Our single sample ROC model for early recurrence (Fig. 7) suggests that tumor markers might be useful. Specific recommendations would require a prospective serial sample study aided by ROC and including an analysis of diagnostic lead time. We have such a study in progress. Even so, application of tumor markers for this clinical problem would only be warranted if further intervention were contemplated for asymptomatic early recurrence.

The greatest clinical application for tumor markers is in the serial monitoring of patients with established disease. An increasing or decreasing CEA correlates with disease progression or regression, respectively, in about 85% of cases (1). From Fig. 8 we might anticipate that MCA and CA 15.3 would perform at least as well. Our finding that CEA as a single marker did less well than CA 15.3 is in keeping with others (29), although in our study this was not statistically significant.

ROC and logistic regression analyses are powerful tools for the assessment of tumor markers. When one uses the AUC method for comparing curves, ROC is similar to logistic regression analysis in providing a global assessment of assay performance independent of any fixed critical value. Logistic regression analysis is modeled on the optimal mathematical discrimination, which is, in turn, influenced by the distribution of the 2 samples. This may not accurately reflect decisions based on fixed decision criteria. A similar potential shortcoming is perhaps more evident for ROC in which it is graphically apparent that all areas of the curve are not equally important. However, with ROC the clear visual correlate of performance can be used to calculate either the best discriminating critical value or select a critical value based on the most clinically acceptable compromise between an FPF and TPF. For example, in Fig. 7 the optimal discriminating value for CA 15.3, as determined by the point of the curve closest to the upper left hand corner (100% TPF, 0% FPF) (30), would be at an FPF of 28%. Yet the decision maker might decide that a 10% FPF value would be more clinically useful. Having selected the best decision criterion for a given clinical problem, one can then compare the discriminating power at this point rather than relying on AUC comparisons (31). In our study the resulting values were not remarkably different from AUC comparisons and have not been reported.

While ROC analysis can identify a marker with significantly improved discriminating power in a pairwise analysis, the great power of logistic regression analysis is in identification of significant independent predictors that may be useful in combination. For example, in the early diagnosis model (Fig. 4) ROC analysis identified CA 15.3 as significantly better than MCA as a single predictor. Yet logistic regression analysis identified MCA as a significant independent predictor.

In an effort to combine the merits of both methods of analysis we have used the logistic regression function derived from all three assays to define an additional ROC curve. In each case the combined function was best but not significantly better than the best single marker.

In our analysis CA 15.3 has been consistently the best overall discriminator. However, it is perhaps more interesting that when placed on an equal footing by ROC analysis, the assays are remarkably similar. A larger study would be required to reveal more subtle differences. Our studies illustrate the need for clinical modeling when comparing assays and the importance of selecting critical values based on such models. ROC analysis with these relatively simple models would be most valuable in assessing the spate of competing tumor markers now becoming widely available. In fact, the general methods could be applied to a great variety of clinical and research decision-making problems.

ACKNOWLEDGMENTS

The authors wish to thank Dr. J. Hanley of the Department of Epidemiology and Biostatistics, McGill University, and Dr. C. Metz of the University of Chicago Medical Center for their generosity in providing current ROC analysis programs and discussion of appropriate application. We also thank the technical staff in the Tumor Marker Laboratory, Kim Kieler for her help in running computer programs, Dr. P. Rebeck for the provision of clinical material, and Linda Wood for secretarial assistance.

REFERENCES


Relative Operating Characteristic Analysis and Group Modeling for Tumor Markers: Comparison of CA 15.3, Carcinoembryonic Antigen, and Mucin-like Carcinoma-associated Antigen in Breast Carcinoma


Updated version
Access the most recent version of this article at:
http://cancerres.aacrjournals.org/content/51/7/1904

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.