CANCER
CYTOPATHOLOGY

Computer-Derived Nuclear Features Compared with Axillary Lymph Node Status for Breast Carcinoma Prognosis

William H. Wolberg, M.D.[1]
W. Nick Street, Ph.D.[2]
Olvi L. Mangasarian, Ph.D.[3]

[1] Department of Surgery, University of Wisconsin, Madison, Wisconsin.

[2] Computer Science Department, Oklahoma State University, Stillwater, Oklahoma.

[3] Computer Sciences Department, University of Wisconsin, Madison, Wisconsin.

Presented as a poster and computer demonstration at the American College of Surgeons meeting, San Francisco, California, October 7–10, 1996.

Supported in part by National Institutes of Health INRSA Fellowship 1 F32 CA 68690-01, Air Force Office of Scientific Research Grant AFOSR F49620-94-1-0036, and National Science Foundation Grant CCR-9322479.

Address for reprints: William H. Wolberg, M.D., Department of Surgery, University of Wisconsin, 600 Highland Avenue, H4/750 CSC, Madison, WI 53792.

Received January 30, 1997; revision received March 18, 1997; accepted March 20, 1997.

BACKGROUND. Both axillary lymph node involvement and tumor anaplasia, as expressed by visually assessed grade, have been shown to be prognostically important in breast carcinoma outcome. In this study, axillary lymph node involvement was used as the standard against which prognostic estimations based on computer-derived nuclear features were gauged.

METHODS. The prognostic significance of nuclear morphometric features determined by computer-based image analysis were analyzed in 198 consecutive preoperative samples obtained by fine-needle aspiration (FNA) from patients with invasive breast carcinoma. A novel multivariate prediction method was used to model the time of distant recurrence as a function of the nuclear features. Prognostic predictions based on the nuclear feature data were cross-validated to avoid overly optimistic conclusions. The estimated accuracy of these prognostic determinations was compared with determinations based on the extent of axillary lymph node involvement.

RESULTS. The predicted outcomes based on nuclear features were divided into three groups representing best, intermediate, and worst prognosis, and compared with the traditional TNM lymph node stratification. Nuclear feature stratification better separated the prognostically best from the intermediate group whereas lymph node stratification better separated the prognostically intermediate from the worst group. Prognostic accuracy was not increased by adding lymph node status or tumor size to the nuclear features.

CONCLUSIONS. Computer analysis of a preoperative FNA more accurately identified prognostically favorable patients than did pathologic examination of axillary lymph nodes and may obviate the need for routine axillary lymph node dissection. ***Cancer (Cancer Cytopathol)* 1997;81:172–9.** © *1997 American Cancer Society.*

KEYWORDS: breast carcinoma, prognosis, nuclei, lymph nodes, computer analysis, machine learning.

Axillary lymph node involvement is generally considered to be the strongest prognostic factor for breast carcinoma outcome. The metastatic status of the axillary lymph nodes is sufficiently important prognostically that these lymph nodes are routinely removed as part of the initial treatment of most breast carcinomas. Many medical and personal decisions are based on the number of these lymph nodes that are found to contain carcinoma. For example, adjunctive chemotherapy is usually given to lymph node positive and not lymph node negative patients. However, axillary lymph node removal is not devoid of morbidity; the ipsilateral arm remains susceptible to infection and lymphedema.

Prognostic data analysis presents unique estimation problems

TABLE 1
Case Distribution by Involved Axillary Lymph Nodes in the Current Series

Positive lymph nodes	No. of patients[a]	Percentage
0	87	44.8%
1–3	56	28.9%
≥4	51	26.3%

[a] No axillary lymph node dissection was performed in four patients who had small tumors, favorable histology, and no palpable adenopathy.

because not all cases are followed to the endpoint being analyzed (i.e., censored), in this case time to distant recurrence (TTR) or to death. In the authors' work, nuclear features are generated by digital image analysis of breast fine-needle aspirates (FNA). Previously, classic statistical analysis of the authors' data showed that these computer-derived nuclear features were prognostically more significant than tumor size and lymph node status and including tumor size and lymph node status with the nuclear features did not increase the accuracy of prognostic estimations.[1,2]

In this article, the authors used a new prognostic estimation method to compare the prognostic accuracy of lymph node status with that obtained by computer analysis of breast FNA cytology. They approached the prediction TTR as a function estimation problem, using techniques of mathematic optimization and machine learning to produce a multivariate predictive model.

PATIENTS AND METHODS

Patients and Aspirates

The cell samples used in this study were FNAs acquired since 1984 by one of the authors (W.H.W.) to diagnose a consecutive series of 198 patients with invasive breast carcinoma. Aspirates were classified as carcinoma based on histologic confirmation by surgical biopsy. An additional 15 in situ carcinomas and 22 invasive primary carcinomas with distant metastasis diagnosed by FNA were eliminated from this analysis. The case distribution by cancerous axillary lymph nodes is seen in Table 1.

To prepare an FNA, a small drop of viscous fluid was aspirated from breast masses by making multiple passes with a 23-gauge needle while negative pressure was applied to an attached syringe. The aspirated material was expressed onto a silane-prepared glass slide and the aspirate was spread when a similar slide was applied face-to-face and the slides were separated with a horizontal motion. Preparations were immediately fixed in 95% ethanol, stained with hematoxylin and eosin, and processed.

Image Preparation

The area on the slide to be imaged was visually selected. The image selected for digital analysis was generated by a JVC TK-1070U color video camera mounted atop an Olympus microscope and the image projected into the camera with ×2.5 ocular and a ×63 objective lens. The image was captured by a ComputerEyes/RT color framegrabber board (Digital Vision, Inc., Dedham MA) as a 640 × 400, 8 bit/pixel Targa file. An 8 bit/pixel gray scale image was used for the image analysis because the authors did not analyze the color of the nuclei. The conversion to grayscale was performed with software using a standard conversion algorithm.

The number of nuclei available for analysis by Xcyt is limited by the number of cells captured in the microscope field because Xcyt stores the data obtained from a single image. One field per case was selected for image analysis. Nuclei that were markedly distorted during preparation and those that were significantly overlapping were not selected for analysis. Thus, the current data were obtained from analysis of approximately 10–20 nuclei.

The User Interface (Xcyt)

The first step in analyzing the digital image is to specify the exact location of each cell nucleus. A graphic user interface allows the user to input the approximate location of nuclei. The interface was developed using the X Windows System and the Athena WidgetSet. A mouse button was used to trace a rough outline of each visible cell nucleus. Beginning with this user-defined approximate border, the actual boundary of the cell nucleus was located by an active contour model known as a ''snake,''[3,4] a deformable spline that seeks to minimize an energy function defined over the arc length of a curve. The energy function is defined in such a way that the snake, in the form of a closed curve, conforms itself to the boundary of a cell nucleus. The snake program creates a series of equally spaced points around the nuclear contour. The mathematic aspects of the snake calculations are described elsewhere.[5] The use of the snake program to directly determine nuclear boundaries using the computer distinguishes Xcyt from the methods used in other studies. A graphic illustration of the Xcyt interface is presented in Figure 1.

Once the nuclei to be analyzed have been outlined, the program calculates ten real-valued nuclear features for each nucleus.[5] Features were verified using idealized phantom cells.[6] The computed features are as follows:

1. **Radius** is computed by averaging the length of radial line segments.

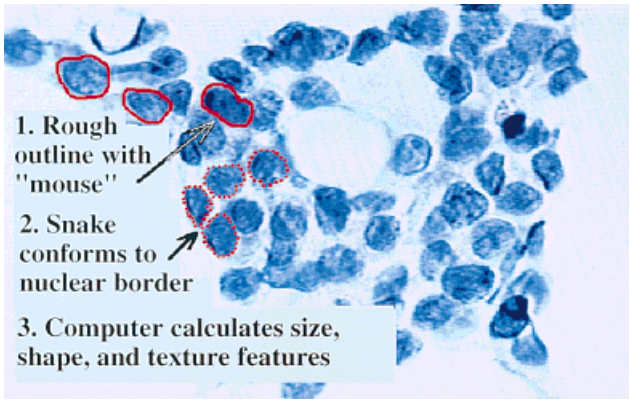

FIGURE 1. A diagrammatic illustration of the user-computer interface used for nuclear feature analysis.

2. **Perimeter** is measured as the sum of the distances between consecutive snake points.
3. **Nuclear area** is measured by counting the number of pixels on the interior of the snake and adding one-half of the pixels in the perimeter, to correct for digitization error.
4. Perimeter and area are combined to give a measure of the **compactness** of the cell nuclei using the formula perimeter^2/area. This dimensionless number is minimized for a circle and increases with the irregularity of the boundary.
5. The **smoothness** of a nuclear contour is quantified by measuring the difference between the length of a radial line and the mean length of the two radial lines surrounding it. If this number is relatively small, the contour is smooth in that region.
6. **Concavity** is captured by measuring the size of any indentations in the cell nucleus. Chords are drawn between nonadjacent snake points, and the extent to which the actual boundary of the nucleus lies on the inside of each chord is measured and averaged. A circular or elliptical nucleus would show no concavity.
7. **Concave points:** this feature is similar to concavity but counts only the number of snake points that lie on concave regions of the contour rather than the magnitude of such concavities.
8. **Symmetry** is measured by finding the relative difference in length between pairs of line segments perpendicular to the major axis of the cell nucleus contour. The major axis is determined by finding the longest chord that passes from a snake point through the center of the nucleus. The segment pairs are then drawn at regular intervals.
9. **Fractal dimension:** The fractal dimension of a cell is approximated using the ''coastline approximation'' described by Mandelbrot.[7] The perimeter of the nucleus is measured using increasingly larger ''rulers.'' As the ruler size increases, decreasing the precision of the measurement, the observed perimeter decreases. Plotting these values on a log scale and measuring the downward slope gives the negative of an approximation to the fractal dimension.
10. **Texture:** The texture of the cell nucleus is measured by finding the variance of the gray scale intensities in the component pixels.

The mean value, largest extreme (worst) value, and standard error of each feature are computed for each image. To reduce possible noise, the three largest values are averaged in computing the extreme values. The extreme value features are the most intuitively useful for the problem at hand, because only a few malignant cells may occur in a given sample. The authors have found that the mean value features appear to be stable after the inclusion of 10–20 nuclei, suggesting that this is an appropriate sample.

TABLE 2
Case Distribution by Involved Axillary Lymph Nodes in the SEER Series

Positive lymph nodes	No. of patients	Percentage
0 (best)	19,072	54.3%
1–3 (intermediate)	7500	21.3%
≥4 (worst)	8564	24.4%

SEER: Surveillance, Epidemiology, and End Results.

Surveillance, Epidemiology, and End Results Database

The Surveillance, Epidemiology, and End Results (SEER) Program of the National Cancer Institute has collected and maintained data on the cancer survival experience of more than 24,000 women newly diagnosed with breast carcinoma between 1977 and 1982. From the SEER database, the authors selected women with invasive breast carcinoma who had no evidence of distant metastases at the time of diagnosis. For comparison with Table 1, the SEER case distribution by cancerous axillary lymph nodes is seen in Table 2.

Recurrence Surface Approximation

Analysis of survival data has long been addressed in the statistics literature[8] with methods that estimate hazard or survival curves, such as Cox regression.[9] Instead, the authors formulated prognosis as a function approximation problem, predicting TTR using a parametric model of the input features. Details of the recurrence surface approximation (RSA) method are

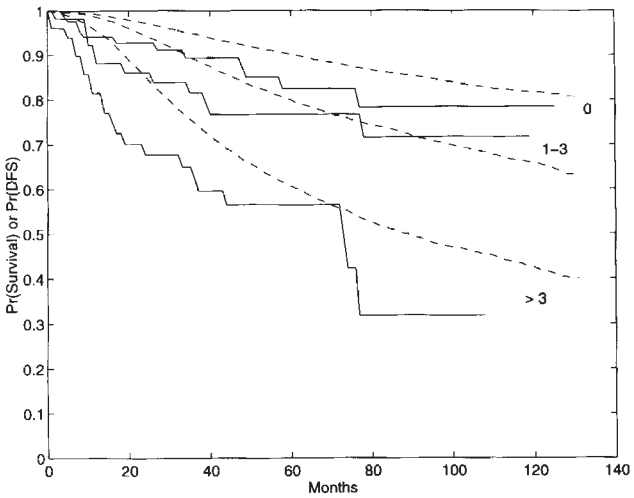

FIGURE 2. Kaplan–Meier plot of disease free survival stratified by lymph node involvement for the patients in the current series (solid line) compared with the survival of Surveillance, Epidemiology, and End Results patients (dashed line).

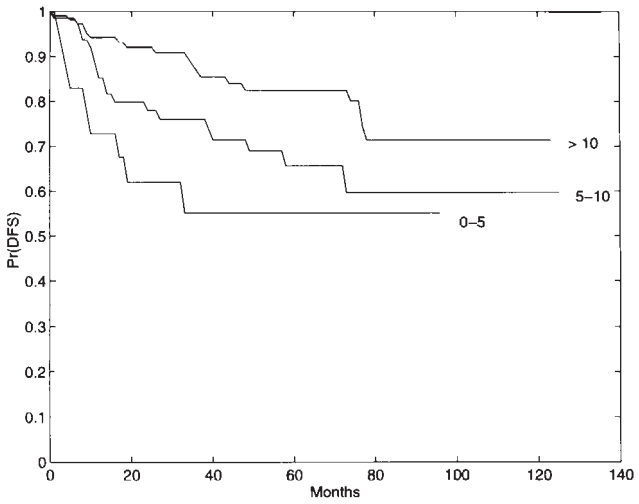

FIGURE 3. Kaplan–Meier plot of disease free survival for the patients in the current study stratified by nuclear features. The nuclear feature analyses are expressed as predicted time to recurrence. The patients are grouped in this figure by predicted time to recurrence between 0 and 5 years (worst), between 5 and 10 years (intermediate), and >10 years (best). Compare this figure with the solid line in Figure 2. Pr: predicted; DFS: disease free survival.

given elsewhere[10] and are summarized in the Appendix. The model attempts to fit (in a least absolute error sense) the observed recurrences as closely as possible, while overestimating the censoring times in some reasonable way. The resulting predictive model gives an expected TTR for each case, in contrast to Cox regression, which estimates a recurrence function for each case. Expected disease free recurrence functions are obtained from recurrence predictions by grouping together a set of examples with similar predicted TTRs and using their actual outcomes to produce a Kaplan–Meier estimate.[11]

TABLE 3
Distribution in the Current Series by Estimated Time to Recurrence Based on Computer-Derived Features

Estimated TTR	No. of patients	Percentage
>10 years (best)	109	55.1%
5–10 years (intermediate)	65	32.8%
<5 years (worst)	24	12.1%

TTR: time to recurrence.

Data Analysis

The median follow-up was 55 months for patients in whom recurrence has not been observed (censored cases), and the median TTR (noncensored cases) was 16 months.

To compare the results of the authors prediction method with the lymph status stratification, a method that estimates the accuracy of the new model on unseen cases was needed. Any sufficiently flexible parametric (or nonparametric) model might appear to explain a given set of training cases; however, the prediction rule may generalize poorly to cases on which it was not trained, particularly if it is overly complex. To obtain a fair, unbiased estimate of future accuracy, a leave-one-out cross-validation[12] was performed in the following automatic fashion. One case was first set aside. The RSA mathematic model was then solved with the remaining 197 cases and the resulting surface tested on the case that was originally set aside. Each case was individually removed and used as a ''test case.'' Once all cases were tested in this manner, they were rank ordered, based on the test case result, into three groups, as in the TNM lymph node stratification. Patients were grouped by predicted time to recurrence between 0–5 years (worst), between 5–10 years (intermediate), and >10 years (best).

Data are presented as Kaplan–Meier curves[11] and group differences are compared by Gehan's generalized Wilcoxon test.[13]

RESULTS

The patients in the current study and the SEER patients had similar extent of lymph node invasion. Furthermore, the current study patients' outcomes (Fig.

TABLE 4
Nuclear Features Stratification in the Current Series

Categories	Wilcoxon score	Z	*P* value
>10 years vs. 5–10 years (best vs. intermediate)	832	2.226	0.0260
>10 years vs. <5 years (best vs. worst)	660	3.364	0.00074
5–10 years vs. <5 years (intermediate vs. worst)	258	1.689	0.0913

TABLE 5
Lymph Node Stratification in the Current Series

Categories	Wilcoxon score	Z	*P* value
0 vs. 1–3 (best vs. intermediate)	260	1.029	0.3034
0 vs. ≥4 (best vs. worst)	1017	3.631	0.00028
1–3 vs. ≥4 (intermediate vs. worst)	516	2.322	0.0202

2, solid line) were similar to those of the SEER patients (Fig. 2, dashed line) when compared according to the extent of lymph node involvement, and when allowance was made for different endpoints; the current study's was distant recurrence whereas the SEER endpoint was death (TTR was not available for SEER patients).

Case distribution in the current study by estimated TTR based on computer-derived features is given in Table 3.

The current study data could be stratified either by lymph node status (solid line in Fig. 2) or by computer-derived nuclear features (Fig. 3). The group differences are compared by Gehan's generalized Wilcoxon test[13] in Tables 4 and 5. Nuclear feature stratification better separates the prognostically best from the intermediate group, whereas lymph node stratification better separates the prognostically intermediate from the worst group.

Prognostic accuracy did not improve when the number of cancerous lymph nodes or tumor size were added to the nuclear features during the RSA prediction process.

The results did not change significantly when patients in the current study were ranked and grouped 50:25:25 (best:intermediate:worst) to obtain the same percentage distribution as resulted from stratification by number of involved lymph nodes.

DISCUSSION

The accuracy of prognostic estimations based on nuclear features was compared with those based on axillary lymph node involvement because the extent of axillary lymph node involvement is generally considered to be the strongest prognosticator of breast carcinoma behavior. Axillary lymph node involvement is prognostically stratified at 0, 1–3, and ≥4 positive lymph nodes.[14,15] In addition, immediate axillary lymph node removal (i.e., removal performed at the time of initial breast carcinoma surgery) as opposed to delaying removal (i.e., deferred removal until such time as the lymph nodes become palpably enlarged) was shown not to affect survival.[16,17] This finding led Fisher et al. to posit that positive axillary lymph nodes were not the predecessor of distant tumor spread but rather were a manifestation of dissemination.[16] Therefore, axillary lymph nodes are removed as part of the initial breast carcinoma surgery for prognostic staging rather than cure. After complete axillary lymph node dissection, the ipsilateral arm is at increased risk for infection and there is a 20% incidence of arm lymphedema.[18,19]

Both axillary lymph node involvement and tumor anaplasia, as expressed by grade, have been shown to be of prognostic importance. The prognostic value of nuclear grade of breast carcinoma has been well established since Black et al.[20] first described the relationship between prognosis and nuclear anaplasia in 1955. Calibrated oculars, projection microscopy, and graphic tablets have been used for many years to measure nuclear features. Objective cell image analysis using computers has become increasingly sophisticated during the past 30 years.[21] The results of computer-based analyses are reproducible and correlate closely with visual assessments.[22] Such techniques demonstrate that larger nuclear size is associated with a poor prognosis.[23–26] Two studies[23,24] also found variation in nuclear size, as reflected in the standard deviation of nuclear size features, to be prognostically unfavorable. The results of the current study may be compared with those of Aubele et al.,[27] who used image analysis and found nuclear morphometric and textural parameters to be prognostically significant but not as strongly as axillary lymph status. Their study was performed with Feulgen-stained samples and their parameters emphasized nuclear DNA characteristics whereas the parameters in the current study emphasized morphology and minimized nuclear textural features. In addition, Aubele et al.[27] worked only with pT1 classified carcinomas and used 5-year distant recurrence free survival as an endpoint, whereas in the current study the authors worked with carcinomas of all sizes and analyzed distant recurrence free survival for >5 years. Currently, a large trial is underway evaluating morphometry (nuclear area and axes ratio) relative to other prognostic factors.[28]

In contrast to previously cited studies, the results presented here were cross-validated to protect against presenting overly optimistic results that resulted from arbitrary cutpoints and the use of unprocessed retrospective data. These cross-validated results may be expected to be robust when subjected to prospective studies, because in all cases the example used for testing was not used in creating the predictive model. The authors have found cross-validation to be a powerful tool for projecting future results. For example, in 1994 they predicted that the Xcyt diagnostic system was capable of classifying benign and malignant cytology with 97.5% accuracy. Since then, Xcyt has correctly classified 97.9% of 192 new cases.[29]

In contrast to the methods used in other studies, segmentation (i.e., determination of nuclear boundaries) is determined automatically by the computer ''snake'' program. Furthermore, the authors' studies use the cellular smear type preparations in which nuclear detail is better preserved than in the histologic preparations used in many of the previous studies. Cells obtained by FNA are preserved intact whereas histologically processed cells are cut in various planes. Despite these technical differences, the authors' prognostic accuracy is almost identical to that reported by Komitowski and Janson.[30] They used projection microscopy and a digitizing tablet to determine size, shape, and texture features in 60 breast carcinoma patients. They achieved 85% prognostic accuracy; inclusion of tumor size increased accuracy to 92%. Pienta and Coffey[31] found that nuclear pleomorphism as measured by both nuclear area and intrasample variation increased with invasive histology and with axillary lymph node involvement with metastatic breast carcinoma.

In the current study, nuclei for analysis were selected by an operator from an area deemed to be the most atypical. Such selection is subject to operator bias compared with a random selection process. In a series of breast carcinomas, Baak et. al. studied the results of nuclear size measurements made selectively by an operator who chose maximally atypical areas with those made randomly.[32] They found the measurements obtained by both selection processes to be closely correlated and concluded that, even when the most atypical area was selected for measurement, the results were comparable to those obtained by systematically randomly measuring over the entire slide. Moreover, in the authors' analyses, the computer program calculates both mean values and ''worst'' values (i.e., the mean of the three largest nuclear values), so ''worst'' values are obtained automatically from the three most atypical nuclei. Pearson correlation coefficients[33] between mean and worst size values are very high: 0.903 for mean and worst area, and 0.919 for mean and worst radius. Therefore, the authors believe that their results are not dependent on subjective selection of a special measurement area.

Robinson et al. reported a significant relation between visual grading of cytologic features of FNAs and histologic grading of 286 ductal carcinomas,[34] which implies prognostic significance of the cytologic features. The authors' work extends these observations and relates objectively assessed cytologic features directly with outcome.

The current study confirms the prognostic importance of nuclear features and shows that computer-derived nuclear features are prognostically strong compared with axillary lymph node status. In fact, nuclear feature stratification better separates the prognostically best from the intermediate group. The computer-assisted nuclear grading superseded the classic lymph node negative versus lymph node positive grouping in identifying the clinically important group of patients with a favorable prognosis. Only 20% of the most favorable patient group, identified either by lymph node negative status or by nuclear feature analysis, experienced distant metastases within 5 years. Such patients, when identified as being lymph node negative after axillary lymph node dissection, comprise approximately 50% of the total patients and have a sufficiently favorable prognosis that they usually are spared adjunctive chemotherapy. Now, a comparably sized group of patients with similarly favorable prognosis can be more accurately identified based on computer-assisted nuclear grading. The ability to do so without surgery becomes very important when neoadjuvant chemotherapy is considered. Since adding lymph node status and tumor size to nuclear feature analysis did not increase the authors' prognostic accuracy, prognostications may be based on computer-derived nuclear features alone and may obviate the need for prognostic axillary lymph node dissections. Even without knowledge of the axillary lymph node status, adjunctive chemotherapy or hormonal therapy could be based on computer-assisted nuclear grading and axillary surgery could be delayed until such time as axillary lymph nodes might become palpable.

In this study, the authors showed that their patients represented the global breast carcinoma population based on similarities according to lymph node stratification and outcome. They reported cross-validated rather than simple observed results to avoid the overoptimistic claims frequently made for newly developed prognosticators. Therefore, the authors anticipate that the prognostic importance of computer-derived nuclear features and their superiority to axillary lymph node status will be confirmed in larger studies.

In summary, computer analysis of a preoperative FNA more accurately identifies prognostically favorable patients than pathologic examination of axillary lymph nodes.

APPENDIX

The Recurrence Surface Approximation Method

The authors approached the prediction of TTR as a function estimation problem, a mapping of an n-dimensional input of cytologic and other features to a one-dimensional TTR output. Complicating the problem is the fact that TTR was known for only a subset of patients; for the remaining patients, only the time of their last check-up or disease free survival time (DFS) was known. By exploiting the straightforward manner in which inequalities are handled in linear programming, the authors were able to include all available cases in a least-error regression model to build accurate, robust predictors.

Intuitively, the authors wanted to fit the observed recurrences as closely as possible, and use the DFS of the censored cases as a lower bound on the TTR of that patient. Predictions for censored cases should be consistent with the observed recurrence rate for the given training set. These assumptions can be formulated into the following linear program:

$$\underset{w,\gamma,\nu,y,z}{\text{minimize}}\ e^T y + e^T z + \sum_{i=1}^{k} (s - er)_*^T H\nu_i$$

$$-y \le Mw + \gamma e - \log t \le y$$

$$-\mathrm{Nw} - \gamma \mathrm{e} + \log r \le z$$

subject to

$$-\nu_i \le e(N_i w + \gamma) - \log s \le \nu_1, \quad i = 1, \ldots k$$

$$\nu,\mathrm{y},\mathrm{z} \ge 0$$

The purpose of this linear program was to learn the weight vector w and the constant term γ that determine a recurrence surface $\mathrm{s} = \epsilon^{\mathrm{xw}+\gamma}$. Here ϵ is the base of the natural logarithm, x is the n-dimensional row vector of measured features, and s is the surface (in this case, an exponential surface defined on the feature space) that predicted TTR. The matrix M is an $m \times n$ matrix of the m recurrent points, with TTRs given by the m-dimensional column vector t. Similarly, the k censored points were collected in the $k \times n$ matrix N, and their last known DFS times were in the k-dimensional vector r. The vectors y and z represent the errors for recurrent and nonrecurrent points, respectively. Any difference between observed TTR and predicted time to $\mathrm{s} = \epsilon^{\mathrm{xw}+\gamma}$ was an error, wherease predicting a TTR smaller than an observed DFS was also an error. The objective minimizes the sum of these errors, using e^T, the transpose of a column vector of 1's of appropriate dimension.

The fitting error for censored cases was also measured by the term $\Sigma_{i=1}^{k} (s - er)_*^T H\nu_i$.

The *hazard curve h(t),* or instantaneous rate of recurrence, was constructed from all the training points and sampled at regular intervals (e.g., every 12 months) and stored in the vector h. The times of these samples are represented by the vector *s:* specifically, $s = [12, 24, 36, \ldots, 120]^{\mathrm{T}}$ and $h = [0.10, 0.07, 0.06, \ldots, 0.05]^{\mathrm{T}}$. The matrix H is defined as *diag(h),* that is, a square matrix with the elements of h along the diagonal and zeros elsewhere (Because the time of their study was necessarily limited, the authors estimated recurrences beyond 10 years by adding a point probability at 20 years, which accounted for the remaining recurrence probability and assumed a 30% cure rate). The step function (ζ) is defined as 1 if $\zeta > 0$ and 0 otherwise. The term $(\mathrm{s} - \mathrm{er})^{\P}$ H results in a row vector similar to h^T but eliminates the values of the hazard curve measured at times less than the observed DFS r, which did not count toward the observed error. The authors then took the dot product of that vector with the prediction errors at each time point in s. This term, together with the $e^T z$ term in the objective, gives the expected error of an overestimated nonrecurrent point.

Inclusion of irrelevant or redundant features in the predictive model can reduce its accuracy, making it too dependent on the particular cases on which it was trained and less useful in new cases. Therefore, the authors incorporated an automatic feature selection method[35] into the RSA mathematic program. A new term is added to the above objective function that minimizes the number of nonzero elements in the weight vector w, thereby reducing the number of input features in x that affect the prediction. This term is given a small positive weight. Therefore, the resulting objective selects among all discriminators the one that reduces the complexity of the model, and therefore produces predictive models that more accurately predict TTR in previously unseen cases.

REFERENCES

1. Wolberg WH, Street WN, Heisey DM, Mangasarian OL. Computerized breast cancer diagnosis and prognosis from fine-needle aspirates. *Arch Surg* 1995;130:511–6.
2. Wolberg WH, Street WN, Heisey DM, Mangasarian OL. Computer-derived nuclear grade and breast cancer prognosis. *Anal Quant Cytol Histol* 1995;17:257–64.
3. Kass M, Witkin A, Terzopoulos D. Snakes: active contour models. *Int J Comp Vision* 1988;1:321–31.
4. Williams DJ, Shah M. A fast algorithm for active contours. In: Anonymous. Proceedings of the Third International Conference on Computer Vision. Los Alamitos, CA: IEEE Computer Society Press, 1990:592–5.

5. Street WN, Wolberg WH, Mangasarian OL. Nuclear feature extraction for breast tumor diagnosis. Proceedings IS&T/SPIE International Symposium on Electronic Imaging 1993; 1905:861–70.
6. Wolberg WH, Street WN, Mangasarian OL. Machine learning techniques to diagnose breast cancer from image-processed nuclear features of fine needle aspirates. *Cancer Lett* 1994; 77:163–71.
7. Mandelbrot BB. The fractal geometry of nature. New York: W. H. Freeman and Company, 1977.
8. Lee ET. Statistical methods for survival data analysis. New York: John Wiley and Sons, 1992.
9. Cox DR. Regression models and life tables. *J R Stat Soc* 1972; 34:187–220.
10. Mangasarian OL, Street WN, Wolberg WH. Breast cancer diagnosis and prognosis via linear programming. *Operations Res* 1995;43:570–7.
11. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc* 1958;53:457–81.
12. Stone M. Cross-validatory choice and assessment of statistical predictions. *J R Stat Soc* 1974;36:111–47.
13. Gehan EA. A generalized Wilcoxon test for comparing arbitrarily single-censored samples. *Biometrika* 1965;52:203–23.
14. Fisher B, Ravdin RG, Ausman RK, Slack NH, Moore GE, Noer RJ. Surgical adjuvant chemotherapy in cancer of the breast. *Ann Surg* 1968;68:337–55.
15. Fisher B, Bauer M, Wickerham L, Redmond CK, Fisher ER. Relation of number of positive axillary nodes to the prognosis of patients with primary breast cancer. *Cancer* 1983; 52:1551–7.
16. Fisher B, Montague E, Redmond C, Barton B, Borland D, Fisher E, et al. Comparison of radical mastectomy with alternative treatments for primary breast cancer. A first report of results from a prospective randomized clinical trial. *Cancer* 1977;39:2827–39.
17. Abe O, Abe R, Asaishi K, Enomoto K, Hattori T, Iino Y, et al. Effects of radiotherapy and surgery in early breast cancer: an overview of the randomized trials. *N Engl J Med* 1995;333:1444–55.
18. Kissin MW, Querci della Rovere G, Easton D, Westbury G. Risk of lymphoedema following the treatment of breast cancer. *Br J Surg* 1986;73:580–4.
19. Aitken RJ, Gaze MN, Rodger A, Chetty U, Forrest APM. Arm morbidity within a trial of mastectomy and either nodal sample with selective radiotherapy or axillary clearance. *Br J Surg* 1989;76:568–71.
20. Black MM, Opler SR, Speer FD. Survival in breast cancer cases in relation to the structure of the primary tumor and regional lymph nodes. *Surg Gynecol Obstet* 1955;100:543–51.
21. Wied G, Bartels P, Bibbo M, Dytch H. Image analysis in quantitative cytopathology and histopathology. *Hum Pathol* 1989;20:549–71.
22. VanDiest PJ, Risse EKJ, Schipper NW, Baak JPA. Comparison of light microscopic grading and morphometric features in cytological breast cancer specimens. *Pathol Res Pract* 1989; 185:612–6.
23. Baak JPA, Kurver PHJ, Snoo-Niewlaat AJE, Graef S, Makkink B. Prognostic indicators in breast cancer: morphometric methods. *Histopathology* 1982;6:327–39.
24. Baak JPA, VanDop H, Kurver PHJ, Hermans J. The value of morphometry to classic prognosticators in breast cancer. *Cancer* 1985;56:374–82.
25. Stenkvist B, Westman-Naeser S, Vegelius J, Holmquist J, Nordin B, Bengtsson E, et al. Analysis of reproducibility of subjective grading systems for breast carcinoma. *J Clin Pathol* 1979;32:979–85.
26. Wittekind C, Schulte E. Computerized morphometric image analysis of cytologic nuclear parameters in breast cancer. *Anal Quant Cytol Histol* 1987;9:480–4.
27. Aubele M, Auer G, Falkmer U, Voss A, Rodenacker K, Rutquist L, et al. Improved prognostication in small (pT1) breast cancers by image cytometry. *Br Cancer Res Treat* 1995;36: 83–91.
28. Baak JPA, VanDiest PJ, Stroet-Van Galen C, Wisse-Brekelmans ECM, Matze-Cok E, Littooy JG. Data processing and analysis in the multicenter morphometric mammary carcinoma project. *Pathol Res Pract* 1989;185:657–63.
29. Wolberg WH, Street WN, Mangasarian OL. Computerized diagnosis of breast fine-needle aspirates. *Breast J.* In press.
30. Komitowski D, Janson C. Quantitative features of chromatin structure in the prognosis of breast cancer. *Cancer* 1990;65: 2725–30.
31. Pienta KJ, Coffey DS. Correlation of nuclear morphometry with progression of breast cancer. *Cancer* 1991;68:2012–6.
32. Baak JPA, Ladekarl M, Sorensen FB. Reproducibility of mean nuclear volume and correlation with mean nuclear area in breast cancer: an investigation of various sampling schemes. *Human Pathol* 1994;25:80–5.
33. Wilkinson L, Hill MA, Welna JP, Birkenbeuel GK. SYSTAT for Windows:Statistics. Evanston, IL: SYSTAT, Inc. 1992.
34. Robinson IA, McKee G, Nicholson A, D'Arcy J, Jackson PA, Cook MG, et al. Prognostic value of cytological grading of fine-needle aspirates from breast carcinomas. *Lancet* 1994; 343:947–9.
35. Bradley PS, Mangasarian OL, Street WN. Feature selection via mathematical programming. Madison (WI): University of Wisconsin Computer Sciences Department; 1995 Dec. Report No.: 95-21.