Computer-Derived Nuclear "Grade" and Breast Cancer Prognosis

William H. Wolberg, M.D., W. Nick Street, Ph.D., Dennis M. Heisey, Ph.D., Olvi L. Mangasarian, Ph.D.

Departments of Surgery, Human Oncology, and Computer Sciences
University of Wisconsin, Madison, WI

Dr. Wolberg is Professor of Surgery and Human Oncology, University of Wisconsin, Madison, WI
Dr. Street is Assistant Researcher in the Departments of Surgery and Computer Sciences, University of Wisconsin, Madison, WI
Dr. Heisey is Assistant Scientist in the Department of Surgery, University of Wisconsin, Madison, WI
Dr. Mangasarian is John von Neumann Professor of Mathematics and Computer Sciences, University of Wisconsin, Madison, WI

Address reprint requests to:
William H. Wolberg, M.D.
Department of Surgery, University of Wisconsin Clinical Sciences Center, 600 Highland Avenue, Madison, WI 53792
FAX 608 263 7652

This study was supported in part by Air Force Office of Scientific Research grant AFOSR F49620-94-1-0036 and National

Running title: Computer-derived prognosis

Keywords: Breast cancer, image processing, cytology, prognosis, machine learning

Abstract:

Visual assessments of nuclear grade are subjective, yet still prognostically important. Now, computer-based analytical techniques can objectively and accurately measure size, shape, and texture features that constitute nuclear grade.

The cell samples used in this study were obtained by fine needle aspiration (FNA) during the diagnosis of 187 consecutive patients with invasive breast cancer. Regions of FNA preparations to be analyzed were digitized and displayed on a computer monitor. Nuclei to be analyzed were roughly outlined by an operator using a mouse. Next, the computer generated a "snake" that precisely enclosed each designated nucleus. Ten nuclear features were then calculated for each nucleus based on these snakes. These results were analyzed statistically and by an inductive machine learning technique that we developed and call Recurrence Surface Approximation (RSA).

Both the statistical and the RSA machine learning analyses demonstrate that computer-derived nuclear features are prognostically more important than are the classical prognostic features, tumor size and lymph node status.

The prognostic value of nuclear grade of breast cancer has been well established since Black et al. [8] first described the relationship between prognosis and nuclear features in 1955. He described a grading system based on nuclear dissimilarity, coarseness of nuclear chromatin, presence of nucleoli, and presence of mitoses. His grading system [9] was modified by Bloom and Richardson [10] and subsequently many other investigators [10,12,15,22,25,28,33,35] confirmed the relationship between nuclear grade and prognosis. However, visual grading systems are vulnerable to intra- and interobserver variation [30]. Despite this problem, recent studies continue to demonstrate the prognostic importance of visually assessed nuclear grade [14, 29]. The individual components of nuclear grade are also prognostically significant [29, 40]. Recently, we developed and now use computer-based image analytic techniques for accurately and objectively quantitating such nuclear features [32].

Computer methodology allows measurements of the nuclear size, shape, and texture features that were previously visually assessed as "grade". Although size, texture, and some shape features have been objectively measured for many years [7,11,19] newer computer image analysis techniques permit accurate measurement of additional shape features [26,41]. The purpose of this paper is to evaluate the prognostic importance of a variety of computer-derived nuclear features analyzed statistically and by inductive machine learning.

Materials and Methods

Patients and Aspirate

The cell samples used in this study were obtained by fine needle aspiration (FNA) in order to diagnose a consecutive series of 187 patients with invasive breast cancer. The age distribution of the patients is given in Figure 1 and the tumor size and number of cancerous axillary lymph nodes are given in Table 1. Nuclear features of these samples were determined by our digital image analysis program, Xcyt [32].

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 being 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.

In our analysis, aspirates were classified as cancer based on histologic confirmation by surgical biopsy.

Image Preparation

The imaged area on the aspirate slides is visually selected for minimal nuclear overlap. The image used for digital analysis is generated by a JVC TK-1070U color video camera mounted on top of an Olympus microscope and the image is projected into the camera

with a 63 X objective and a 2.5 X ocular. The image is captured by a ComputerEyes/RT color framegrabber board (Digital Vision, Inc., Dedham MA) as a 640 x 400, 8-bit-per-pixel Targa file. An 8-bit-per-pixel grey scale image is used for the image analysis, since we are interested only in the darkness of the nuclei and their internal contrast rather than their color.

The User Interface (Xcyt)

The first step in successfully analyzing the digital image is to specify the exact location of each cell nucleus. A graphical computer program called Xcyt was developed that allows the user to input the approximate location of enough nuclei (10 to 20) to provide a representative sample. The program was developed using the X Window System and the Athena WidgetSet on a DECstation 3100.

A mouse is used to trace a rough outline of cell nuclei on the computer monitor. From this rough outline, the actual boundary of the cell nucleus is located by an active contour model known as a "snake"[17,38]. The "snake" is a deformable spline that seeks to minimize an energy function defined over the arclength 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 mathematical aspects of the snake calculations are described elsewhere [32].

Nuclear Features

Once the nuclei to be analyzed have been identified by the operator and have been enclosed by the computer-generated snakes, the computer calculates ten real-valued nuclear features for each nucleus [32]. These features are constructed such that higher values are typically associated with malignancy. Features were verified using idealized phantom cells [41]. Nuclear size is expressed by the Radius and Area features. Nuclear shape is expressed by Smoothness, Concavity, Compactness, Concave Points, Symmetry and Fractal Dimension features. Both size and shape are expressed by the Perimeter feature [41]. Nuclear Texture is measured by finding the variance of the grey scale intensities in the component pixels. The mean value, worst (mean of the three largest values), and standard error of each feature are computed for each image, resulting in a total of thirty features.
The pixel dimensions are one micron equals 3.11 pixels vertically and 3.85 pixels horizontally, so in our size calculations we used one square micron equals 12.25 pixels. Not all of the features are rotation-independent, but this should have a small effect on the computations and any effects are balanced out by the random orientation of nuclei in the training set. Additionally, all images are handled in the same manner, thereby canceling any biases.

Statistical analysis, RSA and cross validation

The mean time for distant recurrence in those patients who recurred was two years. A detailed statistical analysis of

the relation between the computer-generated nuclear features and presence or absence of distant metastases within 24 months of diagnosis was done in a 146 patient subset of these patients. This subset consisted of 118 patients who did not manifest distant recurrence within 24 months and 28 who recurred; excluded from analysis were the 41 of the original 187 patients who have not been followed a minimum of 24 months.

For our statistical analyses, the appearance of distant metastases by two years was the end point. However, for machine learning, prognostic estimations do not fit into the usual classification paradigm. RSA is our proposed solution to this estimation problem [23]. RSA uses linear programming to determine a linear combination of the input features that accurately predicts time to recurrence (TTR). The assumptions for the RSA approach are that: a)Recurrences actually took place at some point in time previous to their detection (observed TTR), although the time between recurrence and detection is probably small, and b)Observed disease-free survival time (DFS) is a *lower bound* to the recurrence time of that patient. These assumptions lead to a linear program to be solved for a given training set is as follows.

$$\underset{w,\gamma,\upsilon,y,z}{\text{minimize}} \qquad \frac{1}{m}e^{T}y + \frac{1}{k}e^{T}z + \frac{\delta}{m}e^{T}v$$

$$\text{subject to} \qquad \begin{array}{rcl} -v \le Mw + \gamma e - t & \le & y \\ -Nw - \gamma e + r & \le & z \\ v & \ge & 0 \\ y & \ge & 0 \\ z & \ge & 0 \end{array}$$

The purpose of the linear program is to learn the weight vector w and the constant term γ. These parameters determine a recurrence surface $s = wx + \gamma$, where x is the vector of measured features and s is the surface which fits the observed recurrence times and overestimates the disease free times. Here M is an m x n matrix of the m recurrent points, with recurrence time vector t. Similarly, the k non-recurrent points are collected in the k x n matrix N, and their last known disease free survival times in the vector r. The vectors y and z represent the errors for recurrent and non-recurrent points, respectively; overestimating the TTR of recurrences is considered and error, while predicting a TTR which is shorter than an observed DFS is also an error. The objective averages the errors over their respective classes. The v term, weighted by an appropriately small δ, forces underestimated recurrent points closer to the surface. Note that e is a vector of 1's of appropriate dimension.

As in classification, it is important to choose the right subset of features to get the best generalization. We choose an appropriate feature set in the following automatic fashion. A tuning set [21] constituting one tenth of the training cases is first set aside. The RSA linear program is then solved using all of the input features, and the resulting surface is tested on the tuning set. Features are then removed, one by one, by setting the smallest (in magnitude) element of the coefficient vector *w* to zero. Each new problem is solved and the result tested on the tuning set, until only one feature remains. Using the features which showed the best performance on the tuning set, we then re-optimized using all the training data (e.g., restore the tuning set). In this manner, we can use the tuning set to select the complexity of the model without paying the penalty of losing some of the training set.

In order to evaluate the effectiveness of the RSA, it was fit into a cross-validation [31] framework, which estimates the accuracy of the learned models on future data. As is typical in the machine learning community, the cross-validation procedure is used only for estimation of performance, not for parameter choosing.

Statistical analyses were done with Systat [37] and SAS [1] software.

Results:

t-Test analysis of significance of features relative to distant disease recurrence at two years in the 147 patient subset shows the unimportance of the number of cancer-containing axillary lymph nodes and tumor size relative to computer-derived nuclear features (Table 2). The computer-derived features of mean Radius, mean Perimeter, mean Area, mean Fractal dimension, worst Radius, worst Perimeter, standard error of Area, worst Area, and worst Fractal dimension were stronger predictors of two year distant recurrence than were tumor size and lymph node status.

A three factor principal components analysis component loadings of features relative to distant disease recurrence at two years was done because many features were highly correlated. Principal component analysis identifies closely correlated features. Component analysis into three groups are reported in Table 3. Most of the sum of the total variance is explained by Component 1 which consists of the sum of nuclear size features. Lesser amounts of the total variance are explained by Component 2 which consists of the sum of nuclear shape features, and Component 3 which consists of the sum of Tumor size and Number of nodes involved with cancer.

The relation between the appearance of distant metastases by two years with the components taken independently are shown by

the Univariate t test results (Table 4). Univariate analysis indicates that Component 1 (nuclear size features) is the only set of features that relates significantly to distant recurrence within two years. When these components are taken together in a discriminant analysis, the multivariate analysis of variance F statistic 13.246, $p<0.001$

Machine learning with ten-fold cross validation was used to select the most useful features for prognostic models. Computer-derived nuclear features were selected 97.8% of the time, whereas tumor size was selected only 1.7% of the time, and lymph node status only 0.5% of the time (Table 5). The Kaplan-Meier [16] plot of the RSA data derived by cross validation is shown as Figure 2. The time during which distant metastases were predicted to occur by RSA are divided into three groups; 54 patients predicted to recur within two years, 98 between two and five years, and 35 after five years. The RSA average error for the appearance of distant metastases with our computer-generated nuclear "grade" was 16.5 months.

In a Cox analysis, a nuclear size feature was selected as the single most prognostically important feature in each of the five best single-feature models (Table 6). After a nuclear size feature was selected, either the number of metastatic lymph nodes or tumor size was chosen in each of the five best two-feature models (Table 6). In the Cox proportional hazards analysis for the two-feature model, the p-value was 0.0001 for the largest

nuclear area and was 0.0069 for the number of metastatic axillary lymph nodes.

Discussion:

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 [36]. The results of computer-based analyses are reproducible and correlate closely with visual assessments [34]. Such techniques demonstrate that larger nuclear size is associated with a poor prognosis [3,6,30,39]. Two studies [3,6] also found variation in nuclear size, as reflected in the standard deviation of nuclear-size features, to be prognostically unfavorable. Currently, a large trial is underway evaluating morphometry (nuclear area and axes ratio) relative to other prognostic factors [5].

In contrast to the methods used in other studies, segmentation (i.e.determination of nuclear boundaries) are determined automatically by the computer "snake" program. Furthermore, our studies use the cellular smear-type preparations in which nuclear detail is better preserved than in the histological preparations used in many of the previous studies. FNA-obtained cells are preserved intact whereas histologically processed cells are cut in various planes. Despite these technical differences, our prognostic accuracy is almost identical to that reported by Komitowski and Janson [20]. They

used projection microscopy and a digitizing tablet to determine size, shape, and texture features in 60 breast cancer patients. They achieved 85% prognostic accuracy; inclusion of tumor size increased accuracy to 92%. Pienta and Coffey [27] 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 cancer.

In our 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 as compared to a random selection process. In a series of breast cancers, Baak et. al. studied the results of nuclear size measurements made selectively by an operator who chose maximally atypical areas with those made randomly [4]. 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 our 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 [37] 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, we believe that our results are not dependent on subjective selection of a special measurement area.

Our observations corroborate those of others that nuclear morphometric features provide prognostic information independent of that derived from tumor size and axillary lymph node involvement with metastatic disease. Mittra and MacRae [24] did a meta-analysis of the relationship between prognosis and an assortment of clinical, microscopic, and biological prognostic factors. A general interrelationship was found between each of the seven biological prognostic factors and between these factors and tumor grade. However, the biological factors and tumor grade were not correlated with the clinical prognostic factors of axillary lymph node status and tumor size. In the Nottingham index, grade emerged as prognostically more powerful than tumor size or lymph-node status and was the only prognostically significant factor on multivariate analysis [14].

Now, objectively and accurately assessed nuclear features are found to be better indicators than are other prognostic features, such as tumor size and lymph node status. Our entire analytical process, including generation of the patient's specific Kaplan-Meier curve for probability of distant recurrence, takes about ten minutes. Should others confirm findings that computer-derived nuclear grade are as prognostically accurate as is lymph node status, patients can be spared axillary surgery for prognostic purposes. Not only would the magnitude of breast surgery be reduced, but the 20% incidence of arm lymphedema following axillary dissection [2,18] would be avoided.

We conclude that computer-generated nuclear features, and particularly size features, are prognostically more important than are tumor size, and lymph node status. The methods described in this paper provide the basis for utilizing computer programs to determine prognosis from computer-generated nuclear "grade".

References:

1. SAS Institute Inc. editor.SAS/STAT User's Guide, Version 6. 4th ed. Cary, NC: SAS Institute Inc. 1989;

2.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.76:568-571, 1989.

3.Baak JPA, Kurver PHJ, Snoo-Niewlaat AJE, Graef S, Makkink B. Prognostic Indicators in Breast Cancer-Morphometric Methods. Histopathology.6:327-339, 1982.

4.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. Pathology.25:80-85, 1994.

5.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. Path Res Pract.185:657-663, 1989.

6.Baak JPA, VanDop H, Kurver PHJ, Hermans J. The Value of Morphometry to Classic Prognosticators in Breast Cancer. Cancer.56:374-382, 1985.

7.Bartels P, Bahr G, Bibbo M, Wied G. Objective Cell Image Analysis. The Journal of Histochemistry and Cytochemistry.20:239-254, 1972.

8.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.100:543-551, 1955.

9.Black MM Speer FD. Nuclear structure in cancer tissues. Surg Gynecol Obstet.105:97-102, 1957.

10.Bloom H Richardson W. Histological grading and prognosis in breast cancer. Br J Cancer.19:359-377, 1957.

11.Cornelisse CJ, deKoning HR, Arentz PW, Raatgever JW, van Heerde P. Quantitative analysis of the nuclear area variation in benign and malignant breast cytology specimens. Analy Quant Cytol and Hist.3:128-134, 1981.

12.Fisher ER, Redmond C, Fisher B, Bass G. Pathologic findings from the National Surgical Adjuvant Breast and Bowel Projedts (NSABP). Prognostic discriminants for 8-year survival for

node-negative invasive breast cancer patients. Cancer.65(supp):2121-2128, 1990.

13.Fleege JC, van Diest PJ, Baak JPA. Computer assisted efficiency testing of different sampling methods for selective nuclear graphic tablet morphometry. Lab Invest.63:270-275, 1990.

14.Galea MH, Blamey RW, Elston CE, Ellis IO. The Nottingham Prognostic Index in primary breast cancer. Breast Cancer Res. Treat.22:207-219, 1992.

15.Henson DE, Ries L, Freedman LS, Carriaga M. Relationship among outcome, stage of disease, and histologic grade for 22,616 cases of breast cancer. Cancer.68:2142-2149, 1991.

16.Kaplan EL Meier P. Nonparametric estimation from incomplete observations. J Am Statist Assoc.53:457-481, 1958.

17.Kass M, Witkin A, Terzopoulos D. Snakes: Active contour models. International Journal of Computer Vision.1:321-331, 1988.

18.Kissin MW, Querci della Rovere G, Easton D, Westbury G. Risk of lymphoedema following the treatment of breast cancer. Br J Surg.73:580-584, 1986.

19.Klaus H, Roth K, Hufnagl P, Wildner GP. Automated microscope image analysis in the cytological tumor diagnostics.Results on borderline-cases of mammary tumors. Arch Geschwulstforsch.55:259-264, 1985.

20.Komitowski D Janson C. Quantitative features of chromatin structure in the prognosis of breast cancer. Cancer.65:2725-2730, 1990.

21.Lang K, Waibel A, Hinton G. A time-delay neural network architecture for isolated word recognition. Neural Networks.3:23-43, 1990.

22.Le Doussal V, Tubiana-Hulin M, Friedman S, Hacene K, Spyratos F, Brunet M. Prognostic value of histologic grade nuclear components of Scraff-Bloom -Richardson (SCR): An improved score modification based on multivaraiate analysis of 1262 invasive ductal breast carcinomas. Cancer.64:1914-1921, 1989.

23.Mangasarian OL, Street WN, Wolberg WH. Breast cancer diagnosis and prognosis via linear programming. University of Wisconsin, Computer Sciences Department, Mathematical Programming Technical Report 94-10. 1994.

24.Mittra I MacRae KD. A Meta-analysis of Reported Correlations between Prognostic Factors in Breast Cancer: Does Axillary Lymph

Node Metastasis Represent Biology or Chronology? Eur J Cancer.27:1574-1583, 1991.

25.Mouriquand J, Gozlan-Fior M, Villemain D, Bouchet Y, Sage JC, Mermet MA, Bolla M. Value of cytoprognostic classification in breast carcinomas. J Clin Pathol.39:489-496, 1986.

26.Partin AW, Walsh AC, Pitcock RV, Mohler JL, Epstein JI, Coffey DS. A Comparison of Nuclear Morphometry and Gleason Grade as a Predictor of Prognosis in Stage A2 Prostate Cancer: A Critical Analysis. J Urol.142:1254-1258, 1989.

27.Pienta KJ Coffey DS. Correlation of nuclear morphometry with progression of breast cancer. Cancer.68:2012-2016, 1991.

28.Rank F, Dombernowsky P, Jespersin NCB, Pedersen BV, Keiding N. Histologic malignancy grading of invasive ductal breast carcinoma. Cancer.60:1299-1305, 1987.

29.Robinson IA, McKee G, Nicholson A, D'Arcy J, Jackson PA, Cook MG, Kissin MW. Prognostic value of cytological grading of fine-needle aspirates from breast carcinomas. Lancet.343:947-949, 1994.

30.Stenkvist B, Westman-Naeser S, Vegelius J, Holmquist J,

subjective grading systems for breast carcinoma. J Clin Path.32:979–985, 1979.

31.Stone M. Cross–validatory choice and assessment of statistical predictions. Journal of the Royal Statistical Society.36:111–147, 1974.

32.Street WN, Wolberg WH, Mangasarian OL. Nuclear feature extraction for breast tumor diagnosis. Proceedings IS&T/SPIE International Symposium on Electronic Imaging.1905:861–870, 1993.

33.Todd JG, Dowle C, Williams MR, Elston CW, Ellis IO, Hinton CP, Blamey RW, Haybittle JL. Confirmation of a prognostic index in primary breast cancer. Br J Cancer.56:489–492, 1987.

34.VanDiest PJ, Risse EKJ, Schipper NW, Baak JPA. Comparison of light microscopic grading and morphometric features in cytological breast cancer specimens. Path Res Pract.185:612–616, 1989.

35.Wallgren A Zajiecek J. The Prognostic Value of the Aspiration Biopsy Smear in Mammary Carcinoma. Acta Cytol.20:479–485, 1976.

36.Wied G, Bartels P, Bibbo M, Dytch H. Image Analysis in Quantitative Cytopathology and Histopathology. Human Pathology.20:549-571, 1989.

37.Wilkinson L, Hill MA, Welna JP, et al.SYSTAT for Windows:Statistics. 5th ed. Evanston, IL: SYSTAT, Inc.; 1992;

38.Williams DJ, Shah M. A fast algorithm for active contours. Proc. Third Int. Conf. on Computer Vision. AnonymousLos Alamitos, CA: IEEE Computer Society Press, 1990;p. 592-5.

39.Wittekind C Schulte E. Computerized morphometric image analysis of cytologic nuclear parameters in breast cancer. Analy Quant Cytol and Hist.9:480-484, 1987.

40.Wolberg WH, Bennett KP, Mangasarian OL. Cell analysis enhances diagnosis and predicts outcome. [Abstract] Proc Amer Assoc Cancer Research 1992;33:59

41.Wolberg WH, Street WN, Mangasarian OL. Machine learning techniques to diagnose breast cancer from image-processed nuclear features of fine needle aspirates. Cancer Letters.77:163-171, 1994.

Table 1: Tumor size and number of cancerous axillary lymph nodes in the 187 patients. The values in the chart are the number of patients in each category.

	Positive nodes			
Tumor size (cm.)	0	1–3	4–9	>9
<1	4	2	0	0
1–3	70	42	5	5
>3	11	10	9	6

Table 2: p values for the relationship between features and distant disease recurrence by two years. 118 patients had not recurred and 28 had.

p<0.001	0.001<p<0.01	0.01<p<0.06
mean Radius	std.error Area	Tumor size
mean Perimeter	worst Area	Lymph node status
mean Area	worst Fractal dim.	mean Concave points
mean Fractal dim.		mean Symmetry
worst Radius		std.error Radius
worst Perimeter		std.error Perimeter
		worst Concave points

Table 3: Principal components analysis, three factors, component loadings of features relative to distant disease recurrence at two years. Principal component analysis identifies closely correlated features. Most of the sum of the total variance is explained by Component 1 which consists of the sum of nuclear size features. Lesser amounts of the total variance are explained by Component 2 which consists of the sum of nuclear shape features, and Component 3 which consists of the sum of Tumor size and Number of nodes involved with cancer.

Principal factor analysis, 3 components, component loadings

	Component 1	Component 2	Component 3
Worst perimeter	0.969	0.019	0.023
Worst radius	0.966	−0.053	0.026
Worst area	0.953	−0.055	0.005
Mean area	0.951	−0.153	0.112
Mean perimeter	0.947	−0.095	0.133
Mean radius	0.939	−0.166	0.127
Std error area	0.875	0.100	−0.265
Std error radius	0.785	0.178	−0.340
Std error perimeter	0.754	0.227	−0.320
Mean concave points	0.680	0.573	0.144
Mean fractal dim.	−0.332	0.886	0.030
Mean symmetry	−0.033	0.800	−0.075
Worst fractal dim.	−0.348	0.787	0.206

Worst concave points	0.420	0.743	0.160
Nodes involved	0.124	−0.041	0.775
Tumor size	0.210	−0.118	0.739
PERCENT of TOTAL VARIANCE EXPLAINED	52.155	19.369	9.862

Table 4: Univariate t test of the relation of the three components and the appearance distant metastases by two years. The relation between these three components taken independently and distant metastases by two years are shown by the Univariate t test results.

Univariate t tests

Variable	t	p
Component 1	5.00	<0.001
Component 2	1.11	0.269
Component 3	1.80	0.072

Table 5: Number of times each feature was used to select the best RSA prognostic model:

Feature	mean	std error	worst
Radius	64	36	66
Perimeter	60	21	31
Fractal dimen.	57	21	24
Area	44	33	50
Compactness	28	35	39
Concavity	28	28	34
Smoothness	20	10	13
Concave points	17	13	10
Symmetry	15	20	13
Tumor size	15		
Texture	9	7	7
Lymph nodes	4		

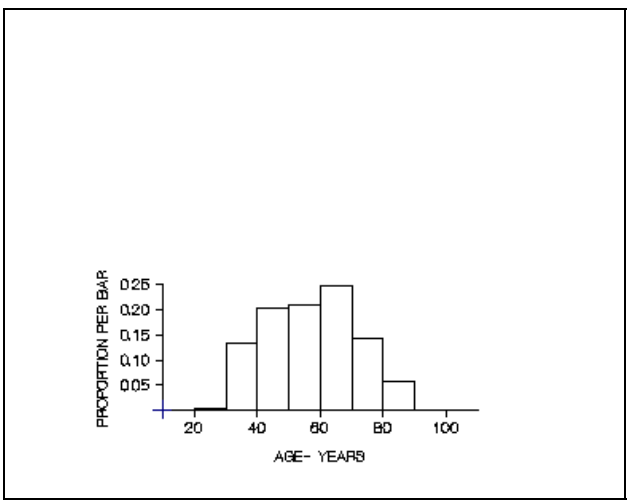
Table 6: Cox Regression Models Selected by Score Criterion

Number of Features	Score Value	Features
1	16.870	Mean Area
	16.706	Worst Area
	16.612	Worst Perimeter
	16.099	Worst Radius
	14.948	Mean Perimeter
2	23.540	Worst Area, Lymph nodes
	23.534	Worst Perimeter, Lymph nodes
	23.054	Worst Radius, Lymph nodes
	22.723	Mean Area, Lymph nodes
	21.933	Worst Area, Tumor size
3	27.186	Mean Symmetry, Worst Perimeter, Lymph nodes
	27.029	Mean Radius, Mean Area, Lymph nodes
	26.581	Mean Radius, Mean Area, Tumor size
	26.175	Mean Symmetry, Worst Area, Lymph nodes
	26.002	Mean Symmetry, Worst Radius, Lymph nodes
4	29.331	Mean Symmetry, Std error Symmetry, Worst Perimeter, Lymph nodes
	29.085	Mean Radius, Mean Area, Mean Fractal dimension, Lymph nodes
	28.859	Mean Radius, Mean Area, Mean Symmetry, Lymph nodes
	28.562	Mean Radius, Mean Area, Tumor size, Lymph nodes

28.367 Mean Radius, Mean Area, Mean Fractal dimension, Tumor size

Legend for figure 1

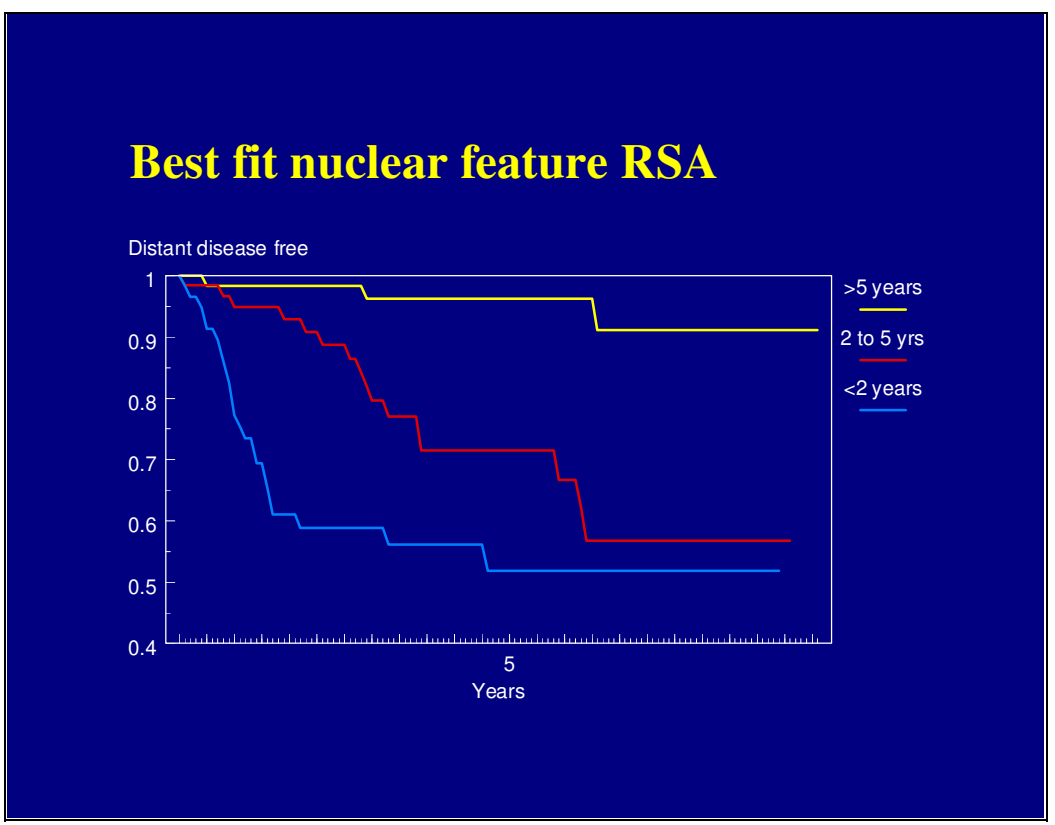
Age distribution of patients.



Legend for figure 2

Kaplan-Meier survival curve derived from RSA analysis. Predicted survivals are divided into three groups:

BLUE distant disease-free survival predicted to be less than two years (54 cases),

RED distant disease-free survival predicted to be between two and five years (98 cases),

and *YELLOW* distant disease-free survival predicted to be greater than five years (35 cases).