Cancer Chemother. Pharmacol. 2, 147–158 (1979)

Cancer Chemotherapy and Pharmacology

General Review

An Analysis of Predictor Variables for Adjuvant Treatment of Breast Cancer

S. Kister[1], J. Aroesty[2], W. Rogers[2], C. Huber[2], K. Willis[2], P. Morrison[2], G. Shangold[1], and T. Lincoln[2]

[1] Columbia University, College of Physicians and Surgeons, Department of Surgery, New York New York 10032, USA

[2] The Rand Corporation, Engineering and Applied Sciences Department 1700 Main Street, Santa Monica, California 90406, USA

Summary. *Modern computer-based statistical analysis of a well-documented data base can facilitate the selection of breast cancer patients for adjuvant chemotherapy. Traditional selection criteria are dominated by the number of positive axillary lymph nodes found by pathologic examination. However, patients of the same nodal status (0 +, 1–3 +, 4–7 +, > 7 +) are still heterogeneous with regard to risk of metastatic recurrence and benefits of adjuvant treatment. A pilot study of 103 patients of similar nodal status (1–3 +) followed for up to 10 years was undertaken to determine whether data already in the patient's record at the time of pathologic examination and derived from the patient's history as well as from clinical, surgical and pathologic examination could supplement nodal status in predicting disease recurrence. Preliminary processing and screening for promising variables were performed with* CLINFO*; maximum-likelihood procedures were then used to relate the probability of disease recurrence to those variables that appeared to be significant. Parametric models of hazard rate for the individual patient were employed corresponding to both exponential and Wiebull distributions of disease-free interval. The hazard rate was related log-linearly to a set of prognostic variables, and model parameters were determined by fitting to the data. Factors that favor longer disease-free intervals (in quantitative order of importance) are: (1) Nipple involved clinically at presentation; (2) Lesion had soft or rubbery consistency on palpation; (3) Disease discovered by physician; (4) Homolateral lymph nodes not involved clinically; (5) Margin from tumor to fascia > 1 cm; (6) No maternal history of breast cancer; (7) Increasing age of patient; (8) Presence of specialized histology.*

Based on the findings of this pilot study, a quantitative summary of personal (SK) and institutional experience is developed in which the probability of recurrence for the individual patient and the associated confidence intervals are used to classify patients with regard to risk of recurrence.

Introduction

In cancer of the breast and other neoplastic diseases that appear locally resectable, the probability of disease recurrence still remains significant and the time to recurrence, as measured by overt metastases, is highly variable. Clinical and pathologic staging procedures can be introduced to characterize the natural history of such diseases and to fix the patient at a particular stage in disease progression. The intention of staging is to reduce the variability in prognosis and to facilitate treatment decisions. In diseases such as carcinoma of the breast with limited nodal involvement, staging procedures are not yet precise. The potential for recurrence in the individual patient and the time to such recurrence still remains difficult to estimate. Recently (Bonadonna et al., 1977; Fisher et al., 1977), adjuvant chemotherapy has been employed to prolong the disease-free interval and to improve the prognostic outlook for each patient. However, such adjuvant therapy is expensive and both physiologically and psychologically stressful, and has the potential of introducing severe clinical complications in some patients. If we could improve our ability to predict relapse for a particular patient, then adjuvant chemotherapy could be given on a selective basis to those who are at significant risk for recurrence.

Recent advances in computer-based multivariate statistical analysis and computation make it possible to identify a set of prognostic variables that improve patient-specific predictions concerning the course of disease. Such variables can be identified by the analysis of a well-organized and documented data base even though a study is still in progress and the data are right-

Reprint requests should be addressed to: J. Aroesty

0344-5704/79/0002/0147/$ 02.20

censored. The best estimate of the risk of recurrence can be calculated for each patient's profile in the study, representing a formalization of the experience within the study group. These estimates can be updated easily as the study progresses. In this way a quantitative staging classification for the recurrence of overt metastases can be used to identify patients at high and low risk of recurrence, and those patients for whom the data are inconclusive.

The explicit numerical relations used in the estimate derive from multivariate maximum-likelihood calculations. These calculations may be based upon parametric or nonparametric representations of the time-dependent hazard rate[1] for disease recurrence and a parametric dependence of hazard rate on a set of prognostic variables. Computer programming capacity no longer lags behind statistical theory in this area: calculations using classic and clinically useful distributions and many prognostic variables can be carried out efficiently and inexpensively in the presence of significant data censorship.

A series of computer programs have recently been developed at Rand to analyze the situation where the probability of disease-free interval for the individual patient is represented by either an exponential or a Wiebull distribution in time. The former corresponds to a constant hazard rate and the latter corresponds to a hazard rate that may either increase or decrease with elapsed time since mastectomy. A recent analysis (Mueller and Jeffries, 1975) of data obtained from a multi-institutional study of over 10,000 breast cancer patients suggests that the hazard rate for survival decreases with time since surgery. This result was based on aggregate data without considering individual patient characteristics, and may not apply when patient-specific characteristics are included in the determination of parameters. Nevertheless, the constant hazard rate is the simplest plausible model for preliminary investigations with limited data and follow-up. This is especially true if the extent of disease recurrence is small (as in the present series) and if individual patient characteristics are considered. In all our analyses, the relationship between hazard rate and prognostic variables is assumed to be a log-linear one.

The subject of prognostic factors in breast cancer is not without controversy (Haagensen, 1971; Rosencweig and Heuson, 1975; Fisher, E. et al., 1975). However, it is generally agreed (Haagensen, 1974; Fisher, B. et al., 1975) that the number of positive axillary nodes found by pathologic examination is the single most important prognostic factor for recurrence of disease. It was our conjecture, at the start of this analysis, that data already contained in the individual patient record and taken at the time of nodal examination and surgery could be used to supplement nodal status (0, 1–3, 4–7, $\geqslant$8 positive nodes) in predicting relapse. In the future, as clinical measures (Sklarew et al., 1977; McGuire et al., 1975; Coombes et al., 1977) of proliferative, biochemical, hormonal, and immunologic status become more effective, the augmented patient profile will yield even further precision in staging and prediction. The present analysis, however, is based solely on variables taken from patient history, clinical examination, surgery, and pathology, all of which are available at the time the breast specimen is examined.

As part of a pilot study, we have analyzed a personal (SK) series of 106 treated breast cancer patients of similar nodal status. These patients have (1) undergone a Halsted-Haagensen (Haagensen, 1971) radical mastectomy for carcinoma of the breast, and (2) had nodal involvement limited to one to three positive lymph nodes in the cleared and completely examined breast specimen. Patients meeting these criteria correspond to about 20% of all mastectomies performed by the senior author during the period 1967–1976.

Methods

Clinical

All patients were worked up according to the expanded protocol of Haagensen (Haagensen, 1974). Clinical examination, staging, surgery, and follow-up were performed by the senior author. Clinical staging according to the Columbia Clinical Classification system was employed. Two patients in the current series were staged as Columbia Classification C: the remaining patients were all A or B. (In a similar earlier series of Stage A and B patients, 63% survived at least ten years and 24% were found to have one to three positive lymph nodes [Haagensen, 1974].)

Each mastectomy specimen was examined for lymph nodes according to the method of Pickren (Pickren, 1956). This approach yields approximately 50% more nodes than a careful conventional examination. The three-sample-cut method (Pickren, 1961) increases the number of positive nodes by an average of more than 20% over the number found on examination of one section per node.

We found 106 cases with one to three positive nodes in the original data records. Data on these patients were collected at the times of initial examination, surgery, pathologic examination, and regular follow-up (see Appendix A). A vigorous follow-up procedure and the assistance of data personnel has achieved over 98% follow-up and an equivalent completeness of data files. Similar procedures have been in effect at Columbia University under Haagensen since 1943, and complete pathology records are available since 1935.

Data were substantially incomplete for three patients, who were omitted from further analysis. Since the patient entry rate has been nearly uniform, the data base is highly right-censored. Thirty-six patients have been followed for more than 5 years and three for more than 10 years. Of the 103 patients in this subgroup, 22 have exhibited overt recurrence of disease within 10 years.

1 The hazard rate at time t is the instantaneous probability of recurrence per unit time, conditional on surviving to time t without disease recurrence

Data Management and Preliminary Analysis

Coding sheets were used to abstract 118 variables for each patient, and the data were entered interactively into CLINFO (Groner et al., 1978), a data management and analysis computer system designed for clinical research. Exploratory data analysis and heuristic screening of all variables for possible prognostic significance were performed with the aid of CLINFO and STATLIB, a statistical software computer package developed at Rand and Bell Telephone Laboratories (Bradsford and Relles, 1975). Scatter plots were prepared and used to identify the individual variables that appeared to correlate with either recurrence or long disease-free intervals. In this way, a series of potential predictors was identified for further analysis by maximum-likelihood methods.

Analytical and Statistical

The primary statistical model used in the analysis assumes that the hazard rate for recurrence of overt metastases is independent of elapsed time since surgery, and depends on the characteristics of the individual patient[2]. For this model, the hazard rate for the i'th individual is λ_i, and the time interval to recurrence is exponentially distributed with expectation $1/\lambda_i$ (Fiegl and Zelen, 1965; Glasser, 1967; Prentice, 1973). The hazard rate λ_i is presumed to depend log-linearly on a set of patient-specific predictor variables, X_{ki}. Further details of the analytical and statistical methods are given in Appendix B1.

Forecasting Accuracy

Two aspects of uncertainty arise in estimating forecast accuracy for a particular patient. The first occurs because of the inherent randomness that must be ascribed to a particular individual, and is reflected in the probabilistic model of an exponential or Wiebull distribution for disease-free interval. The second occurs because the parameters of the particular distribution, represented by the vector β, are estimated from a small sample. Uncertainty about the true values of β may be translated into uncertainty about the probability of disease-free interval exceeding a specific duration. For this data base, at this time, 5 years appears to be a suitable interval for such considerations of uncertainty. A predictive horizon of over 5 years would be useful clinically (Haagensen, 1974). However, both the extent of right-censorship and the limited number of patients presently discourage attempts at longer-term forecasts.

The appropriate theory and equations for confidence statements are outlined in Appendix B2.

Results

Probability of Recurrence-Free Interval — All Patients

The Kaplan-Meier life-table estimate for the probability of disease-free interval is shown in Figure 1. Standard errors in this probability are ± 0.015 at 12 months and ± 0.06 at 84 months. The probability of a disease-free interval greater than 60 months is 0.73; the distribution is approximately exponential between 0 and 60 months and the average hazard rate is approximately 5%/year.

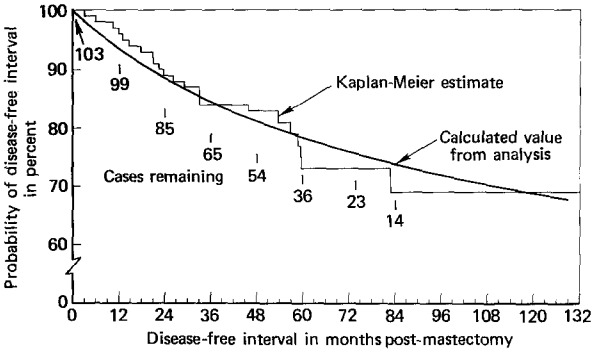

Fig. 1. Probability of disease-free interval versus time post-mastectomy (patients with 1–3 + axillary lymph nodes)

Variable Screening and Results of Maximum-Likelihood Calculations

Preliminary data screening indicated that six binary variables[3], method of discovery (DISCOVERY) *23*; consistency of lesion on palpation (CONSISTENCY) *30*; nipple changes or involvement at presentation (NIPPLE) *33*; margin from tumor to fascia (MARGIN) *57*; homolateral lymph nodes clinically involved (HOMOLATERAL) *35*; and histological type (HISTOLOGY) *58*; were particularly promising. An additional 17 variables, age at time of surgery (AGE) *3*; maternal history of breast cancer (MOTHER) *22*; paragravida *9*; age at first pregnancy *10*; age of youngest child at time of surgery *12*; history of past breast disease *20*; Columbia Clinical classification *40*; tumor-Node-Metastasis classification *41*; any postoperative radiotherapy *51*; associated pathologic findings *73*; pathologic size of lesion, greatest diameter *55*; fraction of axillary vein nodes that are positive *70*; fraction of central nodes that are positive *70*; fraction of scapular nodes that are positive *70*; fraction of Rotter's nodes that are positive *70*; number of positive axillary lymph nodes, total *66*; and menopausal status *11* were either plausible or deserved further study.

Recurrence was not seen in any patient who had any of the following characteristics (see Table 1): DISCOVERY = 1 (method of discovery: medical or clinical examination); CONSISTENCY = 1 (consistency of lesion on palpation: soft or rubbery); NIPPLE = 1 (nipple changes or involvement at presentation: positive includ-

2 Additional preliminary calculations were also performed with a Weibull distribution. These calculations and others to be described later suggest the suitability of the exponential model for these data

3 Numbers refer to *Patient Data Record* (Appendix A)

Table 1. Regression coefficients for constant hazard rate model

	Prognostic variable	β_k Coefficient	Estimated SD of the β_k	t-Statistic
0	CONSTANT	β_0 = 3.8786		
1	HOMOLATERAL	β_1 = -1.4314	0.5317	-2.692
2	MARGIN	β_2 = -2.0903	0.4793	-4.361
3	MARGIN-NA	β_3 = 2.0020	1.0719	1.868
4	AGE	β_4 = .0335	0.0183	1.831
5	HISTOLOGY	β_5 = 1.0568	0.5951	1.776
6	MOTHER	β_6 = -1.9734	0.7372	-2.677
7	DISCOVERY	β_7 = ∞	∞	+3.9
8	CONSISTENCY	β_8 = ∞	∞	+2.8
9	NIPPLE	β_9 = ∞	∞	+2.0

$$-\ln \lambda_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{4i} + \beta_5 x_{5i} + \beta_6 x_{6i} + \beta_7 x_{7i} + \beta_8 x_{8i} + \beta_9 x_{9i}$$

Code for Prognostic Variables

HOMOLATERAL (Homolateral lymph nodes clinically involved) 0: None, 0.13*: Not available, 1: Apparently involved; MARGIN (Margin from tumor to fascia) 0: ≥ 1 cm, 0.25*: Not available, 1: < 1 cm; MARGIN-NA (Answer to MARGIN not available) 0: MARGIN given, 1: MARGIN not given; AGE (Age at time of surgery—in years); HISTOLOGY (Histologic type) 0: No special type (including schirrous or invasive intraductal carcinoma), 0.25*: Not available, 1: Specialized histology (intraductal, papillary, solid circumscribed or medullary, small cell, Pagets, inflammatory, lobular carcinoma, multiple types, predominantly intraductal with foci of invasion, other); MOTHER (Maternal history of breast cancer) 0: No, 0.11*: Not available, 1: Yes; DISCOVERY (Method of discovery) 0: Self-examination, 1: Medical or clinical examination; CONSISTENCY (Consistency of lesion on palpation) 0: Firm, hard, gritty, 1: Soft, rubbery, other; NIPPLE (Nipple changes or involvement at presentation) 0: No nipple changes, 1: Positive including retraction, discharge, inversion (if other side normal), other

*Average values

ing retraction, discharge, inversion if other side normal, other changes).

The values of β_k in Table 1 for each of these variables (k = 7, 8, 9) are thus infinite, and patients with any of these favorable characteristics have zero hazard rate for recurrence. Confidence statements about the probability of 5-year disease-free intervals can still be computed, however. Maximum likelihood calculations were then performed to obtain β values for the remaining variables. The three in the first set (MARGIN, HOMOLATERAL, HISTOLOGY) were analyzed simultaneously and were found to be at least moderately significant.

The variables in the second group were then added to the analysis in sets of three or four at a time. Signifi-

Table 2. Group 1. Patients who have not recurred in ≥ 60 months

CASE NO.	DFI	DISCOVERY	CONSISTENCY	NIPPLE	MARGIN	HOMOLATERAL	HISTOLOGY	AGE-SURGERY	MOTHER	P5YR	LC95P5YR	UC95P5YR	CLASSIFICATION
2	129	0	0	0	0	0	0	52	0	0.80	0.63	0.90	0
7	123	0	1	0	NA	0	0	68	0	1.00	0.99	1.00	++
8	121	0	0	1	1	0	1	40	0	1.00	0.71	1.00	+
10	114	0	0	1	0	0	0	73	0	1.00	0.93	1.00	++
11	110	0	0	0	0	0	0	59	0	0.84	0.67	0.93	0
13	107	1	0	0	0	0	0	54	0	1.00	0.97	1.00	++
16	98	1	0	0	1	0	0	38	0	1.00	0.75	1.00	+
17	94	0	0	0	0	0	1	40	0	0.89	0.69	0.97	+
18	91	1	1	0	0	0	0	53	0	1.00	>0.99	1.00	++
19	90	0	0	0	0	0	1	63	0	0.95	0.78	0.99	+
23	89	0	0	0	NA	0	0	45	0	0.94	0.57	0.99	+
21	89	0	1	1	NA	NA	0	49	0	1.00	>0.98	1.00	++
22	89	0	1	0	0	0	0	51	0	1.00	0.93	1.00	++
24	87	0	0	0	NA	0	1	37	0	0.97	0.75	1.00	+
34	83	0	0	0	NA	0	0	41	0	0.93	0.52	0.99	+
26	83	1	0	0	1	0	1	46	1	1.00	0.75	1.00	+
35	83	0	0	0	0	0	0	27	0	0.61	0.23	0.84	-
37	77	0	1	0	0	0	0	70	0	1.00	0.96	1.00	++
29	76	0	0	0	0	0	0	68	0	0.88	0.69	0.96	0
30	76	0	NA	0	0	0	1	55	0	0.93	0.76	0.98	+
32	74	0	0	0	0	0	0	47	0	0.77	0.58	0.89	0
33	74	0	NA	0	0	0	0	48	0	0.78	0.59	0.89	0
1	70	0	1	0	1	0	0	74	0	1.00	0.80	1.00	++
40	70	0	1	0	0	0	1	65	0	1.00	0.99	1.00	++
41	70	0	0	0	0	1	1	56	0	0.76	0.30	0.94	0
43	69	0	0	0	NA	0	1	28	1	0.76	0.08	0.97	0
45	68	0	1	1	NA	0	0	37	0	1.00	>0.98	1.00	++
46	66	0	1	NA	0	0	0	33	0	1.00	0.89	1.00	++
52	65	1	0	1	0	NA	0	73	0	1.00	>0.99	1.00	++
47	65	0	0	0	0	0	0	41	0	0.73	0.50	0.87	0
48	64	0	0	NA	0	0	0	61	1	0.31	0.02	0.72	--
53	63	1	1	0	0	0	0	59	0	1.00	>0.99	1.00	++
50	62	0	0	0	0	0	0	67	0	0.88	0.69	0.95	0
54	61	1	0	0	0	0	0	67	NA	1.00	0.98	1.00	++
51	61	0	0	0	0	0	0	44	0	0.75	0.54	0.88	0

Notes:
- See Table 1
- DFI: Disease-free interval (censored 25 Jan. 1977)
- P5YR: Probability of 60-month disease-free interval (DFI)
- UC95P5YR: 95% upper confidence limit for P5YR
- LC95P5YR: 95% lower confidence limit for P5YR
- >: Lower confidence limit too conservative (See text)
- NA: Not available

Table 2. Group 2. Patients who have not recurred in < 60 months

CASE NO.	DFI	DISCOVERY	CONSISTENCY	NIPPLE	MARGIN	HOMOLATERAL	HISTOLOGY	AGE-SURGERY	MOTHER	P5YR	LC95P5YR	UC95P5YR	CLASSIFICATION
55	59	0	0	0	0	0	0	55	0	0.82	0.65	0.92	0
70	57	0	1	0	1	0	1	54	0	1.00	0.90	1.00	++
56	57	0	1	0	0	0	0	57	1	1.00	0.74	1.00	+
57	56	0	0	1	0	0	0	41	0	1.00	0.82	1.00	++
58	55	0	0	0	0	0	0	36	0	0.69	0.41	0.86	0
59	54	0	0	0	0	0	NA	73	0	0.92	0.75	0.98	+
61	53	0	0	0	0	0	1	46	0	0.91	0.73	0.97	+
62	53	0	1	1	0	0	1	52	0	1.00	>0.99	1.00	++
63	51	0	1	0	0	0	0	72	0	1.00	0.96	1.00	++
65	51	0	0	0	0	0	0	49	0	0.79	0.60	0.89	0
66	51	0	0	0	0	0	0	50	0	0.79	0.61	0.90	0
67	50	0	0	0	0	0	0	75	0	0.90	0.71	0.97	+
69	50	0	0	0	NA	1	0	80	0	0.92	0.39	0.99	+
84	48	0	0	0	1	NA	1	83	0	0.77	0.21	0.96	0
71	47	0	1	0	1	0	1	47	0	1.00	0.88	1.00	++
72	44	0	0	0	0	0	1	43	0	0.90	0.71	0.97	+
73	43	0	1	0	0	0	0	58	0	1.00	0.94	1.00	++
75	43	0	0	0	0	1	0	37	0	0.22	0.01	0.63	--
77	41	0	0	0	0	0	0	55	0	0.82	0.65	0.92	0
79	40	1	1	0	1	0	1	51	1	1.00	>0.92	1.00	++
80	40	1	0	0	0	0	0	43	0	1.00	0.96	1.00	++
83	37	0	0	0	0	1	0	59	0	0.49	0.15	0.76	--
49	37	0	0	0	0	0	0	52	0	0.80	0.63	0.90	0
100	34	0	NA	0	0	0	0	48	0	0.78	0.59	0.89	0
89	34	1	0	0	0	0	0	67	0	1.00	0.98	1.00	++
88	33	0	0	1	0	0	0	78	0	1.00	0.94	1.00	++
93	31	0	0	0	0	0	1	55	0	0.93	0.76	0.98	+
90	31	1	0	0	NA	0	0	44	0	1.00	0.99	1.00	++
92	30	1	0	0	0	0	0	49	1	1.00	0.85	1.00	++
91	30	0	1	1	1	0	1	48	1	1.00	>0.63	1.00	+
99	29	1	0	0	0	0	0	37	0	1.00	0.96	1.00	++
102	29	1	0	0	0	1	0	81	0	1.00	0.95	1.00	++
94	29	0	0	0	0	0	1	70	0	0.96	0.79	0.99	++
95	28	0	0	0	0	0	0	77	0	0.91	0.71	0.97	+
96	27	0	1	0	NA	0	1	58	0	1.00	1.00	1.00	++
97	27	0	0	0	1	0	0	78	0	0.48	0.12	0.78	--
85	27	1	0	0	NA	0	0	60	0	1.00	0.99	1.00	++
87	26	0	0	1	0	0	0	72	0	1.00	0.93	1.00	++
86	25	0	0	0	1	0	1	25	0	0.22	0.00	0.66	--
106	21	0	0	0	1	0	1	49	0	0.51	0.13	0.80	-
107	20	0	0	0	NA	1	0	70	1	0.44	0.00	0.92	-
108	18	1	0	0	NA	1	0	52	0	1.00	0.96	1.00	++
109	17	0	0	0	0	0	0	60	0	0.85	0.67	0.93	0
110	17	0	0	0	NA	1	0	70	0	0.89	0.32	0.99	+
103	17	0	1	0	0	0	0	72	0	1.00	0.96	1.00	++
104	13	0	1	0	0	0	0	46	0	1.00	0.92	1.00	++

Table 2. Group 3. Patients who have recurred in ≥ 60 months

CASE NO.	DFI	DISCOVERY	CONSISTENCY	NIPPLE	MARGIN	HOMOLATERAL	HISTOLOGY	AGE-SURGERY	MOTHER	P5YR	LC95P5YR	UC95P5YR	CLASSIFICATION
4	82	0	0	0	1	0	0	53	0	0.18	0.03	0.42	--

Table 2. Group 4. Patients who have recurred in < 60 months

CASE NO.	DFI	DISCOVERY	CONSISTENCY	NIPPLE	MARGIN	HOMOLATERAL	HISTOLOGY	AGE-SURGERY	MOTHER	P5YR	LC95P5YR	UC95P5YR	CLASSIFICATION
44	60	0	0	0	0	0	1	49	1	0.55	0.08	0.87	-
38	60	0	0	0	0	1	1	49	0	0.71	0.24	0.92	0
25	59	0	0	0	NA	0	0	38	0	0.92	0.49	0.99	+
28	56	0	0	0	1	0	0	65	0	0.32	0.09	0.59	--
5	53	0	0	0	0	0	0	53	0	0.81	0.63	0.91	0
20	46	0	0	0	0	0	0	42	0	0.74	0.51	0.87	0
64	32	0	0	0	0	1	0	69	0	0.60	0.21	0.84	-
60	32	0	0	0	1	0	0	68	0	0.36	0.10	0.64	--
9	29	0	0	0	1	0	0	50	0	0.15	0.02	0.39	--
12	26	0	0	0	1	0	1	29	0	0.27	0.01	0.68	--
74	23	0	0	0	0	0	0	71	1	0.44	0.05	0.80	-
3	21	0	0	0	0	0	0	52	0	0.80	0.63	0.90	0
36	20	0	0	0	1	1	0	63	0	0.01	0.00	0.17	--
14	20	0	0	0	1	0	0	41	0	0.08	0.00	0.34	--
15	18	0	0	0	1	0	0	70	0	0.38	0.10	0.67	--
98	14	0	0	0	0	0	0	65	1	0.36	0.03	0.75	--
31	12	0	0	0	0	1	0	48	0	0.35	0.06	0.68	--
81	11	0	0	0	1	0	1	45	0	0.46	0.10	0.77	--
76	10	0	0	0	0	1	0	54	0	0.43	0.11	0.72	--
39	6	0	0	0	0	0	0	26	0	0.59	0.21	0.84	-
78	2	0	0	0	1	0	0	45	0	0.11	0.01	0.36	--

cant changes in the magnitude of the log likelihood indicated that two additional variables, age at time of surgery (AGE) and maternal history of breast cancer (MOTHER), deserved inclusion as predictors, albeit not very strong ones. The parameter vector β was then calculated with these five variables: MARGIN, HOMOLATERAL, HISTOLOGY, AGE, and MOTHER. Missing data were usually replaced by average values of the available data. In the case of one variable, MARGIN, resolution of missing data by averaging was clearly not satisfactory. Of the 15 cases with missing values, 14 had not recurred, and it appeared inappropriate to replace these missing data points with average values. A missing value for the variable suggests a favorable prognosis and may follow from the tendency to record variables more constantly when they are considered significant. In this interpretation, a small distance between tumor margin and fascia would often be regarded as an unfavorable sign and would be documented, while larger distances would more frequently go unrecorded. A special analytical treatment of missing values for this variable was employed, in which a new variable was defined, MARGIN-NA, corresponding to a missing value. The variable was then employed as an additional prognostic variable in the analysis.

Table 1 presents parameter estimates, t-statistics, and pseudo-t statistics for the three special variables, DISCOVERY, CONSISTENCY, and NIPPLE. Table 2 summarizes values for the nine variables that enter into the prognostic profile for each patient as well as follow-up and outcome data. Table 2 also presents a summary of the calculated values for the probability of a 5-year disease-free interval for each patient and associated confidence intervals. The approximation used for estimating confidence intervals for patients with favorable values of any of the special characteristics (DISCOVERY = 1, CONSISTENCY = 1, NIPPLE = 1) assumes only one favorable indicator. For cases where two or more favorable indicators exist, the estimate for the lower confidence interval is too low. This is prognostically important for only one patient, No. 91.

The calculated values for the distribution of the probability of disease-free interval are exhibited in Figure 1, where they are shown to compare satisfactorily with the Kaplan-Meier estimate. The validity of the exponential model and the results of the Weibull calculation are described in Appendix C.

Discussion

Accuracy of Representation

The key element of this analysis is the calculated probability of the time to disease recurrence for the individual patient. This is derived from data contained within the patient's own case record as well as data from similar cases that have been treated in comparable ways. It is important to consider not only the point estimate of this probability of recurrence, which is predicted from the regression relations (see Table 1), but also the confidence statements associated with this estimate. The confidence intervals are large, but they can still be used to sharpen the usefulness of the prediction. It is also important to reiterate that these results derive from a pilot study of the natural history of breast carcinoma following radical mastectomy by a single surgeon. The data are limited to the 103 patients found, since 1966, to have one to three positive axillary lymph nodes on very complete examination by a single pathology team. In this group, nearly half (47) the patients have been followed for less than 5 years, and there have been only 22 cases of metastatic recurrence. The limited size of the data base, the limited number of recurrences, the right-censoring of the data, and the large number of significant prognostic variables all increase the width of the confidence intervals. This in turn leads to difficulty in refining the accuracy of these regressions.

If the data base were larger, the prognostic variables fewer, the study interval longer or more complete, and all the variables discrete, an appropriate estimate could be obtained by grouping patients according to combinations of prognostic variables. Each subgroup could then be compared by performing a separate life-table analysis. This approach has been used to analyze and eva-

luate results obtained from nearly 2000 cases of prostatic cancer obtained from a large multi-institutional study (Byar et al., 1974). It is clear that there is a tradeoff between the consistency arising from the observation, treatment, and follow-up by a single team and the statistical advantages of larger numbers. Because of the diversity of clinical approaches to radical mastectomy for carcinoma of the breast, and because of the wide variation in both the pathologic examination of the specimen and interaction between pathologist and surgeon at the time of surgery, the narrower focus is appropriate to this study[4].

Corresponding to this narrower focus is a less comprehensive but still satisfactory procedure for evaluating the accuracy of the analysis. A fixed disease-free interval (DFI) is used as a criterion for comparison. Here the data permit the conventional 5-year interval. (Given the slow evolution of this disease and the appearance of a significant number of recurrences up to 10 years after diagnosis, a longer interval would be useful, but the data will not support it.) In this data base, 36 patients have been followed for at least 5 years without recurrence, but only three have been followed for 10 years. Of the 22 recurrences, 21 occurred within 5 years, 13 of whom could have been followed for at least 5 years. The entire analysis is based on the complete set of 103 cases, but evaluation of the accuracy of the numerical representation must be restricted to these 57 cases whose 5-year follow-up is complete.

From Table 1, it can be seen that the mean value overall for the calculated probability of completing a 5-year DFI is 0.785 for the 103 cases. The mean value for the calculated probability for the 21 patients who recurred within 5 years is 0.45. The mean value is 0.47 for the 13 patients who could have been followed for at least 5 years if they had not recurred. This contrasts with a mean value of 0.875 for those patients who survived 5 years without recurrence. Thus there is a substantial separation in calculated probability between those who recurred and those who did not. Of course, some separation is to be expected on the basis of fitting of data.

For a preliminary evaluation of the prediction, the patients are classified into two groups: those whose computed 5-year probability is greater than the mean (0.785) and those whose computed probability is equal to or less than the mean. On this basis, 18 of the 21 patients who have actually recurred within 5 years are in the below-average category, and 27 of the 36 patients who have survived for 5 years without recurrence are in the above-average category. Thus 86% of those who recur are accurately represented as having a below-average prognosis, and 75% of those who have not recurred are designated as having an above-average prognosis. These comparisons demonstrate that the calculated (or predicted) probability of a 5-year DFI is a promising parameter for classifying patients with regard to risk of recurrence. Although the separation of patients into only two prognostic categories correctly classifies 80% of the patients, the scheme is both too coarse and too dichotomous for clinical decisions.

A Preliminary Clinical Classification Scheme

It is premature to propose definitive rules for the use of this analysis to facilitate patient selection for adjuvant chemotherapy. There is no unique correspondence among predicted disease-free interval, confidence intervals, and clinical risk status, and any approach to staging must reflect the judgment of the investigator and the data upon which it is based. In the present instance, the predictors are based on data derived from all 103 patients, but comparisons and stages are based only on those patients with requisite 5-year histories. Thus, the proposed classification is a preliminary one, to be modified as further data are obtained. We adopt a five-stage classification scheme for expressing predictions in a clinically useful way: patients who are at very low (+ +) or very high (– –) risk, patients with predictors which tend toward low (+) or high (–) risk, and patients in whom the prognosis appears average or equivocal (0). To identify those patients whose prognosis can be predicted with a relatively high level of confidence, we define as very high risk (– –) those patients whose 95% upper confidence limit (UC95 P5YR) is equal to or less than the mean (0.785), and as very low risk (+ +) those patients whose 95% lower confidence limit (LC95 P5YR) is greater than the mean. In this way we classified as very-high-risk cases 12 patients who recurred within 5 years and only two[5] of those who survived for 5 years without recurrence. Similarly, we classified as very-low-risk patients 14 of those who survived for 5 years without recurrence, and none of those patients who recurred.

Classification of the remaining 29 patients is more subtle, and is sensitive to limited data size, individual judgment, and any inaccuracies in the model itself. The predictions for these 29 patients are divided into three categories: tend toward high risk (–), tend toward low risk (+), and equivocal (0). In this fashion the predictions may influence but should not dominate therapeutic

4 The diversity associated with large multicenter studies can influence the analysis of treatment outcome. Methods to account for this diversity in the context of colo-rectal cancer trials have recently been proposed (Fielding et al., 1978). That analysis suggests that traditional patient-related predictor variables should be supplemented by variables reflecting the experience, technique, complication and mortality rates, etc. for the individual surgeon

5 One of these subsequently metastasized

judgment about these patients. The average or equivocal category should buffer the trends toward high or low risk and reduce the sensitivity of the analysis to missing data or judgment about a single variable. Empirically, and somewhat arbitrarily, we suggest that a reasonable separation of the remaining cases occurs when those patients whose computed probability lies in the interval $\pm$ 0.1 about the average values ($0.685 < P5YR < 0.885$) are designated as average or equivocal (0). On this basis 14 cases are considered to be equivocal. An additional five patients then show a trend toward increased risk (−), four of whom have already recurred within 5 years. Of the ten patients showing a trend toward decreased risk (+), only one has recurred within 5 years.

If the patient population included in this data base is typical of future patients, then over 70% of patients who would recur within 5 years would be correctly identified as being high-risk or tending toward high risk, and few such patients would be identified (incorrectly) as being at less than average risk. Moreover, only 10% of those patients who do not recur within 5 years would be designated (incorrectly) as being at an above-average 5-year risk. There is no way at present to determine whether these patients are indeed at a higher than average risk, unless a longer follow-up interval is available. This possibility could be explored on a data base collected and followed over a longer time interval, such as that reported by Haagensen (Haagensen, 1974). Since the same limited data base has been used for both analysis and evaluation, further data, analysis and consideration of risks and benefits of adjuvant therapy are required before a definitive clinical scheme can be proposed.

Applicability of These Relations

The broader applicability of the specific relations and prognostic variables developed during this study depends on two factors: (1) The extent to which the present patient data base is typical with respect to procedures and outcomes for cases of similar nodal status; and (2) How accurately the present data are represented by the analysis.

The second factor has already been considered in part, but further evaluation is planned. This evaluation will be based on a forthcoming analysis of data records obtained from an earlier large patient series, treated in the same institution under comparable criteria and procedures (Haagensen, 1974). These records are especially suitable for analysis, because the patients have been followed for at least 10 years and not more than 2% have been lost to follow-up. Not only does the extended interval of follow-up minimize the impact of data censorship, leading to more accurate predictions at longer intervals, but also, these data will permit a study of the stability of inferences made over a shorter time period. In particular, the current analysis assumes that the hazard rate for recurrence does not depend on the interval since primary treatment. This is the simplest plausible assumption and is consistent with the observation that relapses occur more than 20 years after mastectomy. Our preliminary period of follow-up in the present data is too short to permit a suitable evaluation of this point, particularly as it affects the calculated probability of a 5-year DFI. See Appendix C for further details.

The first factor, the issue of the typicality of the present data base, is even more problematic. If a comparable body of data were obtained for patients treated by other physicians at other institutions, then a similar analysis could be performed, and the results compared with those of the present study. It is unlikely, however, that such a comparable body of data already exists, because each institution tends to examine the patient, select and carry out the procedure, and evaluate the specimen in its own special way. The present results depend upon the consistent and careful evaluation of a patient population that has been treated and followed in an unusually complete and stable fashion over time by a small team of individuals. Nevertheless, the present prognostic variables are largely surrogate expressions of the disease process and cannot be considered as uniquely significant. It is possible that a different but equally suitable set of surrogate variables could be found in another set of case records, if they were also obtained in a consistent and stable manner.

A particularly good example of a surrogate variable is the variable labeled DISCOVERY, which refers to how the disease was initially discovered. In the present sample, an overwhelming number of lesions (87/103) were self-discovered; however, none of the 16 cases that were not self-discovered has recurred to the present time. Thus the variable not only has a very strong weight, but also appears to be counter to expectation. Why should self-discovery have a worse prognosis than discovery by a physician? The answer may be related to special aspects of the present patient population, but it is more likely to depend upon the selection of the subgroup with minimal nodal spread, coupled with differences in patient and physician approaches to breast examination. First, self-discovery by the patient does not necessarily lead to immediate action, and second, the techniques of self-palpation and physician palpation can be expected to discover different lesions. For example, a subtle lesion close to the fascia should be more easily palpable by a patient, whether by accident or intent, while a similar early lesion close to the nipple (with a better prognosis in this series) would be more easily discovered by the physician. Other factors must also play a role, including the judgment as to when a tumor is self-discovered and when it is discovered by the physician. This extreme

example serves to emphasize the possible indirect nature of the prognostic variables. In most cases, however, these prognostic variables can be related in a relatively straightforward way to the pathologic geography and extent of the disease or its biological aggressiveness.

In the present study, the method of pathologic examination has a strong influence on the results. Clearing the specimen and transilluminating it in the search for lymph nodes yields both more nodes for histologic study and more positive nodes per specimen. For this reason, the present classification of nodal status is somewhat more conservative than that employed in other institutions. If less exacting methods of pathology were employed, some patients in the current series would be assigned to the zero-positive-node category, and other patients with four or more positive nodes by the present method of pathologic examination would be considered elsewhere to have three or fewer. Given that there are multiple strong predictive variables, it would appear to be more useful and less arbitrary to consider the number of positive nodes as one predictive variable to be weighted in the same way as the others.

Conclusion

In carcinoma of the breast treated by radical mastectomy, clinical decisions about adjuvant chemotherapy are based primarily on the number of positive lymph nodes found on pathologic examination. If no positive nodes are found, then the patient is considered to be at a sufficiently low risk of recurrence for no adjuvant treatment to be called for (although 25% of these patients will not survive 10 years). If four or more positive nodes are discovered, the patient is considered to be at high risk (although 25% of these patients would survive at least 10 years in the absence of any further treatment). Given four or more positive nodes, the best practice now suggests adjuvant chemotherapy for most patients. Decisions about patients with limited nodal involvement (defined as one to three nodes) pose more difficult problems, because at least 40% will fail within 10 years if they receive no further treatment. On the one hand, if the micrometastasis hypothesis is correct, adjuvant chemotherapy should have a particularly favorable effect in this group. On the other hand, adjuvant therapy is of itself hazardous, expensive, and time-consuming. Thus it is clinically useful to identify further predictive variables that can sharpen the prediction of the risk of recurrence for individual patients so that a better selection of patients at high risk can be made.

In this pilot study a computer-based statistical analysis has been applied to a data base derived from a personal series of patients with one to three positive nodes. This analysis has established the feasibility of improving the prediction of disease recurrence and thus improving the criteria for the selection of patients for adjuvant chemotherapy. In its simplest form the analysis correctly identified 80% of both high- and low-risk patients. The pilot study further indicates that the number of positive lymph nodes discovered should be considered as merely one of the set of predictive variables to be used in the calculation of risk of recurrence.

If an explicit analysis of the risk of recurrence were available for clinical decisions the needs and preferences of the individual patient could be matched more precisely to the spectrum of available treatment. In the long term, the present set of data is expected to be augmented by new measures of hormonal, biochemical, proliferative, and immunologic status. These new measures will result in an even more detailed and accurate prognostic profile for quantification of the natural history of breast carcinoma under treatment. On the basis of these preliminary results, the present approach provides quantitative support to that summary of personal or institutional experience usually prefaced by the phrase 'In our hands . . .'

Acknowledgements: This work was supported in part by U.S. Public Health Service Grants CA-12369, CA-19663, and CA-14088 from the National Institutes of Health-Department of Health, Education, and Welfare, The Helena Rubinstein Foundation, The Rubin Foundation, The Samuel and Anna Jacobs Foundation, Inc., The Shelter Rock Foundation and the Phyllis Rubinstein Special Gift Fund.

Appendix A
Patient data record

1. Code No.: _ _ _ _ _ _ _; 2. Birthdate: _ _/ _ _/_ _; 3. Age: _ _; 4. Birthplace: 1. USA, 2. Canada, 3. Mexico, 4. Russia, 5. Polland, 6. Germany, 7. Czechoslavakia, 8. Italy, 9. Spain, 10. Israel, 11. Austria, 12. Other; 5. Ethnicity: 1. Caucasian, 2. Black, 3. Mongolian, 4. Other; 6. Religion: 1. Jewish, 2. Protestant, 3. Catholic, 4. None, 5. Moslem, 6. Hindu, 7. Syrian-Orthodox, 8. Other; 7. Marital Status: 1. Married, 2. Widowed, 3. Divorced, 4. Separated, 5. Single; 8. Age at Marriage: _ _, Date: _ _/_ _/ _ _; 9. Gravida, Para, Aborta: _ _/_ _/_ _; 10. Age at 1st Birth: _ _; 11. Age at Menopause: _ _; 12. Age of Children: (oldest/youngest) _ _/_ _; 13. Average months nursed per child: _ _; 14. Total months nursed: _ _; 15. Estrogens: 0.Never, 1. Less than 1 year total use, greater than 10 years PTA, 2. Greater than 1 year total, greater than 10 years PTA, 3. Less than 1 year total, from 1 to 10 years PTA, 4. Greater than 1 year total, from 1 to 10 years PTA, 5. Less than 1 year total, less than 1 year PTA, 6. Greater than 1 year total, less than 1 year PTA; 16. Progesterones: 0. Never, 1. Less than 1 year total use, greater than 10 years PTA, 2. Greater than 1 year total use, greater than 10 years PTA, 3. Less than 1 year total, from 1 to 10 years PTA, 4. Greater than 1 year total, from 1 to 10 years PTA,

5. Less than 1 year total, less than 1 year PTA, 6. Greater than 1 year total, less than 1 year PTA; 17. Ethanol: 0. None, 1. Occasional, Social, 2. Regularly, 3. Excessive; 18. Other Drugs: 0. None, 1. Endocrine, 2. Others, 3. 1 and 2; 19. Endocrine Surgery: 0. None, Date, M.D., Hospital, 1. Oopherectomy, Date, M.D., Hospital, 2. Thyroidectomy, Date, M.D., Hospital, 3. Adrenalectomy, Date, M.D., Hospital, 4. Other, Date, M.D., Hospital, 5. Any combination of 1 to 4, Date, M.D., Hospital; 20. Past Breast Disease: 0. None, Date, M.D., Hospital, 1. Fibroadenoma, Date, M.D., Hospital, 2. Cystic Disease, Date, M.D., Hospital, 3. Fibrosis, Date, M.D., Hospital, 4. Cancer, Date, M.D., Hospital, 5. Mammary Duct Ectasia, Date, M.D., Hospital, 6. Other, Date, M.D., Hospital, 7. Any combination of 1 to 6, Date, M.D., Hospital; 21. Past Other Illnesses: 0. None, 1. Endocrinopathy, 2. Others; 22. Family History of Breast Cancer: None, Age, 1. Mother, Age, 2. Maternal Grandmother, Age, 3. Paternal Grandmother, Age, 4. Maternal Aunt, Age, 5. Paternal Aunt, Age, 6. Sister, Age, 7. Other, Age; 23. Discovered by: 1. Self, 2. GYN or LMD (primary care physician), 3. CDH or SJK (follow-up), 4. Routine Screening Mammogram, 5. Other; 24. Interval from Discovery to Biopsy: __ days; 25. Initial Mode of Presentation: 1. Mass or lump in FMG, 2. Pain in FMG, 3. Heavy, dragging or other sensation, 4. Nipple discharge, 5. Skin or nipple changes, 6. Palpable axillary lymph node, 7. Abnormal mammogram, 8. Other; 26. Side: __; 27. Location __; 28. Size: __ cm in greatest diameter; 29. Configuration: 1. Poorly delineated and/or irregular, 2. Discrete (including nodular, lobulated, round, etc.), 3. Others, 4. Not remarked upon: 30. Consistency: 0. Not remarked upon, 1. Firm, hard, or gritty, 2. Soft, 3. Rubbery, 4. Other; 31. Relationship to Environs: 1. Freely moveable, no retraction or fixation, 2. Skin retraction or nipple retraction, 3. Fixed in surrounding breast substance, 4. Fixed to underlying chest wall tissue, 5. Peau d'orange, 6. Other, 7. Any combination of 1 to 6; 32. Skin Changes: 0. None, 1. Ecchymosis, 2. Edema, limited to less than 1/3 of breast surface, 3. Edema, extensive, 4. Peau d' orange, 5. Erythema, 6. Other, 7. Any combination of 1 to 6; 33. Nipple Changes: 0. None, 1. Retraction or inversion (only if other side normal), 2. Discharge (bloody or otherwise), 3. Any combination of 1 to 3; 34. Characteristics of Involved Breast: 0. Non-tender, 1. Tender or painful, 2. Different size from opposite, 3. Pendulous (with or without), 4. Other, 5. Any combination of 1 to 4, 6. Without distinguishing characteristics; 35. Homolateral Lymph Nodes: AX, SC, IM, 0. None, 1. Less than 2.5 cm and freely moveable, 2. Greater than 2.5 cm and freely moveable, 3. Less than 2.5 cm and matted down, 4. Greater than 2.5 cm and matted down; 36. Contralateral Lymph Nodes: AX, SC, IM, 0. None, 1. Less than 2.5 cm and freely moveable, 2. Greater than 2.5 cm and freely moveable, 3. Less than 2.5 cm and matted down, 4. Greater than 2.5 cm and matted down; 37. Neural: 0. No, 1. Yes; 38. Fascial: 0. No, 1. Yes (relative fixation to chest wall); 39. Muscle: 0. No, 1. Yes; 40. Columbia Clinical Classification: __; 41. International TNM System: __; 42. Duration between biopsy of primary tumor and surgical treatment: __ days/months or years; 43. Procedure: 1. Haagensen radical mastectomy, 2. Modified radical mastectomy, 3. Simple mastectomy, 4. Lumpectomy (wide excision); 44. Area of skin removed: __ x __ cm (use greatest dimension); 45. Skin graft size: __ x __ cm; 46. Duration of operation: (hours/minutes __/ __; 47. Transfusion: 0. None, 1. Whole blood __ units, 2. Plasma __ units, 3. Other; 48. Post-operative complications: 0. None, 1. Fluid collection under flaps, etc., 2. Urinary tract infection, 3. Pulmonary tract infection, 4. Thrombophlebitis, 5. Pulmonary Embolus, 6. Cardia or cardio-vascular problem, 7. Depression, 8. Other, 9 Any combination of 1 to 8; 49. Post-operative transfusions: 0. None, 1. Blood, 2. Plasma, 3. Other, 4. Any combination of 1 to 3; 50. Duration of post-op hospital stay: __ days; 51. Radiotherapy: 0. None, 1. Post-op, to IMN chain, 2. Post-op, to axilla and SC area or "regional nodes," 3. Post-op, to chest wall, 4. For recurrence (later) to any area, 5. Post-op, to combination of 1 to 3, 6. Other, 7. For combination of 4 and any other; 52. Chemotherapy: 0. None, 1. Adjuvant, 2. For recurrent disease (later), 3. 1 and 2; 53. Hormonal Manipulation: (if any, give 1 or 2, and 3 to 10), 0. None, 1. Adjuvant (immediately post-op), 2. For recurrent disease (later), 3. Oopherectomy, 4. Adrenalectomy, 5. Oopherectomy and adrenalectomy. 6. Other surgery, 7. Steroids, 8. Androgen analogue, 9. Estrogen or progestogen analogues, 10. Any combination of 3-6 and 7-9; 54. Consistency: 0. Not remarked upon, 1. Firm, hard, or gritty (ignore color and other descriptive data), 2. Soft, 3. Rubbery, 4. Other; 55. Size: __ cm; 56. Margin from tumor to skin: 0. None, skin not involved, 1. None, skin involved, __ cm; 57. Margin from tumor to fascia: 0. None, fascia and muscle not involved, 1. None, fascia or muscle involved, 2. Less than 1 cm, __ cm; 58. Histologic Type: 1. No special type (includes schirrous, invasive, intraductal), 2. Predominantly intraductal with foci of invasion, 3. Intraductal, 4. Papillary, 5. Solid circumscribed or medullary, 6. Small cell, 7. Pagets, 8. Inflammatory, 9. Lobular carcinoma (invasive), 10. Multiple types, with no one predominant, 11. Other; 59. Degree of Differentiation: 0. No special differentiation, 1. Well-differentiated, 2. Moderately to well-differentiated, 3. Moderately differentiated, 4. Moderate to poorly-differentiated, 5. Poorly differentiated; 60. Invasion of blood vessels: 0. No, 1. Yes; 61. Invasion of lymphatics in breast: 0. No, 1. Yes; 62. Invasion of breast stroma: 0. No, 1. Yes; 63. Invasion of axillary fat: 0. No, 1. Yes; 64. Total no. of ALN recovered: __; 65. No. of ALN in each subgroup: __ AV, __ C, __ S, __ R, __ Ap; 66. No. of positive ALN in total: __; 67. Percent of total ALN that are positive: __; 68. No. of positive ALN in each subgroup: __ AV, __ C, __ S, __ R, __ Ap; 69. Percent of total positive ALN that are: __ AV, __ C, __ S, __ R, __ AP; 70. Percent of ALN in each subgroup that are positive: __ AV, __ C, __ S, __ R, __ Ap; 71. Level of visibility of

metastatic foci: 1. Macrometastasis, 2. Micrometastasis; 72. Areas of nodes involved: 1. Capsule and sub-capsular sinus only (micro), 2. Deeper areas of node involved discretely (macro), 3. Node diffusely replaced with tumor (macro); 73. Associated pathologic findings: 0. None, 1. Microscopic cystic disease (including any or all of the following: microscopic cysts, squamous metaplasia, proliferation of duct epithelium, blunt duct adenosis, adenosis, fibrosis). 2. Gross cystic disease, 3. Fibroadenoma (gross or micro), 4. Intraductal papilloma, 5. Papillomatosis, 6. Lobular neoplasia in situ, 7. Chronic lymphadenitis or sinus histiocytosis, 8. Lesions of breast cancer, 9. Other, 10. Any combination of 1 to 9; 74. Date of operation: __/ __/ __; 75. Date of metastasis or recurrence: __/ __ (0 = none); 76. Site of metastasis: 1. Local recurrence (chest wall or skin nodules), 2. Local recurrence (axilla), 3. Regional spread (internal mammary or supraclavicular nodes), 4. Bony metastases, 5. Pulmonary metastases, 6. Hepatic metastases, 7. Brain metastases, 8. Other, 9. Any combination of 1 to 8; 77. Date of death: __/ __/ __ (0 = none); 78. Cause of death: 1. Metastatic disease, 2. Unrelated disease, 3. Other.

Abbreviations

ALN: Axillary Lymph Nodes; AV: Avillary Vein; C: Central: S: Scapular; R: Rotter's; Ap: Apex

Appendix B1

Analytical Methods

Let x be a vector of observed patient characteristics. For the i'th patient the vector is x_i and has components 1, x_{1i}, x_{2i}, ... x_{ki}. Following earlier studies (Fiegl and Zelen, 1965; Glasser, 1967; Prentice, 1973), the hazard rate λ_i is assumed to depend log-linearly on x, and is independent of elapsed times since surgery:

$$-\ln\lambda_i = \beta_0 + \beta_1 x_{1i} + \cdots \beta_k x_{ki} . \quad (1)$$

The expected time to metastases is $1/\lambda_i$, and the median time for the i'th patient is $0.693/\lambda_i$. Since λ_i has units of inverse months, the probability of a 5-year disease-free interval is P5YR = exp $(-60\lambda_i)$. Maximum-likelihood methods were used to determine the parameter vector β by fitting to the data.

In vector notation, $-\ln \lambda_i = \beta' x_i$. The likelihood for those individuals censored at time t_i^* is

$$L_i = \exp\left[-(t_i^* e^{-\beta' x_i})\right] \quad (2)$$

and the likelihood for those individuals observed to recur at time t_i is

$$L_i = \exp\left(\beta' x_i - t_i e^{-\beta' x_i}\right) . \quad (3)$$

When a data entry for a particular x_{ki} was not available, the average value of the k'th variable over all available cases was used to replace the missing value. This procedure is satisfactory in the absence of any additional information about missing data. In one instance, described in the text, this procedure required modification.

The parameter vector β is chosen, which maximizes the likelihood, $\prod_{i=1}^{n} L_i$, where n is the number of patients.

The first two derivatives of ln L are used in the determination of β. They are

$$\frac{\partial \ln L_i}{\partial \beta} = \begin{cases} (t_i^* e^{-\beta' x_i}) x_i' , & \text{if censored} \\ (1 + t_i e^{-\beta' x_i}) , & \text{if uncensored} \end{cases} \quad (4)$$

and

$$\frac{\partial^2 \ln L_i}{\partial \beta^2} = \begin{cases} -(t_i^* e^{-\beta' x_i}) x_i x_i' , & \text{if censored} \\ -(t_i e^{-\beta' x_i}) x_i x_i' , & \text{if uncensored.} \end{cases} \quad (5)$$

It has been observed (Morabito and Marubina, 1976; Walker and Duncan, 1967), that the Newton-Raphson iteration technique for determining a maximum corresponds to the solution of a weighted least-squares problem with weights

$$W_i = \begin{cases} t_i^* e^{-\beta' x_i} , & \text{if censored} \\ t_i e^{-\beta' x_i} , & \text{if uncensored.} \end{cases} \quad (6)$$

With these weights, the dependent variable becomes

$$\frac{\partial \ln L_i}{\partial \beta} \Big/ W_i .$$

It was natural to incorporate these algorithms into a general-purpose statistical software program capable of performing weighted least-squares regression. The STATLIB Program (Bradsford and Relles, 1975) was chosen because of its superior data manipulation facility and efficient cross-product matrix accumulator. The starting point in the iteration sequence was the least-squares solution with ln t as the dependent variable, ignoring censored observations. If this is nonsingular, a finite solution to the censored maximum likelihood is guaranteed. Since the weights in this problem are always positive, it is also assured that the second derivative matrix is negative definite, implying the unique existence of a local maximum.

Appendix B2

Confidence Statements

Based on the theory of maximum likelihood estimators (Wilks, 1962), uncertainty in the parameter vector β,

can be approximated as a subjective normal probability distribution

$$(\hat{\beta} - \beta) \sim \eta\left(0, \frac{-\partial^2 \ln (\Pi L_i)}{\partial \beta^2}\right).$$

For any particular individual with characteristics x_i, the important variable $\beta' x_i$ has variance

$$x_i' \left[\frac{-\partial^2 \ln (\Pi L_i)}{\partial \beta^2}\right] x_i .$$

A normal distribution was used to make confidence statements about $\beta' x_i$ and these were then expressed in terms of P5YR.

The theory for confidence statements is satisfactory when none of the β's is estimated to be infinite. As described in the text, a set of three prognostic variables was identified such that no patient with any of these characteristics was observed to recur. For the 39 patients exhibiting these characteristics, estimates of β are infinite. It can be shown that if the three variables associated with this excellent prognosis actually correspond to zero hazard rate for recurrence, then the maximum likelihood estimator of the hazard including all patients will be the same as the partial maximum likelihood estimator excluding those patients with excellent prognosis. Probability statements and statistical distributions for both cases are then identical. Of the two, the partial maximum likelihood estimator was employed because of its computational simplicity. Insofar as the three variables are predictive of no metastasis, the normal approximation yields both reliable and easily computable confidence statements.

An alternative and theoretically attractive approach is to compute confidence points from the likelihood itself (Cox, 1970). Let $L_\beta = \sum_i ln\, L_i$, with β the true parameter vector. Suppose $f(\beta)$ is a p-dimensional function describing a restriction on possible values of β. (In this case, $f(\beta) = \beta' x$). Generalized likelihood ratio test theory states that

$$2\left(\operatorname{Max}_{\beta} L_\beta - \operatorname{Max}_{\beta : f(\beta) = c} L_\beta\right) = \chi_p^2 .$$

This implies that the lower alpha confidence limit of $\beta' x$ can be obtained by computing the minimum value of $\beta' x$ over possible β

$$\text{subject to } \left(\operatorname{Max}_{\beta^*} L_{\beta^*} - L_\beta = {}^1/_2\, \chi^2_{1,\alpha}\right)$$

where $\chi^2_{i,\alpha}$ is the upper alpha percent point of the χ^2_i distribution. This is feasible but must be repeated for each vector x of required explanatory variables. This exhaustive procedure was not performed, but a close estimate was obtained by a crude search among plausible values of β. This computation of the lower confidence limit is justified by the hypothesis testing interpretation. The value of $\beta' x$ determined by the minimum will not reject the given data at level α, but any value of β leading to a lower value at $\beta' x$ will cause rejection. A similar argument applies to the upper confidence interval.

Appendix C

Validity of the Exponential Model

The validity of the constant hazard rate model was determined in two ways:

1) A Kaplan-Meier estimate of the probability of recurrence-free interval was made with the nondimensional time, $\lambda_i t_i$, where t_i is the observed time to recurrence or censorship and λ_i is the predicted hazard rate, all for the i'th patient. If the assumed model were indeed the correct one, then the resulting estimate of the probability distribution would approximate an exponential distribution $\exp(-\lambda_i t_i)$. This was verified for the region $0 < \lambda_i t_i < 1$, where the data are heavily concentrated. Significant departures from this distribution occur only for values of $\lambda_i t_i > 1$, where data are too sparse ($N = 3$) and error estimates too large to permit meaningful comparisons.

2) The algorithm and computer programs were modified to include hazard rates of the Weibull type, where the hazard rate is proportional to $t\left(\frac{1-\sigma}{\sigma}\right)$ and may either increase ($\sigma < 1$), or decrease ($\sigma > 1$) with time since mastectomy. The exponential relation for the probability of disease-free interval patient was replaced by

$$\exp\left\{-\exp\left(\frac{\ln t - \beta' x}{\sigma}\right)\right\}.$$

The parameter σ was assumed constant for the entire patient set, and a maximum likelihood calculation was performed to determine σ and the coefficient vector β simultaneously by fitting to the data.

A maximum likelihood was obtained when $\sigma = 0.69$, suggesting a hazard function that increased weakly with time. The relative weights of the individual variables were only slightly different from those in the constant hazard case ($\sigma = 1$). In addition, the Weibull estimates for 5-year disease-free interval were usually within 2% of those calculated with the simpler exponential model. Although σ is significantly different from unity, it has little influence on estimates of short-term risk status. For these reasons, and also because inclusion of the Weibull hazard rate adds considerable complexity to description and formulae, we present only the exponential results. Both clinical intuition and a recent study (Mueller and

Jeffries, 1975) of aggregated breast cancer survival data suggest that a value of $\sigma > 1$ is also plausible.

As suggested earlier, the Weibull σ obtains different values when computed on the basis of aggregated data, or on the basis of individual patient characteristics. To illustrate this effect, a maximum likelihood calculation was performed with pooled data from all 103 patients, i.e., neglecting confirmed individual differences. The computed σ was found to be insignificantly different from unity, contrasting with the value of 0.69 described above. Given a larger data base or a longer period of follow-up, the present methodology can easily be used to explore the nature of this time dependence for the individual patient. The issue is still an important one and deserves future study.

References

Bonadonna, G., Rossi, A., Valagussa, P., Banfi, A., Veronesi, U.: The CMF program for operable breast cancer with positive axillary nodes. Cancer **39,** 2904 (1977)

Bradsford, W., Relles, D.: Interactive statistical computing, with applications to forecasting and data analysis. In: Proceedings of computer science and statistics: Eight Annual Symposium on the Interface, p. 530. Los Angeles: Western Periodicals 1975

Byar, D., Huse, R., Bailar, J.: Exponential model relating censored survival data and concomitant information for prostatic cancer patients. J. Natl. Cancer Inst. **52,** 321 (1974)

Coombes, R., Powles, T., Gazet, J., Ford, H., Nash, A., Sloan, J., Hillyard, C., Thomas, P., Keyser, J., Marcus, D., Zinberg, N., Stimson, W., Munro, Neville, A.: A biochemical approach to the staging of human breast cancer. Cancer **40,** 937 (1977)

Cox, D.: The analysis of binary data, p. 88. London: Methuen 1970

Fiegl, P., Zelen, M.: Estimation of exponential survival probabilities with concomitant information. Biometrics **21,** 826 (1965)

Fielding, L., Stewart-Brown, S., Dudley, H.: Surgeon-related variables and the clinical trial. Lancet **1978,** 778

Fisher, B., Slack, N., Katrych, D., Wolmark, N.: Ten-year follow-up results of patients with carcinoma of the breast in a cooperative clinical trial evaluating surgical adjuvant chemotherapy. Surg. Gynecol. Obstet. **140,** 528 (1975)

Fisher, B., Glass, A., Redmond, C., Fisher, E., Barton, B., Such, E., Carbone, P., Economou, S., Foster, R., Frelick, R.: L-Phenylalanine mustard (L-PAM) in the management of primary breast cancer. Cancer **39,** 2883 (1977)

Fisher, E., Gregorio, R., Fisher, B., with the assistance of Redmond, C., Vellios, F., Sommers, S., and cooperating investigators: The pathology of invasive breast cancer, a syllabus derived from findings of the national surgical adjuvant breast project (Protocol No. 4). Cancer **36,** 1 (1975)

Glasser, M.: Exponential survival with covariance. J. Am. Stat. Assoc. **62,** 561 (1967)

Groner, G., Hopwood, M., Palley, N., Sibley, W., Baker, W., Christopher, T., Thompson, H.: An interactive data management and analysis system for clinical investigators. J. Lab. Clin. Med. **92,** 325 (1978)

Haagensen, C. D.: Diseases of the breast, Rev. 2nd ed. Philadelphia: Saunders 1971

Haagensen, C.: Ten-year results with radical mastectomy. Surgery **76,** 685 (1974)

McGuire, W., Carbone, P., Vollmer, E. (Eds.): Estrogen receptors in human breast cancer. New York: Raven 1975

Morabito, A., Marubina, E.: A computer program for fitting linear models when the dependent variable is dichotomous, polichotomous, or censored survival and non-linear models when the dependent variable is quantitative. Comput. Programs Biomed. **5,** 283 (1976)

Mueller, C., Jeffries, W.: Cancer of the breast: Its outcome as measured by the rate of dying and cause of death. Ann. Surg. **182,** 334 (1975)

Pickren, J.: Lymph node metastases in carcinoma of female mammary gland. Roswell Park Memorial Institute Bulletin **1,** 79 (1956)

Pickren, J.: Significant of occult metastases. A study of breast cancer. Cancer **14,** 1266 (1961)

Prentice, R.: Exponential survivals with censoring and explanatory variable. Biometrika **60,** 279 (1973)

Rosencweig, M., Heuson, J.: Breast cancer: Prognostic factors and clinical evaluation. In: Cancer therapy: Prognostic factors and criteria of response. M. J. Staquet (ed.). New York: Raven 1975

Sklarew, R., Hoffman, J., Post, J.: A rapid in vitro method for measuring cell proliferation in human breast cancer. Cancer **40,** 2299 (1977)

Walker, S., Duncan, D.: Estimation of the probability of an event as a function of several independent variables. Biometrika **54,** 167 (1967)

Wilks, S.: Mathematical statistics. New York: Wiley 1962

Received May 22, 1978/Accepted September 28, 1978