

# Receiver operating characteristic (ROC) to determine cut-off points of biomarkers in lung cancer patients

Heidi L. Weiss<sup>a,\*</sup>, Santosh Niwas<sup>b</sup>, William E. Grizzle<sup>c</sup> and Chandrika Piyathilake<sup>d</sup>

<sup>a</sup>*Department of Medicine, Baylor College of Medicine, Houston, TX, USA*

<sup>b</sup>*Medical Statistics Section, Department of Medicine, The University of Alabama at Birmingham, Birmingham, AL, USA*

<sup>c</sup>*Department of Pathology, The University of Alabama at Birmingham, Birmingham, AL, USA*

<sup>d</sup>*Department of Nutrition Sciences, The University of Alabama at Birmingham, Birmingham, AL, USA*

**Abstract.** The role of biomarkers in disease prognosis continues to be an important investigation in many cancer studies. In order for these biomarkers to have practical application in clinical decision making regarding patient treatment and follow-up, it is common to dichotomize patients into those with low vs. high expression levels. In this study, receiver operating characteristic (ROC) curves, area under the curve (AUC) of the ROC, sensitivity, specificity, as well as likelihood ratios were calculated to determine levels of growth factor biomarkers that best differentiate lung cancer cases versus control subjects. Selected cut-off points for p185<sup>erbB-2</sup> and EGFR membrane appear to have good discriminating power to differentiate control tissues versus uninvolved tissues from patients with lung cancer (AUC = 89% and 90%, respectively); while AUC increased to at least 90% for selected cut-off points for p185<sup>erbB-2</sup> membrane, EGFR membrane, and FASE when comparing between control versus carcinoma tissues from lung cancer cases. Using data from control subjects compared to patients with lung cancer, we presented a simple and intuitive approach to determine dichotomized levels of biomarkers and validated the value of these biomarkers as surrogate endpoints for cancer outcome.

## 1. Introduction

The role of biomarkers in disease prognosis continues to be an important investigation in many cancer studies. More specifically, the central role of growth factors and their receptors in the neoplastic process makes evaluation of these biomarkers pivotal in the early detection, diagnosis and development of targeted therapies [1]. Biomarkers have also been used as surrogate endpoints for cancer outcome. For example, many Phase II chemoprevention studies in patients at high-risk for cancer utilize biomarkers as surrogate endpoints

of efficacy [2]. This allows for smaller, less expensive, and faster studies as compared to studies that utilize cancer as the endpoint [3].

Many studies have identified EGF-r<sup>3</sup> and two of its ligands, EGF and TGF- $\alpha$ , and a closely related receptor, p185<sup>erbB-2</sup>, as some of the major growth factors and growth factor receptors involved in neoplastic formations [4]. The role of these biomarkers in lung cancer has been studied by many investigators [5–12]. However, a great heterogeneity in the incidence of overexpression and in the prognostic significance of these growth factors and receptors has been observed both within and between various tumor types, including cancers of the lung.

In order for these biomarkers to have practical application in clinical decision making regarding patient treatment and follow-up, it is common to dichotomize

---

\*Corresponding author. Department of Medicine, Baylor College of Medicine, Room 335A, One Baylor Plaza, MS: BCM 600, Houston, TX, 77030, USA. Tel.: +1 713 798 1630; Fax: +1 713 798 1642; E-mail: hweiss@breastcenter.tmc.edu.

patients into those with low vs. high expression levels [13]. Receiver operating characteristic (ROC) analysis is a widely-used method for comparing diagnostic accuracy of laboratory tests. The ROC curve is constructed by plotting the true positive rate (sensitivity) on the y-axis as a function of false positive rate (1-specificity) on the x-axis. An ideal ROC curve climbs rapidly towards the upper left corner of the graph which indicates good discrimination property. Pina et al. [13] showed the use of ROC in obtaining optimum cut-off points for serum tumor markers in lung cancer patients. They also calculated the area under the ROC curve as a quantitative measure of the discrimination power of tumor markers between lung cancer cases vs. healthy controls or subjects with non-neoplastic-type pulmonary pathologies. Furthermore, Choi has discussed how the slope of the ROC can be used to calculate different types of likelihood ratios [14].

This paper utilizes a case-control study of 60 lung cancer patients and 60 control patients (without lung cancer) to determine levels of biomarkers of growth factors (specifically EGFR, TGF- $\alpha$ , p185<sup>erbB-2</sup>) and fatty acid synthase (FASE), a major component of the de novo fatty acid biosynthetic pathway, that will best differentiate between the two groups. These biomarkers were measured on a semi-quantitative scale using immunohistochemistry techniques [15]. Receiver operating characteristic (ROC) curves, area under the curve (AUC) of the ROC as well as likelihood ratios were calculated to determine levels of biomarkers that best differentiate lung cancer cases and control subjects.

## 2. Methods

### 2.1. Study population

Study subjects were selected from patient populations of the Veterans Administration Hospital and University Hospital in Birmingham, AL. Sixty non-cancer specimens of the lung (controls) were selected at random from a list of lung surgeries performed between 1988–1998 for lung diseases or conditions other than lung cancer. From a patient list of SCC of the lung who had undergone surgery between 1989–1996, sixty cases were selected at random. Formalin fixed, paraffin embedded sections of the lung were obtained from the archival collection. One tissue block from each non-cancer patient was selected to provide a section that contained adequate normal specimen. Tissue blocks from each SCC patient were se-

lected to provide sections that contained uninvolved and carcinoma specimens. Immunohistochemistry scores were evaluated for 4 biomarkers (EGFR, p185<sup>erbB-2</sup>, TGF- $\alpha$ , FASE). Methods of immunohistochemical analysis have been reported previously [15]. Briefly, biomarkers were graded on the intensity of immunostaining (0–4) as well as the percentage of cells at each intensity. The sum of the product of the intensity of immunostaining and percent cells at each intensity were calculated to arrive at a semi-quantitative immunohistochemistry score.

### 2.2. Analysis methods

Descriptive statistics were calculated to summarize immunohistochemistry scores from normal tissues of control subjects and from uninvolved and carcinoma tissues in lung cancer cases. Univariate comparison of scores between lung cancer cases versus controls were performed using Wilcoxon rank-sum tests while within patient comparison of scores were performed using Wilcoxon signed-ranks test.

Continuous levels of immunohistochemistry scores for each biomarker were grouped into multiple categories based on the distribution of biomarker scores. For each biomarker, a logistic regression model was fitted for cases versus controls using categorized levels of the biomarker score as independent variable. ROC curves based on the logistic regression model were constructed. Sensitivity, specificity, likelihood ratios (LR), and AUC of the ROC based on the trapezoidal rule were calculated. Values of AUC closer to 100% indicate better discrimination power of a biomarker between cases versus controls. The slope of the ROC curve represents a likelihood ratio for a diagnostic test [14,16]. The multiple LR [14] was calculated and presented in this paper. Optimal cut-off points for biomarkers were determined based on the best balance of sensitivity and specificity along with larger increases in LR's. Logistic regression models and the corresponding ROC curves were generated for the comparison of normal specimens in controls versus uninvolved specimens from lung cancer cases and for the comparison of normal specimens in controls versus carcinoma specimens from lung cancer cases.

## 3. Results

Table 1 shows mean  $\pm$  standard deviation, median, and range of immunohistochemistry scores in lung can-

cer cases and controls by cellular localization (i.e. basal, luminal, cytoplasm, membrane). Across all biomarkers, lung cancer cases exhibited higher scores compared to control subjects. For example, the mean level of basal p185<sup>erbB-2</sup> score was 0.51 for controls compared to 0.86 from uninvolved tissues of lung cancer cases. Furthermore, an increasing trend of scores from normal tissues of controls to carcinoma tissues of lung cancer cases was evident. For example, mean EGFR cytoplasm scores were equal to 0.38, 0.74, and 1.32 for controls, uninvolved tissues, and carcinoma tissues from lung cancer cases, respectively. A similar trend of increasing mean EGFR scores was observed in the membrane layer. The average of scores from the cytoplasm and membrane layers are also presented. Statistically significant differences in these scores between controls vs. uninvolved tissues from lung cancer cases were observed for all biomarkers in all locations ( $p < 0.0001$ ). These differences also were observed for all biomarkers in all locations when comparing controls vs. carcinoma tissues from lung cancer cases ( $p < 0.005$ ).

Table 2 shows biomarker cut-off points when comparing control versus uninvolved tissues from lung cancer cases. The optimal cut-off points were selected based on levels of each biomarker that provide the best balance of sensitivity, specificity as well as larger increases in LR's. Levels of p185<sup>erbB-2</sup> scores that provide the best discrimination point between lung cancer cases versus controls were equal to 0.70, 1.30, and 0.40 for basal, cytoplasm, and membrane, respectively. These cut-off points provided moderate levels of sensitivity for basal and cytoplasm (60% and 65%, respectively) and a higher sensitivity for p185<sup>erbB-2</sup> in membrane (80%). Correspondingly, the highest AUC was achieved at 89% for p185<sup>erbB-2</sup> membrane. Thus, it appears that p185<sup>erbB-2</sup> membrane levels provide better discrimination power between lung cancer cases and controls. Specificity levels for these cut-off points were at least 85%; while the LR's indicate that p185<sup>erbB-2</sup> levels greater than or equal to the cut-off points were at least 2.0 times as likely to be found in lung cancer cases than in controls. EGFR cut-off points were equal to 1.00, 0.40 and 0.10 for basal, cytoplasm, and membrane, respectively. Moderate levels of sensitivity for these cut-off points were observed for both basal and cytoplasm while sensitivity was equal to 91% for EGFR membrane with a corresponding AUC equal to 90%. Similar to p185<sup>erbB-2</sup>, the best discrimination power was observed for levels of EGFR measured from the membrane layer. Cut-off points for TGF- $\alpha$  and FASE indicate only moderate but slightly better lev-

els of discrimination from measurements in the basal layer (AUC=86% and 78%, respectively) as opposed to carcinoma layer.

Table 3 shows biomarker cut-off points based on comparison between controls versus carcinoma tissues of lung cancer cases. The cut-off point equal to 0.40 for p185<sup>erbB-2</sup> membrane provided 89% sensitivity and 92% specificity with an AUC equal to 93%. A similar cut-off point of 0.40 was obtained for EGFR membrane providing high levels of sensitivity and specificity with an AUC equal to 97%. Interestingly, a cut-off point for FASE equal to 0.70 provided strong discrimination power (AUC = 99.6%) between controls vs. lung cancer cases.

Finally, attributable proportions (AP) were calculated which provide a measure of cases that is attributable to a biomarker [17] (Table 4). Values of AP closer to 1.0 indicate a more valid biomarker as a surrogate endpoint for cancer outcome. Using the optimal cut-off points from Tables 2 and 3, the corresponding AP for each biomarker were calculated (Table 4). As shown in Table 4, all biomarkers exhibited only small to moderate levels of AP. EGFR levels from the membrane exhibited the highest AP equal to 79% while only 63% of lung cancer cases is attributable to high levels of p185<sup>erbB-2</sup> from the membrane. As indicated by Schatzkin et al. [17], individual biomarkers are at best marginal surrogates for cancer outcome. Higher levels of AP were obtained from cut-off points arising from the comparison between controls vs. carcinoma tissues from lung cancer cases. Attributable proportions were equal to 83% and 78% for EGFR membrane and p185<sup>erbB-2</sup> membrane, respectively. In addition to providing strong discrimination power between controls versus tumor tissues from lung cancer patients (Table 3), FASE appears to be a good surrogate for lung cancer outcome providing an AP equal to 96%. A study by Piyathilake et al. [18] suggests that expression of FASE is an early event in the development and progression of SCC of the lung. Combinations of biomarkers also were evaluated in a multiple logistic regression model (data not shown). The combination of p185<sup>erbB-2</sup> and EGFR membrane levels improved AUC to 96% from an AUC of about 90% with either biomarker alone (Table 2) when comparing controls versus uninvolved tissues from lung cancer cases. The combination of FASE and EGFR membrane levels did not improve the AUC of 99.6% from the model with FASE alone when comparing controls vs. carcinoma tissues from lung cancer cases.

Table 1  
Descriptive Statistics for Biomarkers in Lung Cancer Cases versus Controls

Biomarker	Location	Normal Tissues from Controls Mean $\pm$ SD (median), range	Uninvolved Tissues from Lung Cancer Cases Mean $\pm$ SD (median), range	Carcinoma Tissues from Lung Cancer Cases Mean $\pm$ SD (median), range	
p185 <sup>erbB-2</sup>	Basal	0.51 $\pm$ 0.04 (0.50), 0–1.3	0.86 $\pm$ 0.04 (0.80), 0.2–1.6	–	
	Luminal, Cytoplasm	0.99 $\pm$ 0.05 (1.05), 0.08–1.6	1.53 $\pm$ 0.07 (1.50), 0.5–2.8	1.27 <sup>a</sup> $\pm$ 0.44 (1.30), 0.30–2.3	
	Luminal, Membrane	0.23 $\pm$ 0.03 (0.20), 0–1.0	0.68 $\pm$ 0.04 (0.60), 0.1–1.5	0.97 <sup>b</sup> $\pm$ 0.55 (0.80), 0.10–2.3	
	Average (C+M)	0.61 $\pm$ 0.04 (0.65), 0.04–1.30	1.11 $\pm$ 0.05 (1.05), 0.30–2.0	1.12 $\pm$ 0.44 (1.05), 0.20–2.10	
	EGFR	Basal	0.70 $\pm$ 0.07 (0.70), 0–1.8	1.25 $\pm$ 0.08 (1.30), 0.2–2.5	–
EGFR	Luminal, Cytoplasm	0.38 $\pm$ 0.04 (0.30), 0–1.1	0.74 $\pm$ 0.05 (0.70), 0.2–1.6	1.32 <sup>a</sup> $\pm$ 0.52 (1.50), 0.20–2.10	
	Luminal, Membrane	0.09 $\pm$ 0.02 (0.20), 0–0.7	0.55 $\pm$ 0.04 (0.50), 0.01–1.4	1.68 <sup>b</sup> $\pm$ 0.68 (1.80), 0–2.7	
	Average (C+M)	0.24 $\pm$ 0.03 (0.20), 0–0.85	0.65 $\pm$ 0.04 (0.60), 0.15–1.5	1.50 $\pm$ 0.59 (1.65), 0.10–2.35	
	TGF- $\alpha$	Basal	0.20 $\pm$ 0.03 (0.17), 0–0.80	0.60 $\pm$ 0.04 (0.60), 0.05–1.3	–
	Carcinoma	0.61 $\pm$ 0.05 (0.60), 0.03–1.5	1.08 $\pm$ 0.05 (1.03), 0.20–1.9	1.26 $\pm$ 0.42 (1.30), 0.25–2.15	
FASE	Basal	0.07 $\pm$ 0.01 (0.03), 0–0.33	0.25 $\pm$ 0.02 (0.23), 0–0.73	–	
	Carcinoma	0.30 $\pm$ 0.03 (0.30), 0–0.87	0.55 $\pm$ 0.04 (0.52), 0.03–1.17	1.54 $\pm$ 0.43 (1.50), 0.60–2.70	

<sup>a</sup>Carcinoma, cytoplasm.

<sup>b</sup>Carcinoma, membrane.

Table 2  
Cut-off points, sensitivity, specificity, likelihood ratio (LR), and area under the curve (AUC) of ROC. Normal tissues from controls versus uninvolved tissues from lung cancer cases

Biomarker (Location)	Cut-off Point	Sensitivity (%)	Specificity (%)	LR	AUC of ROC (%)
p185 <sup>erbB-2</sup>					
Basal	0.70	60	85	3.14	80
Cytoplasm	1.30	65	85	1.99	82
Membrane	0.40	80	92	7.27	89
EGFR					
Basal	1.00	64	83	3.96	76
Cytoplasm	0.40	78	58	1.22	76
Membrane	0.10	91	75	1.29	90
TGF- $\alpha$					
Basal	0.40	72	89	4.84	86
Carcinoma	0.70	83	67	1.47	80
FASE					
Basal	0.10	74	79	2.90	78
Carcinoma	0.40	60	80	2.04	75

#### 4. Discussion

Selected cut-off points for p185<sup>erbB-2</sup> and EGFR membrane levels appear to have good discriminating power to differentiate controls versus patients with lung cancer. This was true for both comparisons between

controls versus uninvolved tissues from lung cancer cases and between controls versus carcinoma tissues from lung cancer cases. In fact, better discrimination based on higher AUC, was observed for the latter comparison, indicative of the trend of increasing levels of biomarkers from controls to uninvolved tissues to carci-

Table 3  
Cut-off points, sensitivity, specificity, likelihood ratio (LR), and area under the curve (AUC) of ROC. Normal tissues from controls versus carcinoma tissues from lung cancer cases

Biomarker (Location)	Cut-off Point	Sensitivity (%)	Specificity (%)	LR	AUC of ROC (%)
p185 <sup>erbB-2</sup>					
Cytoplasm	1.00	70	50	0.90	67
Membrane	0.40	89	92	6.67	93
EGFR					
Cytoplasm	1.00	74	98	10.2	91
Membrane	0.40	91	98	0.93	97
TGF- $\alpha$	0.70	86	67	1.03	86
FASE	0.70	98	96	2.41	99.6

Table 4  
Attributable proportion (APs) of Biomarkers

Biomarker	Location	AP	
		Uninvolved Tissues From Lung Cancer Cases	AP Carcinoma Tissues From Lung Cancer Cases
p185 <sup>erbB-2</sup>			
	Basal	0.35	–
	Cytoplasm	0.41	0.24 Carcinoma, Cytoplasm
	Membrane	0.63	0.78 Carcinoma, Membrane
EGFR			
	Basal	0.38	–
	Cytoplasm	0.43	0.57 Carcinoma, Cytoplasm
	Membrane	0.79	0.83 Carcinoma, Membrane
TGF- $\alpha$			
	Basal	0.52	–
	Carcinoma	0.58	0.65
FASE			
	Basal	0.50	–
	Carcinoma	0.34	0.96

noma tissues of lung cancer patients. Interestingly, the selected cut-off point for FASE showed superior discriminating power between control versus carcinoma tissues from cases. The study of Piyathilake et al. [18] reported a statistically significant step-wise increase in FASE expression from uninvolved bronchial mucosa to epithelial hyperplasia to SCC suggesting its possible involvement in the development and progression of SCC of the lung.

In this study, we utilized two independent patient groups, i.e. controls and patients with lung cancer to arrive at biomarker cut-off points that differentiate between the two samples. Methods for determining cut points of a prognostic variable in instances wherein only the sample of cancer patients are available include the use of a biological criteria, medians and upper limits of normality (mean  $\pm$  2x standard deviations), and the minimum p-value approach [19]. The latter approach which involves a systematic cut point search leads to inflation of the type I error arising from multiple testing. As such, the use of this method requires adjustment of the inflated p-values; however, methods for adjustment are not frequently used in many prognostic studies. The availability of data from control subjects allowed us to

use a simple and intuitive approach that will best dichotomize biomarker levels into high expression/risk vs. low expression/risk. This provides an alternative approach to the use of arbitrarily defined cut points at the median value or the use of multiple testing to search for cut points. Since the cut-points were selected independent of patient outcome (e.g. response rate, overall survival, disease-free survival), they can then be used as dichotomized prognostic variables without adjustment on significance levels of hypothesis testing.

Our criteria for obtaining cut-off points was based on the best balance between sensitivity and specificity as recommended by Pina et al. [13]. However, our reference population consists of subjects with other types of lung diseases besides lung cancer. As such, we were not able to determine the potential improvement in sensitivity levels from the use of healthy controls [13] in this study. A cohort of healthy controls and additional lung cancer cases is currently being evaluated by authors of this study and will provide for this additional evaluation as well as a serve as an independent patient group to validate the biomarker cut-off points identified in this study.

The use of control subjects also allowed us to validate the value of biomarkers as surrogate endpoints for cancer outcome. With the exception of FASE, this study found only small to moderate proportions of lung cancer cases can be attributed to high levels of p185<sup>erbB-2</sup>, EGFR, and TGF- $\alpha$ . This stresses the importance of validating biomarkers in different study populations and in other cancer sites in order to arrive at a more definitive conclusion regarding the value of a biomarker as a surrogate for cancer outcome.

## References

- [1] M.B. Sporn and G.J. Todaro, Autocrine growth factors and cancer, *Nature* **313** (1985), 745–751.
- [2] G.J. Kelloff, J.R. Johnson and J.A. Crowell et al., Approaches to the development and marketing approval of drugs that prevent cancer, *Cancer Epidemiology, Biomarkers, and Prevention* **4** (1995), 1–10.
- [3] K.A. Cronin, L.S. Freedman, R. Lieberman, H.L. Weiss, S.W. Beenken and G.J. Kelloff, Bayesian monitoring of phase II trials in cancer chemoprevention, *J Clinical Epidemiology* **52** (1999), 705–711.
- [4] A.S. Goustin, E.B. Leof, G.D. Shipley and H.L. Moses, Growth factors and cancer, *Cancer Research* **46** (1986), 1015–1029.
- [5] C.J. Piyathilake, A.R. Frost, U. Manne, H. Weiss, W.C. Bell, D.C. Heimburger and W.E. Grizzle, Differential expression of growth factors in squamous cell carcinoma and precancerous lesions of the lung, *Clinical Cancer Research* **8** (2002), 734–744.
- [6] F.J. Hendler and B.W. Ozanne, Human squamous cell lung cancers express increased epidermal growth factor receptors, *J. Clin. Investig.* **74** (1984), 647–651.
- [7] D. Veale, T. Ashcroft, C. Marsh, G.J. Gibson and A.L. Harris, Epidermal growth factor receptors in non-small cell lung cancer, *Br. J. Cancer* **55** (1987), 513–516.
- [8] V. Rusch, D. Klimstra, I. Linkov and E. Dmitrovsky, Aberrant expression of p53 or the epidermal growth factor receptor is frequent in early bronchial neoplasia and coexpression precedes squamous cell carcinoma development, *Cancer Research* **55** (1995), 1365–1372.
- [9] D. H. Harpole, J.R. Marks, W.G. Richards, J.E. Herndon and D.J. Sugarbaker, Localized adenocarcinoma of the lung: oncogene expression of erbB-2 and p53 in 150 patients, *Clinical Cancer Research* **1** (1995), 659–664.
- [10] P. Pfeiffer, P.P. Clausen, K. Andersen and C. Rose, Lack of prognostic significance of epidermal growth factor receptor and the oncoprotein p185HER-2 in patients with systemically untreated non-small-cell lung cancer: an immunohistochemical study on cryosections, *Br. J. Cancer* **74** (1996), 86–91.
- [11] V. Rusch, J. Baselga, C. Cordon-Cardo, J. Orazem, M. Zaman, S. Hoda, J. McIntosh, J. Kurie and E. Dmitrovsky, Differential expression of the epidermal growth factor receptor and its ligands in primary non-small cell lung cancers and adjacent benign lung, *Cancer Research* **53** (1993), 2379–2385.
- [12] V. Gorgoulis, D. Aninos, P. Mikou, P. Kanavaros, A. Karameris, J. Joardanoglou, A. Rasidakis, M. Veslemes, B. Ozanne and D.A. Spandidos, Expression of EGF, TGF- and EGFR in squamous cell lung carcinomas, *Anticancer Research* **12** (1992), 1183–1187.
- [13] T.C. Pina, I.T. Zapata, J.B. Lopez, J.L. Perez, P.P. Paricio and P.M. Hernandez, Tumor markers in lung cancer: Does the method of obtaining the cut-off point and reference population influence diagnostic yield? *Clinical Biochemistry* **32** (1999), 467–472.
- [14] B.C. Choi, Slopes of a Receiver Operating Characteristic Curve and Likelihood Ratios for a Diagnostic Test, *American Journal of Epidemiology* **148** (1998), 1127–1132.
- [15] W.E. Grizzle, R.B. Myers, U. Manne and S. Srivastava, Immunohistochemical evaluation of biomarkers in prostatic and colorectal neoplasia, in: *John Walker's Methods in Molecular Medicine. Tumor Marker Protocols*, M. Hanausek and Z. Walaszek, eds, Humana Press, Inc. Totowa, NJ, 1998, pp. 143–160.
- [16] W.C. Black and P. Armstrong, Communicating the significance of radiologic test results: the likelihood ratio, *Am J. Roentgenol* **147** (1986), 1313–1318.
- [17] A. Schatzkin, L.S. Freedman, M.H. Schiffman and S.M. Dawsey, Validation of Intermediate End Points in Cancer Research, *Journal of the National Cancer Institute* **82** (1990), 1746–1752.
- [18] C.J. Piyathilake, A.R. Frost, U. Manne, W.C. Bell, H. Weiss, D.C. Heimburger and W.E. Grizzle, The Expression of Fatty Acid Synthase (FASE) is an Early Event in the Development and Progression of Squamous Cell Carcinoma of the Lung, *Human Pathol* **31** (2000), 1068–1073.
- [19] M. Mazumdar and J.R. Glassman, Categorizing a prognostic variable: Review of methods, code for easy implementation and applications to decision-making about cancer treatments, *Statistics in Medicine* **19** (2000), 113–132.