Establishment of a Multi-Analyte Serum Biomarker Panel to Identify Lymph Node Metastases in Non-small Cell Lung Cancer

Jeffrey A. Borgia, PhD,† Sanjib Basu, PhD,‡ L. Penfield Faber, MD,§ Anthony W. Kim, MD,§ John S. Coon, MD, PhD,† Kelly A. Kaiser-Walters, BS,† Cristina Fhied, MS,* Sherene Thomas, PhD,* Omid Rouhi, MD,*‖ William H. Warren, MD,§ Philip Bonomi, MD,‖ and Michael J. Liptay, MD§*

Introduction: In non-small cell lung cancer (NSCLC), the presence of locoregional lymph node metastases remains the most important prognostic factor and significantly guides treatment regimens. Unfortunately, currently-available noninvasive staging modalities have limited accuracy. The objective of this study was to create a multianalyte blood test capable of discriminating a patient's true (pathologic) nodal status preoperatively.

Methods: Pretreatment serum specimens collected from 107 NSCLC patients with localized disease were screened with 47 biomarkers implicated in disease presence or progression. Multivariate statistical algorithms were then used to identify the optimal combination of biomarkers for accurately discerning each patient's nodal status.

Results: We identified 15 candidate biomarkers that met our criteria for statistical relevance in discerning a patient's preoperative nodal status. A 'random forest' classification algorithm was used with these parameters to define a 6-analyte panel, consisting of macrophage inflammatory protein-1α, carcinoembryonic antigen, stem cell factor, tumor necrosis factor-receptor I, interferon-γ, and tumor necrosis factor-α, that was the optimum combination of biomarkers for identifying a patient's pathologic nodal status. A Classification and Regression Tree analysis was then created with this panel that was capable of correctly classifying 88% of the patients tested, relative to the pathologic assessments. This value is in contrast to our observed 85% classification rate using conventional clinical methods.

Conclusions: This study establishes a serum biomarker panel with efficacy in discerning preoperative nodal status. With further validation, this blood test may be useful for assessing nodal status (including occult disease) in NSCLC patients facing tumor resection therapy.

Departments of *Biochemistry, †Pathology, ‡Biostatistics, §Thoracic Surgery, and ‖Medical Oncology, Rush University Medical Center, Chicago, Illinois.

Disclosure: The authors declare no conflicts of interest.

Address for correspondence: Jeffrey A. Borgia, PhD, Department of Biochemistry, Rush University Medical Center, 558 Cohn Research Building, 1735 W. Harrison St, Chicago, IL 60612. E-mail: jeffrey_borgia@rush.edu

ISSN: 1556-0864/09/0403-0338

Key Words: Biomarker, Serum, Non-small cell lung cancer, Nodal status, Luminex.

(*J Thorac Oncol.* 2009;4: 338–347)

Lung cancer is the most common cause of cancer related mortality worldwide. In the USA, over 200,000 new cases will be diagnosed this year with the majority having advanced disease not amenable to potentially curative surgery.[1] Non-small cell lung cancer (NSCLC) accounts for 80 to 85% of all lung cancer cases and is comprised of three major histopathological subtypes, which include adenocarcinoma, squamous cell carcinoma, and large cell carcinoma.

Approximately 25% of NSCLC patients present with localized, early-stage disease. Surgical resection with accompanying lymph node sampling/dissection remains the cornerstone of current treatment for these patients and clearly improves patient outcome.[2] Accurate lymph node staging is crucial to the optimal clinical management of NSCLC and remains the foremost prognostic factor for patient outcome with localized tumors.[3] Unfortunately, micrometastatic disease is often missed by conventional pathologic analysis (hematoxylin and eosin staining). With this, the risk of disease recurrence and death remains substantial even in early-stage patients; with 5-year survival rates for pathologic stage I and II disease of 57 to 85% and 38 to 65%, respectively.[1] The use of adjuvant (postoperative) chemotherapy to obviate disease recurrence for all patients receiving tumor resection is an attractive idea, but remains a controversial subject. That is, although the population of resected patients with stage I-IIIa NSCLC receive a considerable survival advantage from adjuvant chemotherapy,[4] unselected subpopulations of stage I patients do not receive a statistical benefit.[5] Methods that could identify stage I patients with occult micrometastatic disease would provide a means to select for which patients may gain a survival benefit from systemic therapy as part of their postoperative care. New diagnostic tests based on molecular information from the primary tumor are urgently needed to enhance current staging protocols and identify which patients have occult micrometastatic disease.

A variety of approaches has been documented in the literature designed to better stage patients using molecular methods. Perhaps most notable of these were the studies that investigated associations between the patterns of gene expression and specific clinical parameters, such as lymph node metastasis[6–9] or prognosis.[10,11] Relevant to the objectives of this study, Choi et al. were able to identify several multigene expression signatures from microarray data sets that each have value in predicting lymph node metastases in NSCLC patients.[6–9,12,13] These multigene panels survey a range of known tumor-associated biologic pathways (including cell cycle regulation, apoptosis, and angiogenesis) and define a 'molecular phenotype' for the tumor.

There is mounting evidence that multianalyte panels of biomarkers offer clear statistical advantages over individual biomarkers for discerning diagnostic and prognostic features across a variety of disease states, including cancer.[14–17] Not unexpectedly, the early attempts at formulating a multianalyte blood test for NSCLC relied heavily on the well-characterized, prognostic serum biomarkers, such as carcinoembryonic antigen (CEA) and the cytokeratin 19 fragment, CYFRA 21.1.[18–21] Expansion of this CEA/CYFRA 21.1 combination to include auxiliary biomarkers, such as CA125, Neuron specific enolase, Macrophage-colony stimulating factor (M-CSF), and tissue polypeptide antigen[22–24] was also investigated as a means to further improve the prognostic power of biomarker combinations for this purpose. Soon, the benefits of adding biomarkers such as retinol binding protein, α1-antitrypsin, and squamous cell carcinoma antigen to the CEA/CYFRA 21.1 combination became the focus of several test development efforts, with the best of these boasting 72.8% sensitivity and 75.4% specificity[25,26] for correctly diagnosing patients with NSCLC. Despite the extensive development efforts put forth by the scientific community, none of these biomarker combinations ever passed the appropriate validation steps to have a clinical impact. However, a series of recent publications detailing the large-scale evaluation of biomarkers with the Luminex immunoassay platform with new multivariate regression methods have reinvigorated interest in this multivariate approach within the cancer community.[14,16,27]

In this article, we use this multivariate biomarker panel strategy to screen 107 NSCLC patients for 47 serum markers, selected based on either documented relevance to NSCLC progression[8,19,23–25,28–35] or involvement in biologic pathways believed to be important to metastatic progression. Our objective was to develop a multianalyte blood test to discern a patient's pathologic lymph node status preoperatively.

PATIENTS AND METHODS

Patient Populations

Between October 2004 and May of 2008, 107 patients with NSCLC deemed potentially operable were enrolled in this study. All patients included in this study had disease confined to the chest without evidence of distant metastases and any patient who received preoperative chemo- or radiotherapy before our initial blood draw were excluded. The clinicopathologic data for these patients are outlined in Table 1.

TABLE 1. Patient Demographics

Sex		
Male	41	
Female	66	
Age		
Range	14–87	
Median	72	
NSCLC stage		
	Clinical	Pathological
Ia	38	24
Ib	28	44
IIa	4	2
IIb	13	12
IIIa	8	19
IIIb	2	4
IV	2	0
Not determined	12	2
Histopathological subtype		
Adenocarcinoma	62	
Adenosquamous	2	
Squamous cell carcinoma	31	
NSCLC 'other'	12	
Procedure performed		
Lobectomy	71	
Pneumonectomy	9	
Segmectomy	16	
Wedge w/biopsy	1	
Biopsy/other	10	

Clinicopathological characteristics of the tested patient population.
NSCLC, non-small cell lung cancer.

Our standard clinical staging work-up included a computed tomography (CT) scan of the chest (including liver and adrenal glands) and positron emission tomography (PET) scan. Magnetic resonance imaging of the brain was included for bulky disease or symptoms were present. Clinically positive mediastinal lymph nodes identified by CT (short axis dimension $\geq$1 cm), or hilar–mediastinal nodal PET activity warranted nodal biopsies via cervical mediastinoscopy or thoracoscopy.

Patients with pathologically positive mediastinal lymph nodes ($n = 36$) were treated with either induction or definitive chemotherapy with radiation. Those without evidence of metastases to the mediastinal lymph nodes were taken to surgery for a complete anatomic resection of their lung cancer with lymph node dissection (Table 1). Standard lymph node dissections of both hilar and mediastinal nodes consisted of the following: lymph node station levels 2, 4, 7, and 10 for right upper and middle lobes and levels 2, 4, 7, 8, 9, and 10 for right lower lobes. Left-sided resections sampled lymph node station levels 5, 6, 7, and 10 for left upper lobes and levels 5, 6, 7, 8, 9 and 10 for left lower lobes. Surgical pathology routinely performed hematoxylin and eosin staining to identify the presence of metastatic disease with immunostaining performed only as indicated to confirm disease origin.

Patient data, tissue and serum were acquired with written formal consent and in compliance with the Institutional Review Board at Rush University Medical Center. Patients were approximately age and sex matched between the groups for nodal status, with the distribution heavily favoring early-stage disease. Adenocarcinoma was a 2:1 more common histopathological subtype than squamous cell carcinoma; with the number of 'other' NSCLC subtypes representing only a small portion of the total patients enrolled in this study. All pathologic staging data referenced herein were derived from tissue from lymph node sampling at mediastinoscopy or lymph node dissections accompanying tumor resection.

Collection and Storage of Serum Specimens

Peripheral blood was collected from each patient immediately before treatment initiation for NSCLC. Ten milliliter of blood was drawn into standard red-top Vacutainers (without anticoagulants) and coagulated at room temperature for 30 to 40 minutes. Sera was then separated with centrifugation. Yields ranged 4 to 7 ml of serum per 10 ml of whole blood. Sera was then immediately divided into aliquots and archived in an −80°C ultra-low temperature freezer. No specimens were subjected to more than two thaw cycles for this study.

Measurement of Serum Biomarker Concentrations

Luminex xMAP immunobead technology was used whenever possible for the measurement of circulating biomarker concentrations reported in this study. This platform offers most of the advantages of conventional sandwich immunoassays, but with a much higher throughput and increased sensitivity over the conventional enzyme-linked immunosorbent assay (ELISA) assays. The assays were performed according to the respective manufacturer's instructions and were conducted in the following groupings; Interleukin-1β (IL-1β), IL-1ra, IL-6, IL-8, IL-10, tumor necrosis factor-α (TNF-α), and transforming growth factor-α [Millipore; Billerica, MA]; C-reactive protein (CRP) and serum amyloid A [Millipore; Billerica, MA]; epidermal growth factor, basic fibroblast growth factor, vascular endothelial growth factor, and granulocyte colony stimulating factor (G-CSF) [Invitrogen; Carlsbad, CA]; death receptor 5 (DR-5), TNF–receptor I (TNF-RI), and TNF-RII [Invitrogen; Carlsbad, CA]; Regulated on Activation, Normal T Expressed and Secreted, macrophage inflammatory protein-1α (MIP-1α), MIP-1β, monocyte chemotactic protein-1, and eotaxin [Invitrogen; Carlsbad, CA]; matrix metalloproteinases (MMPs) -2, -3, -9, and -13 [R&D Systems; Minneapolis, MN]; soluble intracellular adhesion molecule 1 (sICAM 1), sE-selectin, and sP-selectin [R&D systems; Minneapolis, MN]; IL-2, IL-13, interferon-γ (IFN-γ), interferon-inducible protein 10, and granulocyte monocyte colony stimulating factor (GM-CSF) [Bio-Rad Laboratories; Hercules, CA]; IL-1α, IL-2Rα, M-CSF, stem cell factor (SCF), and stem cell-derived factor 1α (SDF-1α) [Bio-Rad Laboratories; Hercules, CA]; and sEGFR (erb-b1), Her-2 (erb-b2), CA125, CA15–3, CA19–9, CEA, and CYFRA 21.1 were measured at the University of Pittsburgh Cancer Institute's Luminex Core Facility, on a fee-for-service basis. Biomarker concentrations were calculated through a five-parametric curve fit as part of the BioPlex Suspension Array System Software v4.0 (Bio-Rad Laboratories; Hercules, CA). Measurements of TIMP-1 and osteopontin concentrations were conducted using commercially-available ELISA assays and in accordance to the kit directions (R&D Systems; Minneapolis, MN). Data were collected on a BioTek PowerWave XS plate reader using KC Junior (v1.40.3) software package. A four-parametric curve fit was used to calculate the concentrations from the raw absorbance readings. The ELISA platform was employed for these two biomarkers due to the lack of commercially-available Luminex platform assays at the time of study performance. All assays performed for this study were conducted in a blinded fashion.

Statistical Methods

Individual biomarker evaluation

Descriptive statistics (median, range) and graphical displays (histogram, box-plot, normal probability plot) for the range of concentration values associated with each biomarker were obtained using SPSS 15.0 for Windows (SPSS Inc., Chicago, IL). Biomarkers which had more than one-third of their values outside the range of the assay were disqualified from further consideration for panel establishment. For the remaining biomarkers, values which were reported as missing due to being below the lower limit of the assay were replaced by the lowest measured value of that biomarker. A nonparametric Mann-Whitney rank sum test was then used to assess the difference in biomarker concentrations according to the preoperative (pathologic) nodal status (i.e., positive versus negative). Receiver operating characteristic (ROC) curves for classifying nodal status were calculated for each individual biomarker using SPSS 15.0. In nearly all cases, our data compared favorably to previously reported values for the individual biomarkers.[8,19,23–25,28–35] Subgrouping of the pathologic lymph node positive group into either single or multiple station involvement was based on intraoperative nodal dissection with pathologic confirmation. Patients ($n = 5$) deemed inoperable by our initial evaluation (with positive mediastinal biopsies) were excluded from these subgroupings. A nonparametric Mann-Whitney rank sum test was used to evaluate the association of select biomarkers against this surrogate for the extent of nodal involvement.

Multivariate Panel Selection and Analysis

The optimal multivariate panel of biomarkers was chosen based on variable selection algorithms performed within the random forests package in R.[36,37] We applied liberal inclusion criteria for the individual biomarkers (a Mann-Whitney p value smaller than 0.20 or an area under the ROC curve [AUC] higher than 0.60) to ensure that no biomarker with potential value in a multianalyte panel was prematurely excluded from this selection process based on a weak individual performance. Briefly, the Random Forests package selects optimal combinations of biomarkers by growing numerous (1000 in the present study) cross-validated classification trees for each subpanel of biomarkers, with each tree used to predict group membership for each case. These are counted as the tree "votes" for that

group. The forest chooses the group membership having the most votes over all the trees in the forest. Each such tree is grown by cross-validation; where a training set (approximately two-thirds of the values) is randomly selected from the full data and each tree is grown on this training data to the largest extent possible (no pruning). The resultant tree is then used to predict the group membership for the remaining test cases, which is termed as an *out-of-bag* (OOB) prediction. This process is then repeated 1000 times; that is, another training set is randomly selected and a new tree is grown and used to perform another OOB prediction. The classification accuracy of the random forest is measured by the averaged error of the OOB predictions across the entire forest; this is termed the OOB error rate. The OOB error thus uses disjoint subsets of the data for model fitting and validation repeatedly. This cross-validation is also used to compute a variable importance for each biomarker included in the Random Forest analysis. The stepwise selection method sequentially searches for optimal subpanel of markers where the marker with the lowest variable importance score from the Random Forest are removed at each step.

The optimal 6-analyte panel of biomarkers resulting from the Random Forest variable selection process was then used by a Classification and Regression Tree (CART) algorithm to model a classification tree for identifying a patient's true (pathologic) preoperative lymph node status. This analysis was performed using the RPART package of the R statistical software suite.[38] Briefly, classification trees determine a set of binary *if-then* logical (split) conditions that permit accurate classification of (in this case) the patient's nodal status. The CART algorithm discriminates between groups by splitting the range of values measured for each individual biomarker at all of its possible split points. The 'goodness of split criterion' is then used to determine the best split point for each biomarker for predicting nodal status. CART then ranks all of the "best" splits on each biomarker and selects the best biomarker and its split point for the split at the root node. CART then assigns classes to the two split nodes according to a rule that minimizes misclassification error. This process is continued at each nonterminal child node and at each of the successive stages until all observations are perfectly classified or the sample size within a given node is too small to divide ($n \leq$ a user-supplied number; such as 5). The final output of the resulting classification tree is a graphical display of decision criteria for each split, with the resulting predicted group memberships at the terminal nodes (Figure 2). The predicted probabilities of preoperative nodal status from the tree were used to obtain sensitivity and specificity across a range of cut-points for decision rules and the resulting ROC curve. We also appraised several other classification methods (logistic regression, neural networks, and support vector machines) for their ability to correctly predict group classification. These methods work with linear combinations of biomarker concentrations (instead of splits) and were found to be less optimal than CART in our application.

RESULTS

Analysis of Individual Serum Biomarkers According to Pathologic Lymph Node Status

In an effort to establish a multianalyte panel of serum biomarkers capable of accurately detecting preoperative lymph node status in NSCLC patients (relative to the pathologic assessments), we selected an array of 47 biomarkers to screen the sera from 107 NSCLC patients treated at Rush University Medical Center. The clinical and pathologic characteristics of the patients are outlined in Table 1. All biomarkers were selected based either on published reports for each biomarker showing value for at least one of the following functions: NSCLC diagnosis, staging, or prognosis[8,19,23–25,28–35] or involvement in biologic processes implicated in disease progression. Several biomarkers, including IL-1α, IL-1β, IL-2, IL-15, GM-CSF, transforming growth factor-α, DR5, MMP-13, had a large portion of their overall data set fall below the threshold of assay range (using the manufacturer's suggested dilution factor) and were disqualified from further analysis. No trends in the raw data from these biomarker candidates were observed to justify re-evaluation/remeasurement.

Serum concentrations of M-CSF, SCF, IFN-γ, were found to be significantly higher in the nodal negative group (Mann-Whitney two-sided p-values 0.04, 0.01, and 0.01 respectively) whereas the concentration of CRP was found to be significantly higher in the nodal positive group (p value = 0.03), both relative to pathologic staging. The difference in MIP-1α concentration between the nodal negative and positive groups was marginally significant (p value = 0.08). In an effort to avoid the premature exclusion of biomarkers that may benefit a multianalyte panel, but may not be a satisfactory predictor individually, we liberally defined the threshold needed to be considered statistically-relevant to be a ROC analysis "AUC" value greater than 0.60 or significant Mann-Whitney rank sum test statistics (i.e., p value) less than 0.20. Based on these criteria, we found 15 of the original 47 biomarkers to individually possess value for classifying patient groups based on lymph node status. MIP-1α, SCF, G-CSF, M-CSF, sICAM-1, TNF-α, TNF-RI, TNF-RII, IFN-γ, sP-selectin, CEA, IL-2Rα, osteopontin, CRP and IL-1ra exhibited the most promising overall profiles. The results of this analysis are presented in Table 2 and individual analyte profiles displayed as 'box and whisker' plots in Figure 1.

Classifications Based on a Multianalyte Panel for Identifying Preoperative Nodal Status

A classification tree based on a subpanel of six markers (consisting of CEA, MIP-1α, SCF, TNF-RI, IFN-γ, and TNF-α) was used with CART methodology to predict preoperative nodal status for our patient population. The classification tree used in this process is represented in Figure 2. This tree correctly classified 94 of the 107 cases (a correct classification rate of 88%). The predicted probabilities of being nodal positive from the tree were used to obtain sensitivity and specificity across a range of cut-points for decision rules and the resulting ROC curve is shown in Figure 3. The total area under this ROC curve (AUC) was at 0.94, with 88% sensitivity and 87% specificity. As we anticipated, a substantial gain in our ability to predict nodal

TABLE 2. Individual Biomarkers with Efficacy in Classifying Nodal Status in NSCLC

	NSCLC: Node Negative[a] (n = 71)		NSCLC: Node Positive[a] (n = 36)			
Biomarker	**Median[b]**	**Range[b]**	**Median[b]**	**Range[b]**	**AUC**	**Mann-Whitney U**
MIP-1α	0.13	0.09–0.70	0.12	0.05–0.81	0.775	0.081
SCF	52.8	13.2–246.4	46.5	16.7–110.4	0.715	0.011
TNF-α	14.5	4.5–55.0	13.7	2.9–41.9	0.689	0.141
IL-2R$_\alpha$	45.2	2.1–452	47.0	7.0–192.8	0.688	0.598
TNF-RII	1.87	0.32–7.87	1.60	0.22–4.42	0.684	0.223
M-CSF	2.27	0.05–34.7	1.28	0.05–12.0	0.679	0.041
sICAM-1	9608	5148–17,860	10,128	5058–28,376	0.675	0.408
CEA	1392	255–54,741	1,717	255–32,506	0.660	0.191
G-CSF	0.117	0.014–0.070	0.150	0.014–0.580	0.657	0.247
IFN-γ	33.2	5.7–2258	8.2	5.7–1417	0.629	0.011
TNF-RI	1.69	0.62–8.14	1.41	0.57–8.13	0.623	0.546
Osteopontin	25.3	0–195	22.5	0–190.0	0.616	0.640
IL-1ra	290.7	1.2–5796	321.7	1.1–3258	0.600	0.898
CRP	212.6	1.5–1018.5	293	1.5–992.2	0.576	0.027
MMP-2	6006	3293–12,536	5589	897–10,094	0.568	0.118

[a]Based on pathologic staging.
[b]Values expressed as pg/ml.
NSCLC, non-small cell lung cancer; MIP-1α, macrophage inflammatory protein; SCF, stem cell factor; TNF, tumor necrosis factor; IFN-γ, interferon; TNF-RI, tumor necrosis factor-receptor I; MMP, matrix metalloproteinases; CRP, C-reactive protein; G-CSF, granulocyte colony stimulating factor; sICAM-1, soluble intracellular adhesion molecule 1; M-CSF, monocyte colony stimulating factor; CEA, carcinoembryonic antigen; IL, Interleukin, AUC, area under the curve.

status was achieved by comparing our multianalyte ROC curve against the ROC curves observed from individual analytes.

The six biomarker panel used for construction of the classification tree was selected from the 15 biomarker candidates using variable selection by a Random Forests algorithm. The averaged OOB misclassification errors from the range of the 1000 trees of the Random Forest grown for each of their respective subpanels are shown in Table 3. Clearly, we found that the continued 'focusing' of the panel improved our ability to correctly classify patients relative to the pathologic nodal status. However, after the 6th iteration these gains approached negligibility, while the range of biologic processes surveyed started to become limiting. Ultimately, we decided that the optimal balance of test performance with range of biologic pathways evaluated was represented best by the 5th iteration, which provided a 6-analyte panel (consisting of CEA, MIP-1α, SCF, TNF-RI, IFN-γ, and TNF-α) that is used in the subsequent CART assessments for preoperative nodal status.

When we evaluated members of our 6-analyte panel individually against subgroupings of lymph node positive disease (single- versus multiple-nodal station involvement), trends in serum biomarker concentrations associated predictably with increasing nodal involvement (Figure 4). This substudy was performed as a means to confirm the relevance of the associations of the individual biomarkers with increasing nodal involvement.

There were 13 cases in which lung cancer stage was classified differently by our blood test relative to pathologic staging. These consisted of one patient who was erroneously predicted to have a negative nodal status by our methods, but was pathologically determined to be node positive, and 12 patients predicted by our blood test to have a positive nodal status, but were actually (pathologically) node negative. In comparison, our clinical staging protocols, which are based on PET and CT-based criteria (described in the Methods section), provided an 85% accuracy rate for identifying preoperative lymph node status, relative to the pathologic staging. This value is well within the range of published accuracy figures from other investigators.[39,40] Clinical staging (imaging) discrepancies consisted of seven up-staged and eight down-staged patients upon pathologic staging. With this, we observed the clinical protocols to be approximately equally prone to misidentifying a patient's nodal status relative to the pathologic assessment, whereas our multianalyte, molecular approach tended to err towards classifying the patient as node-positive - requiring that nearly all of the misclassified patients be down-staged upon pathologic evaluation. Interestingly, only one misclassification was common to both clinical and molecular staging methods.

DISCUSSION

Surgery remains the mainstay of therapy for localized NSCLC. In this setting, the status of the locoregional lymph nodes is the most important prognostic factor.[41] When lymph nodes are involved, the 5-year survival is markedly reduced compared with stage I node negative patients. In current National Comprehensive Cancer Network guidelines, when lymph node metastases are identified, consideration should be given for adjuvant chemotherapy, as it improves survival.[42]

As a wide-ranging hypothesis for this study, we reasoned that the 'metastatic phenotype' responsible for locoregional lymph node progression in NSCLC patients is implicated in postoperative disease recurrence. This is likely due to the presence of undetected occult metastases at the time of

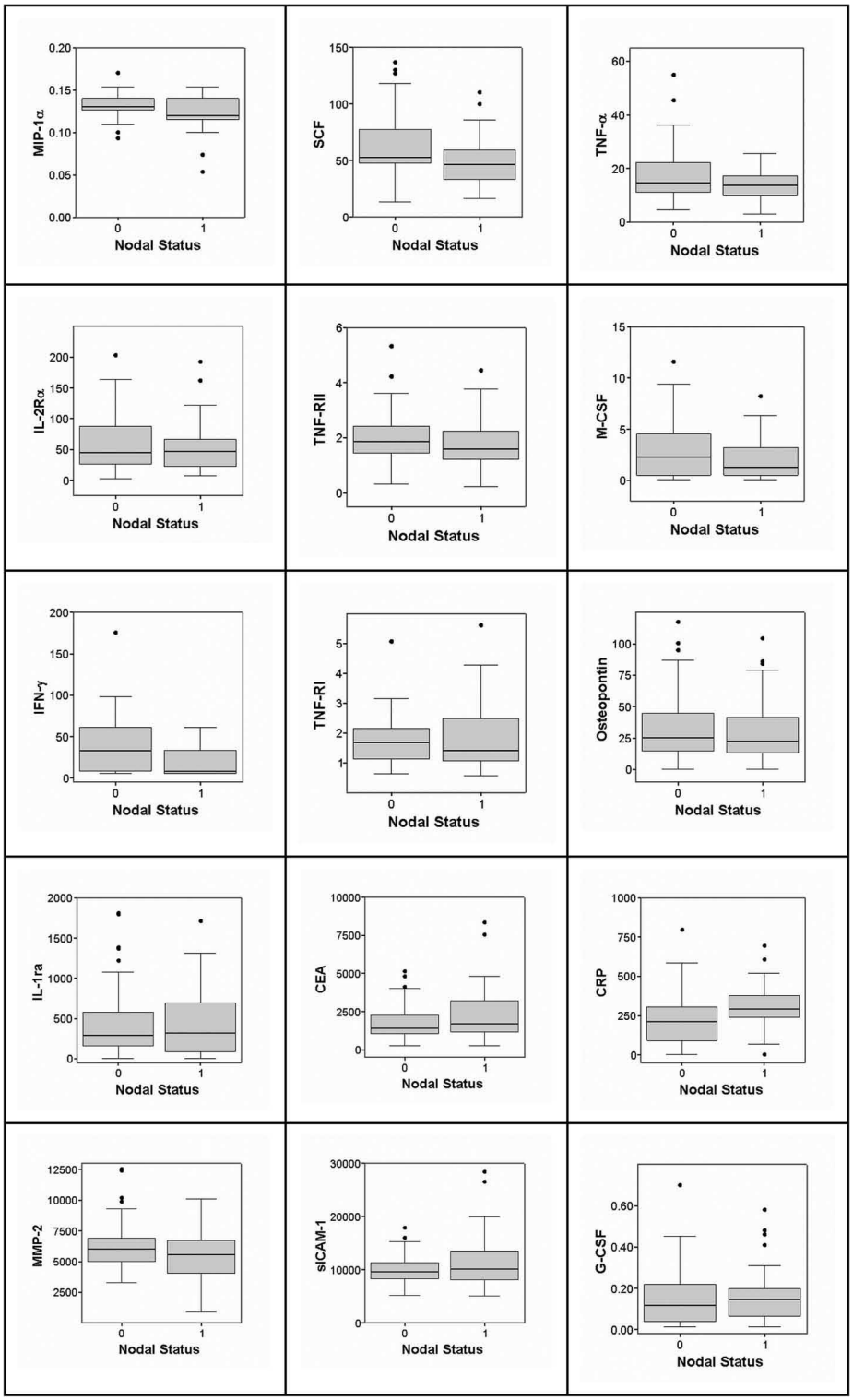

FIGURE 1. Boxplots of biomarker candidates from the Random Forest analysis. Box and Whisker plots were generated for each individual biomarker in the final iteration of the Random Forest analysis. Designations on the abscissa: "0" refers to lymph node negative, whereas "1" refers to lymph node positive (both relative to pathologic stage). The extreme values are hidden.

surgery. To test this hypothesis, we evaluated a range of serum analytes associated with NSCLC progression for value in discriminating preoperative lymph node status in NSCLC patients, relative to the pathologic assessments. Multivariate statistical analyses were then used to both identify the optimal combination of individual analytes relevant for detecting lymph node metastases and modeling a classification tree for determining the true (pathologic) nodal status preoperatively.

The 15 biomarkers which correlated most closely with pathologic nodal status represent a range of biologic pro-

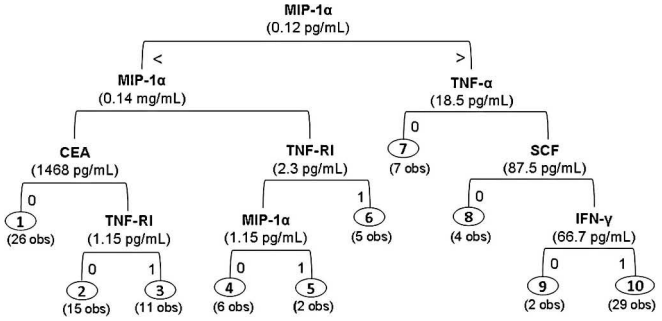

FIGURE 2. Classification and Regression Tree for predicting preoperative nodal status in non-small cell lung cancer (NSCLC) patients. A binary 'if-then' set of decision rules is used to split the data into separate branches of the tree with each node of the tree displaying the appropriate analyte and threshold concentrations. Additional classifications continue to the left if the threshold is not met, and to the right if the value is equal to or exceeding this cutoff. The number of classifications (observations) are listed immediately below each terminal node, with each final arm labeled (0 = node negative; 1 = node positive).

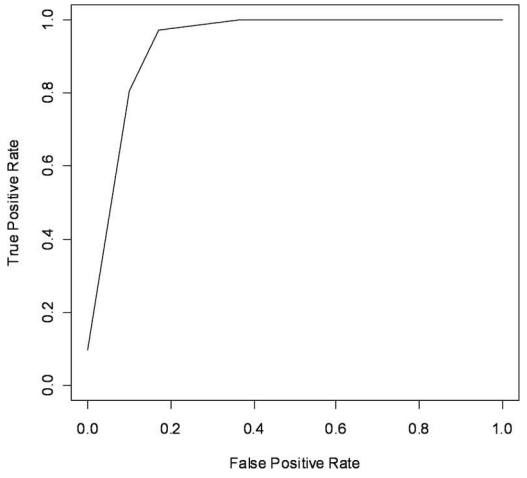

FIGURE 3. Receiver operator characteristic curve for the most robust staging subpanel generated by the Random Forest analysis. Area under the curve = 0.94; Sensitivity = 88%; Specificity = 87%.

cesses well associated with NSCLC progression. For example, a large body of data has been documented that implicates an inflammatory component to the development and progression of cancer, regardless of histologic origin.[24] Many of these inflammatory markers are within our top 15 most relevant biomarkers (e.g., TNF-α, TNF-RI, TNF-RII, IFN-γ IL-2R$_\alpha$, M-CSF, G-CSF, MMP-2, and CRP) and are thought to be used by tumors to modulate specific effector molecules involved in metastatic progression. Some (general) examples of these include IL-8 for neovascularization, MMP and adhesion molecules for tumor cell mobilization and overall invasive potential.[43] In addition, inflammatory molecules can also induce the production of 'secondary' biomarkers in neighboring cell types, such as the production of sICAM-1 by tumor-activated T-cells and lung fibroblasts. Biomarkers of this type typically provide a large degree of signal amplification and, therefore, offer excellent sensitivity relative to primary (tumor-originating) analytes. Although biomarkers associated with these somewhat generalized pathologic processes (such as inflammation) are not expected to lend very much specificity to the detection of tumor metastasis, they still may provide prognostic value for clinicopathological features of the disease. The importance of these "generic" cancer biomarkers may be punctuated by their inclusion in other multianalyte biomarker panels; such as those for cancers of the head and neck, ovaries, and skin.[14,16,17,27] More specifically, Yurkovetsky et al.[16] defined a role of G-CSF, MIP-1α, TNF-α, and TNF-RII in distinguishing high-risk melanoma patients from normal controls, with MIP-1α (and MIP-1β) showing the most promising efficacy in predicting incidence of disease recurrence. Further examples of the possible general role the 15 biomarkers found to have a role in predicting nodal status in NSCLC include: MMP-2, IFN-γ, MIP-1α, IL-1ra, and G-CSF had diagnostic efficacy in head and neck cancer,[27] CEA had efficacy in predicting liver metastases in colorectal cancer,[44,45] IL-2Rα had efficacy in predicting lymph node involvement in breast cancer,[46] and sICAM 1 was revealed from serum proteomics to be able to predict metastases in nasopharyngeal carcinoma.[47]

Interestingly, although many of our 15 most relevant biomarkers have already been well appreciated for their diagnostic or prognostic significance, not all the molecules in our 'top 15′ represent biomarkers documented to have any predictive value for NSCLC, per se. Instead, during the conceptualization of this study we reasoned that testing multiple targets of pathways already implicated in various mechanistic features of NSCLC progression may lead to the unveiling of previously unappreciated candidate biomarkers for diagnostic purposes. For example, members of the active lymphocytic response (ALR) pathway, such as RANTES and MIP-1β, have already been established to be strongly predictive of patient survival.[48] Interestingly, no chemokine in the ALR pathway previously demonstrated to be prognostic for NSCLC (including RANTES, MIP-1β, MCP-1, eotaxin, IP-1, IL-8) was found to have significant value in identifying preoperative lymph node metastases by our examinations. However, MIP-1α, another member of the ALR pathway, was one of our strongest biomarkers for classifying patients according to nodal status, despite the fact that it has no documented record for distinguishing any clinical parameter related to NSCLC. Along these lines, lymphocytes were also recently found to be activated by IL-2Rα after gefitinib administration.[49] This interplay between the platelet and lymphocyte activation mechanisms led us to speculate that IL-2Rα may also have value for assessing nodal status in NSCLC. Indeed, we found IL-2Rα to have considerable value in this capacity and it turned out to be one of the more discriminating analytes tested in our study. In another case, the considerable attention focused on the role of progenitor and stem cells in NSCLC progression[50] tempted us to screen for circulating biomarkers for these cells, including SDF-1α and SCF. Although Mroczko et al.[51] previously established a diagnostic role for SCF, to our knowledge we are the

TABLE 3. Variable Selection Using Random Forests of the 15 Biomarker Candidates

		Biomarkers															
	Variables	**MIP-1α**	**SCF**	**CEA**	**TNF-RI**	**TNF-α**	**IFN-γ**	**M-CSF**	**G-CSF**	**TNF-RII**	**sICAM-1**	**MMP-2**	**CRP**	**IL-2Rα**	**OP**	**IL-1ra**	**OOB**
1	15	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	0.336
2	12	X	X	X	X	X	X	X	X	X	X	X	X				0.308
3	10	X	X	X	X	X	X	X	X	X	X						0.308
4	8	X	X	X	X	X	X	X	X								0.317
5	6	X	X	X	X	X	X										0.308
6	5	X	X	X	X	X											0.289
7	4	X	X	X	X												0.345

MIP-1α, macrophage inflammatory protein; SCF, stem cell factor; TNF, tumor necrosis factor; IFN-γ, interferon; TNF-RI, tumor necrosis factor-receptor I; MMP, matrix metalloproteinases; CRP, C-reactive protein; G-CSF, granulocyte colony stimulating factor; OP, osteopontin; sICAM-1, soluble intracellular adhesion molecule 1; M-CSF, monocyte colony stimulating factor; CEA, carcinoembryonic antigen; IL, interleukin; OOB, out-of-bag.

first to report associations for this biomarker with locoregional metastatic progression in NSCLC. Given the resistance of stem cells to traditional lines of chemotherapy[52] and the apparent relationship circulating SCF has for detecting lymph node metastases, it would be of great interest to further investigate the role of stem cells in the establishment of secondary metastases in NSCLC as well as the risk for disease recurrence.

Our 6-analyte panel was able to accurately identify lymph node involvement in 88% of the NSCLC patients tested, relative to the gold-standard of pathologic staging. Recently, the integration of CT and PET has emerged as the most effective protocol for clinically staging patients preoperatively. Although the accuracy of integrated PET/CT scans to classify patients according to nodal status ranges considerably in the literature, recent comparisons (preoperative clinical staging versus intraoperative histopathological staging) provided figures in the range of 78 to 90%. They have a dramatic loss in accuracy and specificity with nodes ≤1 cm or relevant comorbidities.[39,40] We were able to correctly identify lymph node metastases in 85% of the patients enrolled in this study, which agrees favorably with the documented accuracy rates mentioned above. Although the errors for clinically evaluating our patients were equally divided between up- and down-staging, we observed a strong trend for our 'molecular staging' method to 'up-stage' the patients (i.e., predict a positive nodal status in a pathologically-determined node negative case). No obvious trend towards disease recurrence was observed in these misclassified patients within the framework of the available clinical follow-up data. Such an observation would clearly have been highly suggestive of the presence of occult metastases at the time of surgery. Interestingly, only a single misclassification out of the total patient population enrolled in this study was observed by both imaging and blood-based methods. This observation is highly suggestive that the molecular and clinical staging methods may be complementary, and the combination of a blood-based assay with PET/CT imaging may show the most promise for identification of patients who are candidates for systemic therapy.

Several major concerns are typically raised during the development of multianalyte diagnostic tests that man-

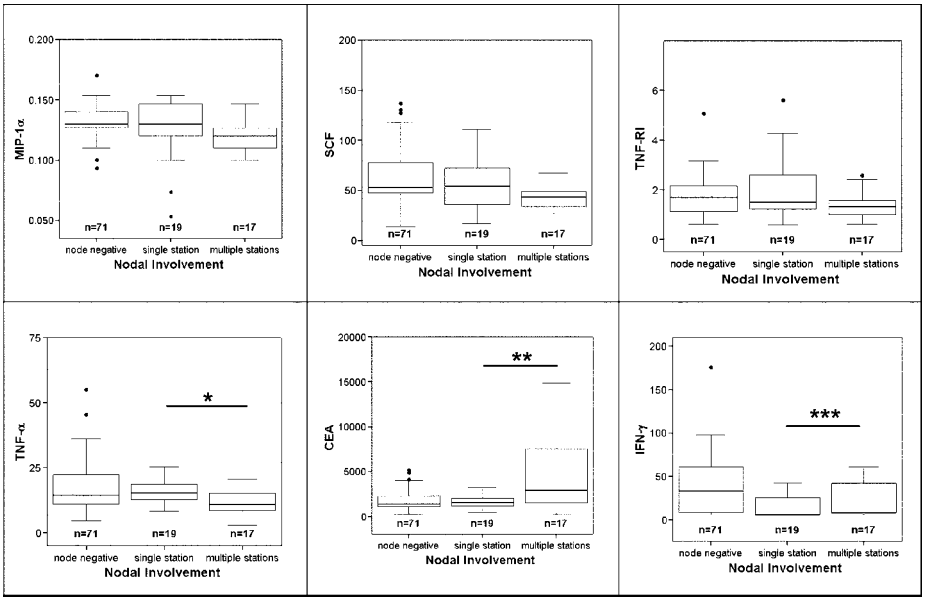

FIGURE 4. Boxplots of the 6-analyte biomarker panel for identifying preoperative nodal status with respect to extent of nodal station involvement. Boxplots for the 6-analytes comprising our most optimal panel for predicting lymph node involvement are represented with the node positive subjects subdivided according to either single or multiple nodal station involvement as a surrogate for total nodal involvement. (Mann-Whitney *p*-values: *, $p = 0.011$; **, $p = 0.07$; **, $p = 0.012$).

date a series of additional validation steps to suitably appraise the value of the test before clinical application. Recently, Brenner et al.[53] defined these factors as the following: (1) over-fitting the data with multiparametric statistical models, (2) bias from inappropriate patient selection, and (3) bias inherent to the specific patient population being studied that is not generally true across the total population. We were careful to obviate concerns 1 and 2 by cross-validating our results across several multiparametric statistical models (as defined in the Results section) and relying on the clinical expertise of our research team to select an appropriate and representative patient population for this study. To address concern 3, Pepe et al.[54] proposed a five-phase validation sequence that specifically defines the aims, specimen selection, and primary outcome measures necessary in each phase of biomarker development. This validation scheme has been adopted by the Early Detection Research Network and, indeed, is required for this diagnostic tool to progress towards clinical application. Specifically, we are planning additional validation against a much larger serum cohort that has been accrued by the NCI-sponsored 'Cancer and Leukemia Group B.'

Once our multianalyte serum biomarker panel has been properly validated against a second, independent serum cohort, we expect that one of the principal outcomes will be an improved modality for the preoperative selection of patients requiring systemic treatment. This may be particularly valuable given that neoadjuvant chemotherapy is generally better tolerated than adjuvant (postoperative) chemotherapy. Therefore, with more accurate staging methods, patients would receive more appropriately administered chemotherapy, thereby decreasing the overall morbidity of NSCLC treatments. In the future, we anticipate developing a series of multianalyte blood tests that will be capable of identifying specific mechanistic details of disease progression and allow for the individualization of patient therapies with existing targeted or multimodal therapies. In addition to providing more effective and better tolerated treatment options for patients, this paradigm is also expected to lead to the development of new targeted therapies through an improved understanding of the biology of the disease. We predict this insight will allow us to more effectively treat patients with NSCLC and, thereby improve survival.

ACKNOWLEDGMENTS

This work was supported by a generous donation from the Sapiente family to the Rush Thoracic Oncology Group and new investigator start-up funds provided by the Falk Foundation (JB).

The authors would like to thank Ms. Elizabeth Avery, for her assistance in the compilation of the data sets for statistical analysis, and Erin C. Farlow, MD, for her editorial assistance in the preparation of this manuscript.

REFERENCES

1. Jemal A, Siegel R, Ward E, et al. Cancer statistics, 2007. *CA Cancer J Clin* 2007;57:43–66.
2. Lu C, Soria JC, Tang X, et al. Prognostic factors in resected stage I non-small-cell lung cancer: a multivariate analysis of six molecular markers. *J Clin Oncol* 2004;22:4575–4583.
3. Sihoe AD, Yim AP. Lung cancer staging. *J Surg Res* 2004;117:92–106.
4. Ou SH, Zell JA, Ziogas A, et al. Prognostic factors for survival of stage I nonsmall cell lung cancer patients: a population-based analysis of 19,702 stage I patients in the California Cancer Registry from 1989 to 2003. *Cancer* 2007;110:1532–1541.
5. Pignon JP, Tribodet H, Scagliotti GV, et al. Lung adjuvant cisplatin evaluation: a pooled analysis by the LACE Collaborative Group. *J Clin Oncol* 2008;26:3552–3559.
6. Choi N, Son DS, Lee J, et al. The signature from messenger RNA expression profiling can predict lymph node metastasis with high accuracy for non-small cell lung cancer. *J Thorac Oncol* 2006;1:622–628.
7. Kikuchi T, Daigo Y, Katagiri T, et al. Expression profiles of non-small cell lung cancers on cDNA microarrays: identification of genes for prediction of lymph-node metastasis and sensitivity to anti-cancer drugs. *Oncogene* 2003;22:2192–2205.
8. Xi L, Coello MC, Litle VR, et al. A combination of molecular markers accurately detects lymph node metastasis in non-small cell lung cancer patients. *Clin Cancer Res* 2006;12:2484–2491.
9. Xi L, Lyons-Weiler J, Coello MC, et al. Prediction of lymph node metastasis by analysis of gene expression profiles in primary lung adenocarcinomas. *Clin Cancer Res* 2005;11:4128–4135.
10. Chen G, Gharib TG, Huang CC, et al. Proteomic analysis of lung adenocarcinoma: identification of a highly expressed set of proteins in tumors. *Clin Cancer Res* 2002;8:2298–2305.
11. Singhal S, Vachani A, Antin-Ozerkis D, et al. Prognostic implications of cell cycle, apoptosis, and angiogenesis biomarkers in non-small cell lung cancer: a review. *Clin Cancer Res* 2005;11:3974–3986.
12. Lancashire LJ, Rees RC, Ball GR. Identification of gene transcript signatures predictive for estrogen receptor and lymph node status using a stepwise forward selection artificial neural network modelling approach. *Artif Intell Med* 2008;43:99–111.
13. Seike M, Yanaihara N, Bowman ED, et al. Use of a cytokine gene expression signature in lung adenocarcinoma and the surrounding tissue as a prognostic classifier. *J Natl Cancer Inst* 2007;99:1257–1269.
14. Gorelik E, Landsittel DP, Marrangoni AM, et al. Multiplexed immunobead-based cytokine profiling for early detection of ovarian cancer. *Cancer Epidemiol Biomarkers Prev* 2005;14:981–987.
15. McIntosh MW, Pepe MS. Combining several screening tests: optimality of the risk score. *Biometrics* 2002;58:657–664.
16. Yurkovetsky ZR, Kirkwood JM, Edington HD, et al. Multiplex analysis of serum cytokines in melanoma patients treated with interferon-alpha2b. *Clin Cancer Res* 2007;13:2422–2428.
17. Yurkovetsky ZR, Linkov FY, E Malehorn D, Lokshin AE. Multiple biomarker panels for early detection of ovarian cancer. *Future Oncol* 2006;2:733–741.
18. Okada M, Nishio W, Sakamoto T, et al. Prognostic significance of perioperative serum carcinoembryonic antigen in non-small cell lung cancer: analysis of 1,000 consecutive resections for clinical stage I disease. *Ann Thorac Surg* 2004;78:216–221.
19. Pujol JL, Molinier O, Ebert W, et al. CYFRA 21–1 is a prognostic determinant in non-small-cell lung cancer: results of a meta-analysis in 2063 patients. *Br J Cancer* 2004;90:2097–2105.
20. Takamochi K, Nagai K, Suzuki K, et al. Clinical predictors of N2 disease in non-small cell lung cancer. *Chest* 2000;117:1577–1582.
21. Tomita M, Matsuzaki Y, Shimizu T, et al. Preoperative prognostic factors for pN2 non-small cell lung cancer. *Ann Thorac Cardiovasc Surg* 2006;12:15–20.
22. Bates J, Rutherford R, Divilly M, et al. Clinical value of CYFRA 21.1, carcinoembryonic antigen, neurone-specific enolase, tissue polypeptide specific antigen and tissue polypeptide antigen in the diagnosis of lung cancer. *Eur Respir J* 1997;10:2535–2538.
23. Hatzakis KD, Froudarakis ME, Bouros D, et al. Prognostic value of serum tumor markers in patients with lung cancer. *Respiration* 2002;69:25–29.
24. Kaminska J, Kowalska M, Kotowicz B, et al. Pretreatment serum levels of cytokines and cytokine receptors in patients with non-small cell lung cancer, and correlations with clinicopathological features and prognosis. M-CSF - an independent prognostic factor. *Oncology* 2006;70:115–125.
25. Molina R, Filella X, Auge JM, et al. Tumor markers (CEA, CA 125, CYFRA 21–1, SCC and NSE) in patients with non-small cell lung cancer as an aid in histological diagnosis and prognosis. Comparison

with the main clinical and pathological prognostic factors. *Tumour Biol* 2003;24:209–218.
26. Patz EF Jr, Campa MJ, Gottlin EB, et al. Panel of serum biomarkers for the diagnosis of lung cancer. *J Clin Oncol* 2007;25:5578–5583.
27. Linkov F, Lisovich A, Yurkovetsky Z, et al. Early detection of head and neck cancer: development of a novel screening tool using multiplexed immunobead-based biomarker profiling. *Cancer Epidemiol Biomarkers Prev* 2007;16:102–107.
28. Boldrini L, Donati V, Dell'Omodarme M, et al. Prognostic significance of osteopontin expression in early-stage non-small-cell lung cancer. *Br J Cancer* 2005;93:453–457.
29. D'Amico TA, Brooks KR, Joshi MB, et al. Serum protein expression predicts recurrence in patients with early-stage lung cancer after resection. *Ann Thorac Surg* 2006;81:1982–1987; discussion 1987.
30. Huang C, Liu D, Masuya D, et al. Clinical application of biological markers for treatments of resectable non-small-cell lung cancers. *Br J Cancer* 2005;92:1231–1239.
31. Kaya A, Ciledag A, Gulbay BE, et al. The prognostic significance of vascular endothelial growth factor levels in sera of non-small cell lung cancer patients. *Respir Med* 2004;98:632–636.
32. Neuner A, Schindel M, Wildenberg U, et al. Prognostic significance of cytokine modulation in non-small cell lung cancer. *Int J Cancer* 2002;101:287–292.
33. Onn A, Correa AM, Gilcrease M, et al. Synchronous overexpression of epidermal growth factor receptor and HER2-neu protein is a predictor of poor outcome in patients with stage I non-small cell lung cancer. *Clin Cancer Res* 2004;10:136–143.
34. Tarro G, Perna A, Esposito C. Early diagnosis of lung cancer by detection of tumor liberated protein. *J Cell Physiol* 2005;203:1–5.
35. Vielh P, Spano JP, Grenier J, et al. Molecular prognostic factors in resectable non-small cell lung cancer. *Crit Rev Oncol Hematol* 2005;53:193–197.
36. Breiman L. Random Forests. *Mach Learn* 2001;45:5–32.
37. Brieman L, Friedman J, Olshen R, Stone C. Classification and Regression Trees. Belmont, California: Wadsworth Co, 1984.
38. Team RDC. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing, 2006.
39. Al-Sarraf N, Gately K, Lucey J, Wilson L, McGovern E, Young V. Lymph node staging by means of positron emission tomography is less accurate in non-small cell lung cancer patients with enlarged lymph nodes: analysis of 1145 lymph nodes. *Lung Cancer* 2008;60:62–68.
40. Yang W, Fu Z, Yu J, et al. Value of PET/CT versus enhanced CT for locoregional lymph nodes in non-small cell lung cancer. *Lung Cancer* 2008;61:35–43.
41. Mountain CF. Staging classification of lung cancer. A critical evaluation. *Clin Chest Med* 2002;23:103–121.
42. Solomon B, Bunn PA Jr. Adjuvant chemotherapy for non-small cell lung cancer. *Cancer Invest* 2007;25:217–225.
43. Waterston A, Bower M. TNF and cancer: good or bad? *Cancer Therapy* 2004;2:131–148.
44. Hatate K, Yamashita K, Hirai K, et al. Liver metastasis of colorectal cancer by protein-tyrosine phosphatase type 4A, 3 (PRL-3) is mediated through lymph node metastasis and elevated serum tumor markers such as CEA and CA19–9. *Oncol Rep* 2008;20:737–743.
45. Levy M, Visokai V, Lipska L, et al. Tumor markers in staging and prognosis of colorectal carcinoma. *Neoplasma* 2008;55:138–142.
46. Hwa HL, Kuo WH, Chang LY, et al. Prediction of breast cancer and lymph node metastatic status with tumour markers using logistic regression models. *J Eval Clin Pract* 2008;14:275–280.
47. Liao Q, Zhao L, Chen X, et al. Serum proteome analysis for profiling protein markers associated with carcinogenesis and lymph node metastasis in nasopharyngeal carcinoma. *Clin Exp Metastasis* 2008;25:465–476.
48. Moran CJ, Arenberg DA, Huang CC, et al. RANTES expression is a predictor of survival in stage I lung adenocarcinoma. *Clin Cancer Res* 2002;8:3803–3812.
49. Kanazawa S, Yamaguchi K, Kinoshita Y, et al. Elevation of soluble interleukin-2 receptor in patients with non-small cell lung cancer treated with gefitinib. *J Cancer Res Clin Oncol* 2006;132:719–725.
50. Donnenberg VS, Landreneau RJ, Donnenberg AD. Tumorigenic stem and progenitor cells: implications for the therapeutic index of anti-cancer agents. *J Control Release* 2007;122:385–391.
51. Mroczko B, Szmitkowski M, Czygier M. [Stem cell factor (SCF) in diagnosis and monitoring of non-small-cell lung cancer]. *Pol Arch Med Wewn* 1999;101:213–218.
52. Vinoth KJ, Heng BC, Poonepalli A, et al. Human embryonic stem cells may display higher resistance to genotoxic stress as compared to primary explanted somatic cells. *Stem Cells Dev* 2008;17:599–607.
53. Brenner DE, Normolle DP. Biomarkers for cancer risk, early detection, and prognosis: the validation conundrum. *Cancer Epidemiol Biomarkers Prev* 2007;16:1918–1920.
54. Pepe MS, Etzioni R, Feng Z, et al. Phases of biomarker development for early detection of cancer. *J Natl Cancer Inst* 2001;93:1054–1061.

