

A Multigene Assay Is Prognostic of Survival in Patients with Early-Stage Lung Adenocarcinoma

Dan J. Raz,¹ M. Roshni Ray,¹ Jae Y. Kim,¹ Biao He,¹ Miquel Taron,³ Marcin Skrzypski,⁴ Mark Segal,² David R. Gandara,⁵ Rafael Rosell,³ and David M. Jablons¹

Abstract Purpose: Clinical staging does not adequately risk stratify patients with early stage non-small cell lung cancer. We sought to generate a real-time PCR (RT-PCR)-based prognostic model in patients with early stage lung adenocarcinoma, the dominant histology of lung cancer in the United States.

Experimental Design: We studied gene expression of 61 candidate genes in 107 patients with completely surgically resected lung adenocarcinoma using RT-PCR. We used crossvalidation methods to select and validate a prognostic model based on the expression of a limited number of genes. A risk score was generated based on model coefficients, and survival of patients with high- and low-risk scores were analyzed.

Results: We generated a four-gene model based on expression of *WNT3a*, *ERBB3*, *LCK*, and *RND3*. Risk score predicted mortality better than clinical stage or tumor size (adjusted hazard ratio, 6.7; 95% confidence interval, 1.6-28.9; $P = 0.001$). Among 70 patients with stage I disease, 5-year overall survival was 87% among patients with low-risk scores, and 38% among patients with high-risk scores ($P = 0.0002$). Among all patients, 5-year overall survival was 62% and 41%, respectively ($P = 0.0054$). Disease-free survival was also significantly different among low- and high-risk score patients.

Conclusions: This multigene assay predicts overall and disease-free survival significantly better than clinical stage and tumor size in patients with early stage lung adenocarcinoma and performs especially well in patients with stage I disease. Prospective clinical trials are needed to determine whether high-risk patients with stage I disease benefit from adjuvant chemotherapy.

Lung cancer is the most common cause of death in the United States and worldwide (1). Despite recent advances, long-term survival remains poor, with no >15% of patients still alive at 5 years after diagnosis. Despite improved understanding of the molecular biology of lung cancer, treatment decisions continue to be guided largely by clinical stage, designated as stages I through IV. Even in the most curable subset of non-small cell lung cancer (NSCLC), stage I disease, 5-year survival averages no >70% (2, 3). Although recent randomized clinical trials have shown improved survival with the use of postoperative

adjuvant chemotherapy in stages II and III NSCLC, they have failed to detect a benefit for adjuvant chemotherapy for patients with stage I disease (4, 5). However, ongoing efforts to identify prognostic genomic biomarkers for risk stratification and predictive markers of therapeutic benefit offer considerable promise to alter current decision making paradigms.

A number of prognostic biomarkers have been proposed for NSCLC, but none has yet been successfully translated into clinical application (6, 7). Although several genome-wide expression microarray-based prognostic models of lung cancer have been reported, such array-based technology may be suboptimal for clinical use due to the need for specialized laboratory facilities and complex statistical analyses (8-12). Prognostic models based on gene expression of a limited number of genes using quantitative real-time PCR (RT-PCR) may be more clinically practical. RT-PCR, considered the gold standard for measurement of gene expression, is highly reproducible and is relatively simple to analyze (13). Two recently reported RT-PCR-based prognostic models of lung cancer were able to risk stratify patients with resected NSCLC, although it is unclear whether they predicted outcome better than clinical factors, such as tumor size, and whether prognostic information differed by NSCLC histology (14, 15).

We set out to generate a prognostic risk score for survival in patients with completely resected lung adenocarcinoma based on genes previously identified in microarray models of prognosis in NSCLC. As different histologic subtypes of NSCLC

Authors' Affiliations: ¹Thoracic Oncology Program, Department of Surgery and ²Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, California; ³Catalan Institute of Oncology, Hospital Germans Trias I Pujol, Barcelona, Spain; ⁴Medical University of Gdansk, Gdansk, Poland; and ⁵Division of Hematology/Oncology, University of California Davis Cancer Center, Sacramento, California

Received 3/3/08; revised 4/22/08; accepted 5/8/08.

Grant support: Private foundation (anonymous donor).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Note: Supplementary data for this article are available at Clinical Cancer Research Online (<http://clincancerres.aacrjournals.org/>).

Requests for reprints: Dan J. Raz, University of California, San Francisco, 513 Parnassus Avenue S-321, San Francisco, CA 94131. Phone: 415-476-1239; Fax: 415-353-9530; E-mail: Dan.raz@ucsf.edu.

©2008 American Association for Cancer Research.

doi:10.1158/1078-0432.CCR-08-0544

are known to have different patterns of gene expression, we chose to restrict our study to patients with adenocarcinoma, the dominant histology of NSCLC in the United States. We planned to separately analyze patients with stage I disease, as we hypothesized that risk among these patients would be stratified best by molecular modeling compared with more advanced stage disease. Moreover, patients with stage I disease who are at high risk for recurrence and death may benefit from adjuvant therapies. Finally, we also sought to compare our model to the five-gene prognostic model recently reported by Chen and colleagues (14).

Materials and Methods

Patients. Between January 1997 and June 2004, 120 patients who had undergone complete surgical resection of lung adenocarcinoma without preoperative chemotherapy or radiation treatment at the University of California San Francisco, and had fresh-frozen tissue banked for genomic analysis, were entered into the study. Eligible patients had undergone surgical resection with curative intent and had adequate mediastinal lymph node staging. Thirteen patients were excluded due to insufficient banked tissue, inadequate RNA quality, or weak expression of housekeeping genes. We did not perform a formal power calculation for this study as our sample size was limited by the availability of banked tissue at our institution and our inclusion criteria, the available sample size was on par with prior molecular marker prognostic studies. Information on clinical variables and patient follow-up were obtained from a prospectively maintained database including all subjects with banked tissue in the study. The primary end point was overall survival (OS). Vital status and date of death was determined by querying the Social Security Death Index using the subject's social security number (available online at Social Security Death Index).⁶ Disease-free survival was defined as the time from surgery until radiographic evidence of recurrent disease or time until the last documented physician follow-up visit in the absence of recurrent disease. Patients consented to tissue specimen collection prospectively, and the study was approved by the University of California, San Francisco, institutional review board (CHR# H8714-28880-01).

Gene selection. Of 217 genes identified from previously published microarray and PCR-based studies of prognosis in early stage lung cancer (8, 10, 12, 16–18), 76 cancer-related genes were identified by study investigators or by literature review. Fifteen genes were excluded due to nonfunctioning primers or very weak expression in tumor tissue by RT-PCR in a pilot study. The remaining 61 genes are listed in Supplementary table.

Sample preparation and analysis. All tissue was frozen in liquid nitrogen at the time of the operation. Tissue was macrodissected into 1 cm³ sections that was ground in liquid nitrogen. RNA was extracted using a TriZol (Invitrogen) extraction protocol. Taqman RT-PCR was done on cDNA in 384-well plates using Prism 7900HT machine (Applied Biosystems). The expression of each gene was assayed in triplicate. Samples were compared with commercially available pooled normal lung RNA (Clontech), and normalized to 18S ribosomal rRNA (Applied Biosystems). Detailed methods on sample preparation and analysis are described in Supplementary methods.

Statistical Analysis. The salient features of our data structure are (a) a right censored survival end point (death) with modest event numbers (47—after subsetting to exclude missings); (b) a multitude (61) of predictors as constituted by the (log) gene expression values obtained from real-time quantitative PCR ($\delta\text{-}\delta$ Cts described elsewhere); and (c) select clinical and demographic covariates. In view of

these features and dimensions, the primary data analytic tool we used was L1 penalized Cox proportional hazards regression (19). This methodology extends the simultaneous coefficient shrinkage and predictor selection that is inherent in L1 penalization (20), where it has proven highly effective (21, 22), to the survival data setting. Earlier extensions, largely motivated by microarray gene expression applications, were either computationally prohibitive (23) or reliant on approximation (24). The tradeoff between model fit, as captured via the Cox partial likelihood, and model complexity [number of included predictors (genes)], as captured via the L1 penalty, is governed by a tuning parameter that weights the contribution of the penalty. Determination of the value of this parameter makes recourse to crossvalidation, as is detailed elsewhere (19, 24). Briefly, the tuning parameter value that achieves optimal fit, as evaluated over unseen data furnished by the crossvalidation data partitioning approach, is selected. This choice then determines which and how many genes to retain. A risk score was then generated for each subject based on model coefficients. Resultant predicted risk scores were dichotomized (at the median), and corresponding Kaplan-Meier survival curves were displayed and compared via the log-rank statistic. All analyses were conducted using the programming language R (Version 2.3.1, 2006; ref. 25).

Results

Clinical and histologic characteristics of patients are listed in Table 1. We first analyzed the entire data set and subsequently restricted the analysis to stage I patients. Crossvalidation, after forced adjustment for age, tumor size, and stage, supported a model containing four genes, selected in the following order:

Table 1. Patient characteristics, test group (N = 107)

	n (%)
Age, years (mean, SD)	68 ± 11
Sex	
Women	65 (61%)
Men	42 (39)
Smoking Status	
Current	26 (24%)
Past	61 (57%)
Never	20 (19%)
Histology	
Pure adenocarcinoma	64 (60%)
Pure BAC	14 (13%)
Mixed adenocarcinoma and BAC	29 (27%)
Race/ethnicity	
Caucasian	77 (72%)
Asian	19 (18%)
African-American	5 (5%)
Hispanic	2 (2%)
Other	4 (4%)
Tumor grade	
Well-differentiated	32 (30%)
Moderately differentiated	46 (43%)
Poorly differentiated	28 (26%)
Tumor size (mean in cm, SD)	3.5 ± 2.0
Cancer stage	
I	70 (66%)
II	12 (11%)
IIIA	12 (11%)
IIIB/IV	13 (12%)
Follow up time among surviving patients, months [median (range)]	61 (36-114)

Abbreviation: BAC, bronchioloalveolar carcinoma.

⁶ <http://ssdi.rootsweb.com>

Table 2. Adjusted Cox regression analysis

	HR	95% CI	P
Risk score*	6.72	1.56-28.91	0.01
Stage [†]	1.32	1.05-1.65	0.02
Tumor size	1.04	0.91-1.19	0.56
Age	1.01	0.98-1.03	0.54

Abbreviations: HR, hazard ratio; CI, confidence interval.
 *Risk score analyzed as a continuous variable.
[†]Hazard ratio for stage is stage I patients compared with all other patients.

WNT3A, *RND3*, *LCK*, and *ERBB3*. OS was analyzed using the risk score for the model using these four genes (with L1 shrunken coefficients). After adjusting for patient age, disease stage, and tumor size, the hazard ratio for the risk score as a continuous variable was 6.7 (95% confidence interval, 1.6-28.9; $P = 0.005$), as listed in Table 2. Other clinical factors, including sex, smoking status, and race/ethnicity, had weak associations with survival and were not included in the multivariate analysis.

Survival curves for the entire study cohort based on dichotomized risk scores from the four-gene model are presented in Fig. 1. Corresponding survival curves for dichotomized risk scores among 70 stage I patients are presented in Fig. 2. Five-year OS was 56% for all patients and 65% for stage I patients. Among all patients, 5-year OS was 62% among low-risk patients and 41% among high-risk patients ($P = 0.0054$). Five-year disease-free survival was 60% among low-risk patients and 28% among high-risk patients ($P = 0.0053$). Among stage I patients, 5-year OS was 87% among low-risk patients and 38% among high-risk patients ($P = 0.0002$), and disease-free survival was 77% among low-risk patients and 35% among high-risk patients ($P = 0.0045$).

Applying the 5-gene model of Chen and colleagues (14) to our study population, the hazard ratio adjusted for age, stage, and tumor size was 2.4 (95% confidence interval, 1.3-4.3; $P = 0.0047$). Among all patients, 5-year OS was 65% among low-risk patients and 43% among high-risk patients ($P = 0.045$). Among stage I patients, 5-year OS was 80% among low-risk patients and 47% among high-risk patients ($P = 0.022$), Fig. 3.

Discussion

We identified a risk score based on the gene expression of four genes by quantitative PCR that is prognostic of long-term survival in patients with completely surgically resected lung adenocarcinoma. This gene expression-based score predicted survival (adjusted hazard ratio, 6.7) and disease recurrence better than clinical stage and tumor size. Our model best predicted risk of death among patients with stage I disease. In our crossvalidated cohort, low- and high-risk scores were associated with 5-year OS of 87% and 38% among stage I patients.

We used crossvalidation both for model selection and for model validation. Rather than splitting data into test and validation sets, crossvalidation uses repeated data-splitting to

prevent model overfitting and to generate accurate estimates of model coefficients. Crossvalidation is operationally equivalent to data splitting with respect to generation of validated model coefficients, but it uses data more efficiently than simple data splitting, making it a compelling statistical technique for both model selection and validation (26).

This prognostic model may have important clinical implications, as a tool to complement clinical staging in risk-stratifying patients with stage I lung adenocarcinoma. The current standard of care for patients with stage I NSCLC is surgical resection alone, typically lobectomy. Identifying patients with stage I NSCLC who are at high risk for disease recurrence from clinically undetectable micrometastases may provide the basis for further study of adjuvant chemotherapy or targeted therapies in this patient population. Conversely, postresection patients at low risk for recurrence may be best served by observation alone, and may even be candidates for less radical lung resection. Because our findings relate to prognostic value only, prospective clinical trials will be needed to evaluate whether high-risk stage I patients identified by this assay benefit

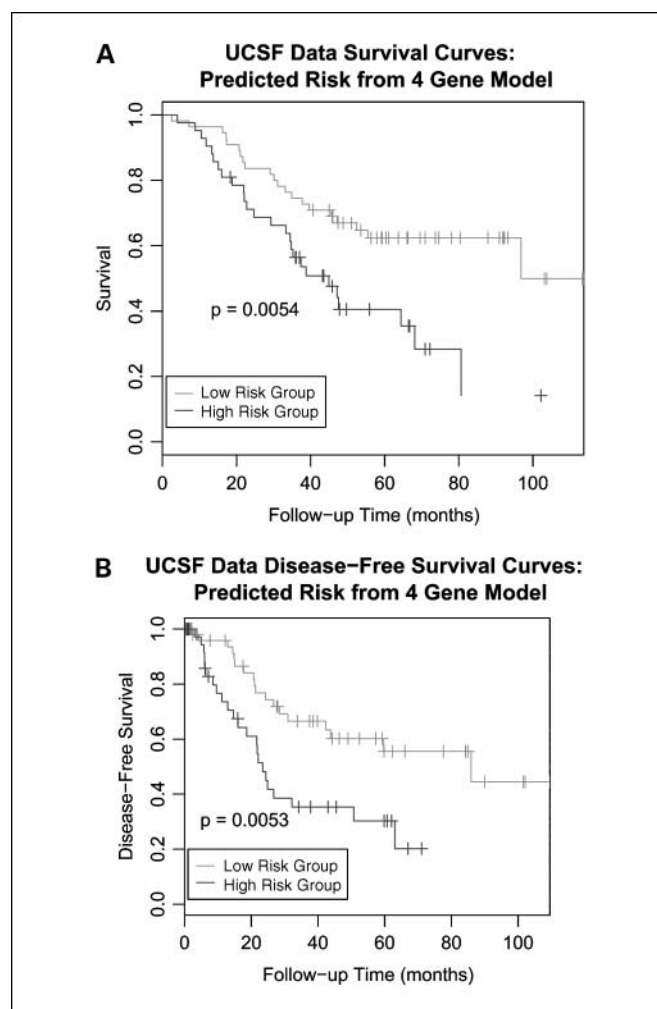


Fig. 1. A, Kaplan-Meier curves comparing OS among patients with low- and high-risk (all stages) using the four-gene prognostic model. B, Kaplan-Meier curves comparing disease-free survival among patients with low- and high-risk scores (all stages).

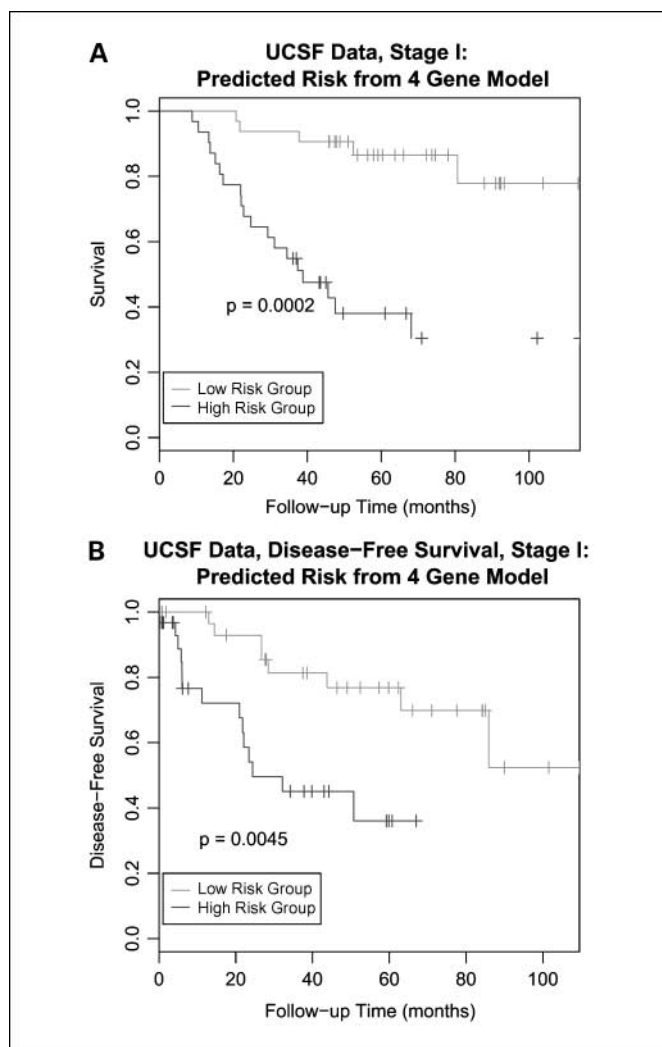


Fig. 2. A, Kaplan-Meier curves comparing OS among stage I patients with low- and high-risk scores using the four-gene prognostic model. B, Kaplan-Meier curves comparing disease-free survival among stage I patients with low- and high-risk scores.

from adjuvant chemotherapy, i.e., to determine the predictive value of this assay.

The genes included in this model have been previously implicated in cancer biology. Two of the genes, *ERBB3* (*v-erb-b2 erythroblastic leukemia viral oncogene homologue 3*) and *LCK* (*lymphocyte-specific protein tyrosine kinase*), overlap with the previously published five-gene RT-PCR model of Chen et al. (14) in resected NSCLC. *ERBB3*, a receptor tyrosine kinase, has been associated with an increased risk of distant metastases in patients with early stage NSCLC (27). *LCK*, also a tyrosine kinase, is involved in regulating apoptosis and regulation of chemotaxis in lymphoma (28). *Wnt3a* (*wingless-type mouse mammary tumor virus integration site family, member 3a*), a proto-oncogene that activates the *Wnt* pathway, is a critical pathway in lung carcinogenesis and progression (29). *RND3* (*Rho family GTPase 3*) is a tumor suppressor with an important role in the regulation of cell proliferation and is down-regulated in prostate cancer development (30). As previously mentioned, the initial 76 study genes were all cancer-related genes. This selection was necessary to make the study feasible. This may be

viewed as a limitation in that novel cancer genes may have been excluded. However, selection based on gene function adds biological plausibility to the gene selection and may decrease the probability of selecting genes that do not have a true prognostic association.

In our patient population, the prognostic value of our model compares favorably to that of Chen and colleagues (14). There are several possible explanations for this finding. First, our model was fit using our data, which raises the concern that model performance was artificially superior to that of Chen and colleagues (14). Yet the use of cross-validation should greatly reduce or eliminate model “over-fitting” and should allow relatively unbiased comparison of models. Next, candidate genes for our model were selected based on previously published genome-wide expression microarray studies (8–12, 16, 31), whereas genes in the Chen model were selected from a more limited custom expression array platform. Additionally, a relatively large set of genes were assessed by RT-PCR before finalizing our gene model. Moreover, our patient follow-up time was relatively long, with a minimum follow-up period of 3 years and a median follow-up of 61 months. This factor may have allowed us to more accurately detect differences in survival between risk groups. We did not test the three-gene model recently reported by Lau and colleagues (15) in our samples, as our study was completed before that data were reported. Nevertheless, the survival curves and adjusted hazard ratio for death from this study compare favorably to Lau’s model, especially in stage I patients. Direct comparison in an independent set of samples is necessary.

We also did not directly compare our RT-PCR-based model with any microarray-based study. A limited RT-PCR-based model has the advantage of being more reproducible and more feasible in the clinical setting than microarray-based models. We found that hardly any genes identified as being prognostic of survival in previously published microarray-based studies of lung cancer overlapped between studies. This implies that many genes may predict outcome in parallel. We sought to identify

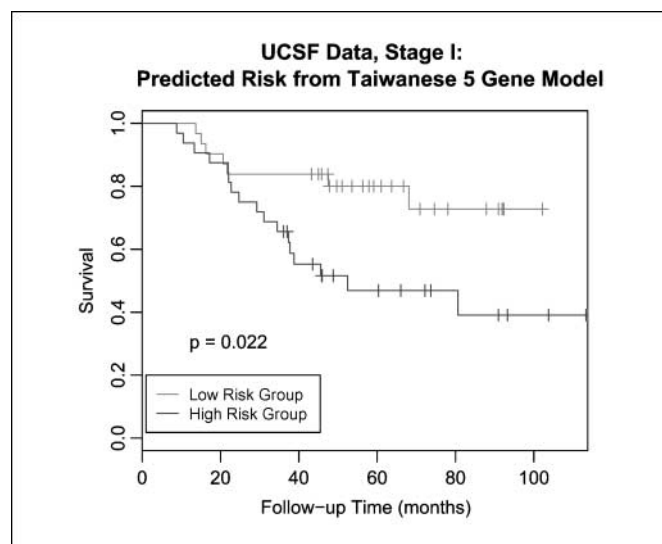


Fig. 3. Kaplan-Meier curves comparing stage I patients with low- and high-risk scores using the 5-gene model published by Chen and colleagues (14).

the sparsest model that still had excellent prognostic discrimination. Independent studies comparing our model to other genomic-based prognostic models in patients with early stage lung adenocarcinoma are needed.

Our study has several limitations. First, we used tumor tissue frozen in liquid nitrogen at the time of surgical resection so as to extract RNA of the highest quality for measurement of gene expression. Although the use of fresh-frozen tissue maximized the internal validity of expression levels, whether our results are generalizable to formalin-fixed tissue and cytology specimens remains to be determined. A pilot study comparing our multigene assay on formalin-fixed tissue specimens from subjects of this study to the assay results on fresh-frozen tissue is currently under way. In addition, we relied on macrodissection of tumor tissue by one of the study investigators. Whereas tumor samples were reviewed by a pathologist before tissue banking, the specific piece of tumor used for RNA extraction was not. Nonetheless, in our experience, it is extremely rare for macrodissection of banked tissue to yield tumor tissue sections with <70% tumor cells.

Next, although the sample size in this study was comparable with other recently published prognostic genomic models, we plan to prospectively study the expression of these four genes in a larger cohort of stage I patients so that multiple strata of risk can be more accurately estimated based on risk score.

Finally, this study was limited to patients with adenocarcinoma. We chose to restrict our study to this histologic subtype to minimize heterogeneity because histologic subtypes of NSCLC are associated with different molecular profiles and because adenocarcinoma has become the dominant NSCLC histology in the United States over the last 25 years (28). As tumor biology differs not only by histology, but also by clinical features such as ethnicity and smoking status, it is possible that the risk stratification capability of prognostic genomic models will vary by clinical subgroup. Although we addressed this issue in part by restricting our subjects to adenocarcinoma histology, our sample size was too small to

have a meaningful assessment of the effect of ethnicity and smoking status on model accuracy.

We found that our model performed better than clinical variables known to be important for risk, including clinical stage, tumor size, and patient age. A clinical/pathologic risk score for early stage NSCLC, incorporating clinical and pathologic variables commonly collected, is needed as a benchmark for genomic models of outcome. A genomic-based model that does not predict long-term survival more accurately than clinical factors is unlikely to be useful in clinical practice. Ultimately, an independent comparison of RT-PCR-based models, microarray-based models, and such a clinical/pathologic risk model is needed.

Ideally, a prognostic tool should provide accurate risk stratification, should be clinically feasible to use in day-to-day practice, and should be cost effective. Such an assay would be of particular benefit to patients with surgically resected stage I NSCLC. The current standard of care for stage I NSCLC is lobectomy and mediastinal lymph node dissection, without adjuvant chemotherapy. Better identification of good prognosis patient subsets might allow lesser surgical procedures to be used with equal survival potential. Conversely, stage I subsets with a poor prognosis could be selected for inclusion into clinical trials testing novel approaches and new therapeutic agents. Considering the current limitations of chemotherapy in stage I disease, a bioassay that is both prognostic and predictive of chemotherapy benefit would be especially beneficial. Lastly, stage I NSCLC is likely to be of increasing importance in the future. Although ~20% of patients currently diagnosed with NSCLC are stage I, this proportion probably will grow due to the recent advent of lung cancer screening by computerized tomography (32).

Disclosure of Potential Conflicts of Interest

D.J. Raz, B. He, D.R. Gandara, R. Rosell, and D.M. Jablons have a minor ownership interest in Pinpoint Graphics.

References

- Jemal A, Siegel R, Ward E, et al. Cancer statistics, 2006. *CA Cancer J Clin* 2006;56:106–30.
- Dominioni L, Imperatori A, Rovera F, Ochetti A, Torrioni G, Paolucci M. Stage I nonsmall cell lung carcinoma: analysis of survival and implications for screening. *Cancer* 2000;89:2334–44.
- Harpole DH, Jr., Herndon JE II, Young WG, Jr., Wolfe WG, Sabiston DC, Jr. Stage I nonsmall cell lung cancer. A multivariate analysis of treatment methods and patterns of recurrence. *Cancer* 1995;76:787–96.
- Arriagada R, Bergman B, Dunant A, Le Chevalier T, Pignon JP, Vansteenkiste J. Cisplatin-based adjuvant chemotherapy in patients with completely resected non-small-cell lung cancer. *N Engl J Med* 2004;350:351–60.
- Winton T, Livingston R, Johnson D, et al. Vinorelbine plus cisplatin vs. observation in resected non-small-cell lung cancer. *N Engl J Med* 2005;352:2589–97.
- Thunnissen FB, Schuurbiens OC, den Bakker MA. A critical appraisal of prognostic and predictive factors for common lung cancers. *Histopathology* 2006;48:779–86.
- Brundage MD, Davies D, Mackillop WJ. Prognostic factors in non-small cell lung cancer: a decade of progress. *Chest* 2002;122:1037–57.
- Beer DG, Kardia SL, Huang CC, et al. Gene-expression profiles predict survival of patients with lung adenocarcinoma. *Nat Med* 2002;8:816–24.
- Bhattacharjee A, Richards WG, Staunton J, et al. Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. *Proc Natl Acad Sci U S A* 2001;98:13790–5.
- Potti A, Mukherjee S, Petersen R, et al. A genomic strategy to refine prognosis in early-stage non-small-cell lung cancer. *N Engl J Med* 2006;355:570–80.
- Raponi M, Zhang Y, Yu J, et al. Gene expression signatures for predicting prognosis of squamous cell and adenocarcinomas of the lung. *Cancer Res* 2006;66:7466–72.
- Wigle DA, Jurisica I, Radulovich N, et al. Molecular profiling of non-small cell lung cancer and correlation with disease-free survival. *Cancer Res* 2002;62:3005–8.
- Wang Y, Barbacioru C, Hyland F, et al. Large scale real-time PCR validation on gene expression measurements from two commercial long-oligonucleotide microarrays. *BMC Genomics* 2006;7:59.
- Chen HY, Yu SL, Chen CH, et al. A five-gene signature and clinical outcome in non-small-cell lung cancer. *N Engl J Med* 2007;356:11–20.
- Lau SK, Boutros PC, Pintiilie M, et al. Three-gene prognostic classifier for early-stage non small-cell lung cancer. *J Clin Oncol* 2007;25:5562–9.
- Garber ME, Troyanskaya OG, Schluens K, et al. Diversity of gene expression in adenocarcinoma of the lung. *Proc Natl Acad Sci U S A* 2001;98:13784–9.
- Miura K, Bowman ED, Simon R, et al. Laser capture microdissection and microarray expression analysis of lung adenocarcinoma reveals tobacco smoking- and prognosis-related molecular profiles. *Cancer Res* 2002;62:3244–50.
- Schneider PM, Praeuer HW, Stoeltzing O, et al. Multiple molecular marker testing (p53, C-Ki-ras, c-erbB-2) improves estimation of prognosis in potentially curative resected non-small cell lung cancer. *Br J Cancer* 2000;83:473–9.
- Park M, Hastie T. An L1 regularization-path algorithm for generalized linear models. *Journal of the Royal Statistical Society* 2007;69:659–77.
- Tibshirani R. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society* 2006;58:267–88.
- Efron B, Hastie T, Johnstone I, Tibshirani R. Least angle regression. *Annals of Statistics* 2004;32:407–99.
- Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning*. New York: Springer-Verlag; 2001.
- Gui J, Li H. Penalized Cox regression analysis in the high-dimensional and low-sample size settings, with applications to microarray gene expression data. *Bioinformatics* 2005;21:3001–8.
- Segal MR. Microarray gene expression data with

- linked survival phenotypes: diffuse large-B-cell lymphoma revisited. *Biostatistics* 2006;7:268–85.
25. Team RDC. R: A language and environment for statistical computing. 2.3.1 ed. Vienna, Austria: R Foundation for Statistical Computing; 2006.
26. Harrell FE, Jr., Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996;15:361–87.
27. Muller-Tidow C, Diederichs S, Bulk E, et al. Identification of metastasis-associated receptor tyrosine kinases in non-small cell lung cancer. *Cancer Res* 2005;65:1778–82.
28. Inngjerdingen M, Torgersen KM, Maghazachi AA. Lck is required for stromal cell-derived factor 1 α (CXCL12)-induced lymphoid cell chemotaxis. *Blood* 2002;99:4318–25.
29. Mazieres J, He B, You L, Xu Z, Jablons DM. Wnt signaling in lung cancer. *Cancer Lett* 2005;222:1–10.
30. Bektic J, Pfeil K, Berger AP, et al. Small G-protein RhoE is underexpressed in prostate cancer and induces cell cycle arrest and apoptosis. *Prostate* 2005;64:332–40.
31. McDoniels-Silvers AL, Stoner GD, Lubet RA, You M. Differential expression of critical cellular genes in human lung adenocarcinomas and squamous cell carcinomas in comparison to normal lung tissues. *Neoplasia* 2002;4:141–50.
32. Raz DJ, Glidden DV, Odisho AY, Jablons DM. Clinical characteristics and survival of patients with surgically resected incidentally detected lung cancer. *J Thorac Oncol* 2007;2:125–30.