Development and Validation of a Quantitative Real-Time Polymerase Chain Reaction Classifier for Lung Cancer Prognosis

Guoan Chen, MD, PhD,* Sinae Kim, PhD,† Jeremy M. G. Taylor, PhD,† Zhuwen Wang, MD,* Oliver Lee, MS,† Nithya Ramnath, MD,‡ Rishindra M. Reddy, MD,* Jules Lin, MD,* Andrew C. Chang, MD,* Mark B. Orringer, MD,* and David G. Beer, PhD*

**Introduction:** This prospective study aimed to develop a robust and clinically applicable method to identify patients with high-risk early-stage lung cancer and then to validate this method for use in future translational studies.

**Methods:** Three published Affymetrix microarray data sets representing 680 primary tumors were used in the survival-related gene selection procedure using clustering, Cox model, and random survival forest analysis. A final set of 91 genes was selected and tested as a predictor of survival using a quantitative real-time polymerase chain reaction-based assay using an independent cohort of 101 lung adenocarcinomas.

**Results:** The random survival forest model built from 91 genes in the training set predicted patient survival in an independent cohort of 101 lung adenocarcinomas, with a prediction error rate of 26.6%. The mortality risk index was significantly related to survival (Cox model \( p < 0.00001 \)) and separated all patients into low-, medium-, and high-risk groups (hazard ratio \( 1.00, 2.82, 4.42 \)). The mortality risk index was also related to survival in stage 1 patients (Cox model \( p = 0.001 \)), separating patients into low-, medium-, and high-risk groups (hazard ratio \( 1.00, 3.29, 3.77 \)).

**Conclusions:** The development and validation of this robust quantitative real-time polymerase chain reaction platform allows prediction of patient survival with early-stage lung cancer. Utilization will now allow investigators to evaluate it prospectively by incorporation into new clinical trials with the goal of personalized treatment of patients with lung cancer and improving patient survival.

**Key Words:** Lung cancer, qRT-PCR, Prognosis.

Lung cancer is the leading cause of cancer-related death, and non-small cell lung cancer (NSCLC) accounts for almost 80% of deaths from lung cancer. Therapy is the major treatment option for patients with early-stage (stages IA and IB) NSCLC, yet as many as 35 to 50% of these patients will relapse within 5 years, indicating that a “high-risk” subgroup of these patients might benefit from adjuvant chemotherapy if properly identified. Although recent studies have demonstrated significantly improved survival using adjuvant chemotherapy in patients undergoing pulmonary resection for stage IB, II, or IIIA NSCLC, these patients are subject to treatment-related toxicity. It remains a critical and unsolved challenge to estimate more precisely the risk for survival or recurrence in individual patients to provide adjuvant therapy to high-risk patients and avoid providing adjuvant therapy to low-risk patients.

The emerging use of gene expression signatures may enable clinicians to make treatment decisions based on specific characteristics of individual patients and their tumor. Many microarray-based gene profiles with gene numbers varying from dozens to thousands have been reported to predict patient survival in lung cancer. Because of the need for specialized laboratory facilities, the large number of contributing genes, and complex statistical analyses, microarray-based gene expression profiles are not very practical for clinical use. Alternatively, a quantitative real-time polymerase chain reaction (qRT-PCR) method may be used more efficiently. Several qRT-PCR profiles have been reported, but none of these gene sets have been refined or tested adequately for clinical use.

In this study, we have combined microarray gene profiles of a large NSCLC data set and a qRT-PCR-based approach, using a novel gene selection method. A novel aspect of the design was to include genes representing diverse biological processes and then to measure gene expression using qRT-PCR. Finally, we developed and then verified a 91-gene qRT-PCR card-based platform survival classifier using an independent cohort of 101 lung adenocarcinomas. The strategy used in this study is shown in Figure 1. This strategy was prospectively defined and executed as planned.
PATIENTS AND METHODS

The methods are briefly described in this article with full descriptions in the Supplementary Methods (Supplemental Digital Content 1, http://links.lww.com/JTO/A115).

Published Microarray Data Collection

Three published Affymetrix microarray data sets representing 680 primary tumors were used in the survival-related gene selection procedure. The primary training data set included 439 lung adenocarcinomas from a consortia study of four centers,13 and a combined 111 lung adenocarcinomas and squamous cell carcinoma (SCC) data set represented test set 18 and a 130 lung SCC data set was used as test set 2.12 The clinical information for these three data sets is provided in Table 1. Other microarray data sets7,9,10,14 were not chosen due to platform differences, sample sizes less than 100, or not having survival information. Our primary outcome was overall survival for all datasets censored at 5 years. The information concerning adjuvant chemotherapy or radiation therapy was provided in the original articles and also summarized in Table 1.

Patients and Tissue Specimens for qRT-PCR Measurements

A subset of 47 of the 439 patients had qRT-PCR measured. In addition, we identified an independent validation set of 101 lung adenocarcinomas procured from patients having pulmonary resection for cancer between February 1992 and November 2007 at the University of Michigan (a total of 120 samples were examined using qRT-PCR, including 12 paired normal lung and tumor tissues and 7 duplicate tumor samples representing a different portion of these tumors). This study was approved by the Institutional Review Board of the University of Michigan. None of the patients received preoperative chemotherapy or radiation therapy. A total of 68 patients were stage 1b or above 1b, and 38 of these 68 patients (56%) received adjuvant chemotherapy or radiation therapy. No adjuvant therapy was provided to the 29 of 33 stage 1a patients. The clinical information for this cohort is presented in Table 1.

FIGURE 1. Strategy for the development and validation of the 91-gene quantitative real-time polymerase chain reaction (qRT-PCR) classifier for lung cancer prognosis.

TABLE 1. Clinical Characteristics of Samples Used in this Study

<table>
<thead>
<tr>
<th>Data Set</th>
<th>Training Set</th>
<th>Bild Test Set</th>
<th>Raponi Test Set</th>
<th>Validation Set</th>
</tr>
</thead>
<tbody>
<tr>
<td>Platform</td>
<td>U133A</td>
<td>U133 + 2.0</td>
<td>U133A</td>
<td>qRT-PCR</td>
</tr>
<tr>
<td>Sample number</td>
<td>439</td>
<td>111</td>
<td>130</td>
<td>101</td>
</tr>
<tr>
<td>Type of cancer</td>
<td>AD</td>
<td>AD/53 SCC</td>
<td>SCC</td>
<td>AD</td>
</tr>
<tr>
<td>Age average (SD)</td>
<td>64.4 (10.1)</td>
<td>64.8 (9.6)</td>
<td>67.5 (9.9)</td>
<td>67.0 (9.6)</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>218 (49.7%)</td>
<td>48 (43.2%)</td>
<td>48 (36.9%)</td>
<td>53 (52.5%)</td>
</tr>
<tr>
<td>Male</td>
<td>221</td>
<td>63</td>
<td>82</td>
<td>48</td>
</tr>
<tr>
<td>Stage</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stage I</td>
<td>276 (62.9%)</td>
<td>67 (63.2%)</td>
<td>73 (56.2%)</td>
<td>59 (58.4%)</td>
</tr>
<tr>
<td>Stage II</td>
<td>104</td>
<td>18</td>
<td>34</td>
<td>16</td>
</tr>
<tr>
<td>Stage III</td>
<td>59</td>
<td>21</td>
<td>23</td>
<td>26</td>
</tr>
<tr>
<td>Differentiation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Well</td>
<td>60</td>
<td>NA</td>
<td>15</td>
<td>28</td>
</tr>
<tr>
<td>Moderate</td>
<td>208</td>
<td>NA</td>
<td>76</td>
<td>38</td>
</tr>
<tr>
<td>Poor</td>
<td>166 (38.3%)</td>
<td>NA</td>
<td>39 (30%)</td>
<td>34 (33.7%)</td>
</tr>
<tr>
<td>Dead (5 yr)</td>
<td>186 (42.4%)</td>
<td>58 (52.3%)</td>
<td>52 (40%)</td>
<td>44 (43.6%)</td>
</tr>
<tr>
<td>Alive</td>
<td>253</td>
<td>53</td>
<td>78</td>
<td>57</td>
</tr>
<tr>
<td>Median survival (mo)</td>
<td>47</td>
<td>31.1</td>
<td>34.5</td>
<td>28.8</td>
</tr>
<tr>
<td>Adjuvant therapy</td>
<td>108</td>
<td>Unknown</td>
<td>48</td>
<td>42</td>
</tr>
<tr>
<td>No adjuvant therapy</td>
<td>329</td>
<td>Unknown</td>
<td>69</td>
<td>58</td>
</tr>
</tbody>
</table>

AD, adenocarcinoma; SCC, squamous cell cancer; NA, not applicable.

Adjuvant therapy includes chemotherapy and/or radiotherapy. Unknown adjuvant therapy were 2, 13, and 1 for Training, Raponi, and Validation sets, respectively.
Custom TaqMan Low-Density Arrays and qRT-PCR

Regions containing a minimum of 70% tumor cellularity were used for RNA isolation. RNA quality was analyzed by 2100 Bioanalyzer (Agilent Santa Clara, CA). Custom TaqMan Low-Density Arrays (384-well micro fluidic cards) were obtained from Applied Biosystems Inc. The primers of survival-related genes including endogenous loading control gene (18S RNA, beta-actin, and glyceraldehyde-3-phosphate dehydrogenase) and blank controls were preapplied to the cards. The preparation and running of the microfluidic cards (qRT-PCR) followed the guidelines of the product protocols. Cycle threshold values were generated for each card by automatic selection of a threshold. The technical performance and repeatability measures were tested (Supplementary Figures S1–S3, Supplemental Digital Content 1, http://links.lww.com/JTO/A115). After prescreening, 13,306 probes were left for further analysis. All genes in training and testing datasets were median-centered and the median absolute deviation (MAD) scaled before use in subsequent analyses.

Statistical Analysis

Initial Microarray Data Processing and Filtering

The preprocessing and filtering steps were identical to those described in Shedden et al.13 (Supplementary Methods, Supplemental Digital Content 1, http://links.lww.com/JTO/A115). After prescreening, 13,306 probes were left for further analysis. All genes in training and testing datasets were median-centered and the median absolute deviation (MAD) scaled before use in subsequent analyses.

Preselection of Survival-Related Clusters and Genes

The first step in the strategy was to select approximately 370 promising genes from the Affymetrix data. Using the 439 training samples, genes were first separated into 300 groups using K-means clustering. A two-stage selection procedure was then implemented; first, selection of clusters and second, selection of genes within each of the selected clusters. The top 73 clusters whose cluster mean was most associated with survival were selected using backward elimination and stepwise regression using a Cox proportional hazard model. Within each of the selected clusters, the second selection identified a subset of genes prognostic for survival based on a combination of various criteria, which are described in Supplementary Methods (Supplemental Digital Content 1, http://links.lww.com/JTO/A115). This approach led to a set of 73 clusters and a total of 368 genes from the selected clusters considered relevant to patient survival of lung cancer. We also tested the survival predictability of these 368 genes using two independent datasets (Bild et al., 110 samples and Raponi et al., 130 samples) with random survival forests (RSFs) (Supplementary Results, Supplemental Digital Content 1, http://links.lww.com/JTO/A115).

Normalization and Imputation of qRT-PCR Values

The second step in the strategy was to measure the preselected 368 genes using qRT-PCR technology on a subset of the 439 training samples. The qRT-PCR measurements on the 47 samples were standardized using the control gene 18S RNA. The qRT-PCR measurements on the remaining 392 patients were treated as missing data, and a multiple imputation (MI) strategy was used to make full use of all 439 patients to build a prediction model. The MI was performed using IVEware.23 The details of MI and alternative strategies of normalizing the microarray data for the training set of 439 samples are described in the Supplementary Methods (Supplemental Digital Content 1, http://links.lww.com/JTO/A115). The Spearman's correlations between qRT-PCR and Affymetrix-based measurements were calculated.

RSFs for Survival Analysis and Prediction

The third step of the strategy was to further refine the selected genes down to approximately 91 genes for final evaluation and validation using qRT-PCR measurements. The details of RSF are described in the Supplementary Methods (Supplemental Digital Content 1, http://links.lww.com/JTO/A115).

To assess the statistical significance of the predictions on the validation dataset, the mortality risk index (MRI) was included as a continuous variable in a univariate Cox model both for all 101 tumors and for stage 1 tumors only. Kaplan-Meier survival analyses are shown using the MRI to separate patients into three risk tertiles (high-, medium-, and low-risk, one-third in each group).

RESULTS

Identification of a Survival-Related 91 Gene Subset

To identify a 91-gene qRT-PCR platform-based classifier obtained from a subset of the 368 genes selected in the Affymetrix platform, three major criteria were considered: correlation between Affymetrix and qRT-PCR, association of gene with survival, and representation of a broad spectrum of biological processes.

First, we defined genes whose qRT-PCR measurement showed high correlation with Affymetrix microarray measurements based on the same 47 samples used in the training set. There were 301 of 368 (82%) genes, which had a significantly high correlation value larger than 0.5 (p < 0.001, Table 2 and Supplementary Figure S5, Supplemental Digital Content 1, http://links.lww.com/JTO/A115).

<table>
<thead>
<tr>
<th>Spearman Correlation</th>
<th>No. of Genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;0.9</td>
<td>67</td>
</tr>
<tr>
<td>0.8–0.9</td>
<td>92</td>
</tr>
<tr>
<td>0.7–0.8</td>
<td>71</td>
</tr>
<tr>
<td>0.6–0.7</td>
<td>46</td>
</tr>
<tr>
<td>0.5–0.6</td>
<td>25</td>
</tr>
<tr>
<td>0.4–0.5</td>
<td>23</td>
</tr>
<tr>
<td>0.3–0.4</td>
<td>15</td>
</tr>
<tr>
<td>0.2–0.3</td>
<td>13</td>
</tr>
<tr>
<td>&lt;0.2</td>
<td>16</td>
</tr>
</tbody>
</table>

qRT-PCR, quantitative real-time polymerase chain reaction.
Second, based on the Affymetrix expression values and the measured qRT-PCR expression values, we imputed the qRT-PCR values for the remaining patients in the training dataset. We performed a RSF using 1000 trees and repeated it 10 times on each of the 20 imputed training data sets. We selected genes that had either: (a) \( p \) values from the Cox model adjusted for stage and age on the imputed PCR data which were less than 0.05 or (b) average variable importance measure (VIMP) from the RSF (mean of 10 VIMPs per dataset) was larger than the “noise” VIMP average from RSF.

The final step was a subjective one of reducing the number of genes to 91, while retaining representation from each cluster if possible and selecting multiple genes from the largest clusters if the cluster and the gene appeared to be strongly associated with survival. A set of 91 genes from 53 clusters were selected.

To compare the relative prediction capability of the 91-gene classifier with the 368-gene classifier based on the Affymetrix data, we performed the similar RSF prediction analysis as done with the 368-gene signature described in Supplementary Results (Supplemental Digital Content 1, http://links.lww.com/JTO/A115). The 91-gene signature gave a similar prediction result when compared with using 368 genes with both of the two test sets. The prediction error rates were 40.7% (33.9% for adenocarcinomas and 43.9% for SCC) and 36.3%, respectively, for the Bild and Raponi test sets (Supplementary Table S2, Supplemental Digital Content 2, http://links.lww.com/JTO/A116). This indicated that the 91-gene signature was comparable with the 368-gene signature in predicting patient survival in lung cancer.

The annotation of the 91 genes is provided in Supplementary Table S3 (Supplemental Digital Content 2, http://links.lww.com/JTO/A116), and the main biological categories are indicated in Supplementary Figure S6 (Supplemental Digital Content 1, http://links.lww.com/JTO/A115). Among these, signal transduction, transcription regulation, cell cycle, cell adhesion, and proliferation are the major biological processes.

**Validation of the 91-Gene Classifier in an Independent Test Set**

To validate the 91-gene classifier for lung cancer prognosis, we used the qRT-PCR card-based platform with a completely independent cohort of 101 lung adenocarcinomas. The qRT-PCR data were normalized as described earlier in the text. The RSF prediction model was built with the 91 genes, tumor stage information, and patient age information using the average of 20 imputed training sets of 439 tumors. The qRT-PCR data obtained from the 101 tumors were then tested with the RSF prediction model. The prediction error rate for the 101 test cohort was 26.6% (Table 3). We then tested the usefulness of predictors used to build the RSF using a univariate Cox model with the MRI as a continuous covariate. The RSF prediction was significant for the 101 patient cohort (likelihood ratio test [LRT] \( p < 0.00001 \)). Using the MRI produced from the RSF, three risk groups were also identified, with patient 5-year survival being significantly different between low-, medium-, and high-risk groups (hazard ratio = 1.00, 2.82, 4.42, \( p = 0.0008 \); Figure 2A and Table 3). For stage I tumors only, this MRI was also significantly related to survival (Cox model LRT, \( p = 0.001 \)) and separated patients into low-, medium-, and high-risk groups (hazard ratio = 1.00, 3.29, 3.78, \( p = 0.04 \); Figure 2B and Table 3). The area under the curves from receiver operating characteristic analyses were both 0.77 for all patients and for stage I only (Supplementary Figure S7, Supplemental Digital Content 1, http://links.lww.com/JTO/A115). A notable feature of the validation shown in Figure 2 is the large separation.

### Table 3. Prediction Results of the 91-Gene qRT-PCR Signature in the Validation set (n = 101)

<table>
<thead>
<tr>
<th></th>
<th>RSF Error Rate</th>
<th>Cox Model</th>
<th>Log-Rank Test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( p )</td>
<td>HR 95% CI</td>
<td>( p )</td>
</tr>
<tr>
<td>Low risk</td>
<td>26.6%</td>
<td>5.21e-09</td>
<td>1</td>
</tr>
<tr>
<td>Medium risk</td>
<td>2.82</td>
<td>1.16–6.88</td>
<td>0.04</td>
</tr>
<tr>
<td>High risk</td>
<td>4.42</td>
<td>1.88–10.42</td>
<td></td>
</tr>
</tbody>
</table>

a RSF prediction model built from 439 training set including 91 genes, stage, and age.

b Mortality risk index (MRI) as continuous value, likelihood ratio test (LRT) was used, univariate Cox model.

c MRI separated test patients to three risk groups (low, medium, and high risk, one-third in each group).

qRT-PCR, quantitative real-time polymerase chain reaction; RSF, random survival forest; HR, hazard ratio; CI, confidence interval.

**FIGURE 2.** Prediction results of the 91-gene quantitative real-time polymerase chain reaction (qRT-PCR) signature in the 101 samples of validation set. A, Kaplan-Meier survival curve using patient mortality risk index (MRI) from the random survival forest (RSF) prediction model built from training set including 91 genes, stage, and age. This predictor could significantly separate high-, medium-, and low-risk groups (one-third in each group, HR = 1.00, 2.82, 4.42, \( p = 0.0008 \)) among all 101 patients (A), and among the 59 stage 1 patient (one-third in each group, HR = 1.00, 3.29, 3.776; B).
between the curves in the first 2 years of follow-up, with almost no patients dying in the first 2 years for the low-risk group, but with significant number of deaths in the first 2 years for the high-risk group.

To confirm the MI strategy, we developed results comparable with other methods. We built a RSF prediction model directly from the Affymetrix microarray data of training set of 439 patients (median centered and MAD scaled) and applied it to the 101 qRT-PCR validation set (similarly MAD scaled). We found that the RSF survival prediction results on the validation set of 101 qRT-PCR patients are similar using both MI (Table 3 and Figure 2A, Supplemental Digital Content 1, http://links.lww.com/JTO/A115) and MAD-scaled microarray data (Supplementary Table S4, Supplemental Digital Content 2, http://links.lww.com/JTO/A116, and Supplementary Figure S8, Supplemental Digital Content 1, http://links.lww.com/JTO/A115) from the training sets of 439 patients.

To evaluate whether the 91-gene set improves the prediction compared with clinical variables, age, and stage in the validation set, we compared two Cox models through LRT; a model with age and stage versus a model with age, stage, and the mortality index. We found that the set of 91 genes improves the prediction capability when compared with age and stage only (LRT $p < 0.0001$) using all 101 patients.

We also compared the effect of adjuvant therapy for these three risk groups defined by the MRI shown in Figure 2A. We did not observe any benefit from adjuvant therapy in the high- and medium-risk groups ($p = 0.8$ and 0.5, respectively). A reduced survival for low-risk patients was observed if adjuvant therapy was given ($p = 0.01$; Supplementary Fig. S10, Supplemental Digital Content 1, http://links.lww.com/JTO/A115). More detailed results and discussion regarded the effect of adjuvant therapy are provided in Supplementary Results (Supplemental Digital Content 1, http://links.lww.com/JTO/A115).

**DISCUSSION**

Lung cancer is a heterogeneous disease, and it is often difficult to accurately predict patient survival using tumor pathological characteristics or staging information only. Several groups, including ours, have postulated that improved estimation of an individual patient’s potential risk for recurrent disease can be achieved by a combination of clinical information and certain molecular markers including gene mutations (e.g., EGFR and KRAS mutations), DNA copy number change (e.g., MET and IGF1R), and gene or microRNA expression. Studies based on gene expression signatures from microarrays or using qRT-PCR assays have been reported as predictive of patient survival in lung cancer, but there is sparse overlap of survival-related genes from different studies. The small amount of gene overlap may reflect sample collection methods, processing protocols, single-institutional subject cohorts, statistical methods, small sample sizes, different analysis platforms, and different probes used. Ein-Dor et al. have suggested that because of biological heterogeneity, thousands of samples may be required to identify robust and reproducible gene subsets for most tumor types. Boutros et al. demonstrated that thousands of different six-gene signatures can predict patient survival if only six genes were used. The largest microarray-based study of lung adenocarcinomas showed that combining a cluster and Cox model-based method for gene selection plus using clinical covariates provided the best overall survival predictive ability. Because the number of genes in each cluster varied from dozens to hundreds, however, it would be very difficult to apply some microarray-based gene classifiers in the clinical setting. These studies indicate that for an optimal survival classifier, a large sample size, including more genes in the signature, and appropriate statistical methods incorporating accurate clinical information are needed. As an approach, qRT-PCR is a more reproducible, simple, efficient, and clinical practicable assay. In this study, we attempted to combine these major criteria to reduce the number of survival-related genes to a reasonable, but not necessarily a very small number. Then, we developed and validated a qRT-PCR card-based 91-gene survival classifier, using the four major procedures shown in Figure 1, for the purpose of developing a clinically practicable qRT-PCR-based assay.

Of the 91 genes in this study, the functional analysis showed that more than 20 different biological processes were involved. Most of these processes were cancer related, and most of these genes have been reported by others as being involved in cancer development or used for cancer diagnosis or prognosis (Supplementary Table S3, Supplemental Digital Content 2, http://links.lww.com/JTO/A116). We have compared our 91-gene list with 20 other qRT-PCR-based or microarray-based studies for lung cancer prognosis. Eleven genes were reported also by seven other studies (Supplementary Table S5, Supplemental Digital Content 2, http://links.lww.com/JTO/A116). No overlapping genes to this study were reported by the other 13 studies (Supplementary Table S6, Supplemental Digital Content 2, http://links.lww.com/JTO/A116). Several genes including DUSP6 and ERBB3 were also used in Chen et al. five-gene signature, and ERBB3 was used in Raz et al. four-gene model. SLC2A1 and MEF2C were presented in Lou et al. 64-gene profile. Interestingly, five genes were also reported by Beer et al. in 2002.

Gene cluster analysis and a risk index created from Cox models have been successfully used before as the statistical approach for gene expression profile-based survival prediction. Genes in the same cluster which are coordinately expressed in a dataset often represent similar biological functions or define similar pathological features. In our approach, we specifically chose genes representative of as many clusters as possible to aid in prediction performance in the likely case of lung adenocarcinoma tumor heterogeneity. We used both Cox models and RSF to aid in the identification of genes and development of the classifier. In general, performances of RSF and Cox model were similar, with RSF being complementary to the Cox model providing genes important for survival prediction based on the VIMP value. To our knowledge, this is the first study to combine clustering, Cox model, and RSF prediction models for survival-related gene
selection and then to predict survival from qRT-PCR data in lung cancer. This is also the first study to impute a large microarray data to a mimic qRT-PCR value in the training set in the prediction study and we found a similar performance (Table 3 and Figure 2) as using Affymetrix microarray-based values (MAD-scaled) (Supplementary Table S4 and Supplementary Figure S8, Supplemental Digital Content 1, http://links.lww.com/JTO/A115).

The strength of this study is that it was prospectively planned and executed as described in Figure 1. The planned strategy was of incrementally refining and reducing the number of genes as the study transitioned from an Affymetrix platform to a qRT-PCR platform. Another advantage of this study is the large sample size used in the training set, with the data being from one uniform study although measured at four centers. Using the power of imputation, we were able to make use of the survival data from all 439 subjects in the training data rather than just the subset of those who had both Affymetrix and qRT-PCR measurements. We find that the prediction results using Affymetrix-based measurements using RSF and the 368 genes were similar to the imputation method thus supporting its use. Further, the successful survival prediction for the 91-gene qRT-PCR platform also included stage 1 cancers. Interestingly, these 91 genes also can predict patient survival with SCC indicating that some of the same biological processes may be shared with lung SCC44 and suggesting a greater utility of the prognostic gene set. The use of prospective clinical trials to test the prediction of benefit from chemotherapy for patient groups defined by the MRI are needed to broadly use this classifier for treatment selection in lung cancer.

ACKNOWLEDGMENTS

Supported by MICHR/CSTA and MIIE UL1RR024986 (D.G.B.) and CTSABD (S.K.).

REFERENCES

32. Lee ES, Son DS, Kim SH, et al. Prediction of recurrence-free survival in


38. Ein-Dor L, Zuk O, Domany E. Thousands of samples are needed to generate a robust gene list for predicting outcome in cancer. Proc Natl Acad Sci USA 2006;103:5923–5928.


