

Validation of the 12-gene Predictive Signature for Adjuvant Chemotherapy Response in Lung Cancer

Yang Xie^{1,2,3}, Wei Lu⁴, Shidan Wang¹, Ximing Tang⁴, Hao Tang¹, Yunyun Zhou⁵, Cesar Moran⁴, Carmen Behrens⁶, Jack A. Roth⁷, Qinghua Zhou⁸, David H. Johnson⁹, Stephen G. Swisher⁷, John V. Heymach⁶, Vassiliki A. Papadimitrakopoulou⁶, Guanghua Xiao^{1,2,3}, John D. Minna^{3,9,10,11}, and Ignacio I. Wistuba⁴



Abstract

Purpose: Response to adjuvant chemotherapy after tumor resection varies widely among patients with non-small cell lung cancer (NSCLC); therefore, it is of clinical importance to prospectively predict who will benefit from adjuvant chemotherapy before starting the treatment. The goal of this study is to validate a 12-gene adjuvant chemotherapy predictive signature developed from a previous study using a clinical-grade assay.

Experimental Design: We developed a clinical-grade assay for formalin-fixed, paraffin-embedded (FFPE) samples using the NanoString nCounter platform to measure the mRNA expression of the previously published 12-gene set. The predictive performance was validated in a cohort of 207 patients with early-stage resected NSCLC with matched propensity score of adjuvant chemotherapy.

Results: The effects of adjuvant chemotherapy were significantly different in patients from the predicted adjuvant chemotherapy benefit group and those in the predicted adjuvant chemotherapy nonbenefit group ($P = 0.0056$ for

interaction between predicted risk group and adjuvant chemotherapy). Specifically, in the predicted adjuvant chemotherapy benefit group, the patients receiving adjuvant chemotherapy had significant recurrence-free survival (RFS) benefit (HR = 0.34; $P = 0.016$; adjuvant chemotherapy vs. nonadjuvant chemotherapy), while in the predicted adjuvant chemotherapy nonbenefit group, the patients receiving adjuvant chemotherapy actually had worse RFS (HR = 1.86; $P = 0.14$; adjuvant chemotherapy vs. nonadjuvant chemotherapy) than those who did not receive adjuvant chemotherapy.

Conclusions: This study validated that the 12-gene signature and the FFPE-based clinical assay predict that patients whose resected lung adenocarcinomas exhibit an adjuvant chemotherapy benefit gene expression pattern and who then receive adjuvant chemotherapy have significant survival advantage compared with patients whose tumors exhibit the benefit pattern but do not receive adjuvant chemotherapy.

Introduction

Lung cancer is the leading cause of death in the United States and around the world, and non-small cell lung cancer (NSCLC) accounts for up to 85% of such deaths (1–3). Large-scale clinical trials have demonstrated the survival benefits of adjuvant chemotherapy and have established adjuvant chemotherapy as the standard of care for completely resected stage II–III NSCLC, and shown the benefits of adjuvant chemotherapy for stage IB patients with tumors larger than 4 centimeters (4–9). However, the effect

of adjuvant chemotherapy on prolonging survival is modest, while such treatment is also associated with serious adverse effects (4, 8, 10, 11). Therefore, the routine use of adjuvant chemotherapy is not justified for all patients with resected NSCLC (9), and it is critical to identify the subgroup of patients who will most likely benefit from adjuvant chemotherapy (12).

Recently, using a cohort of 442 patients with stage I–III NSCLC from the Lung Cancer Director Challenge Dataset (13), we discovered a 12-gene signature that predicts adjuvant chemotherapy response (14) in patients with NSCLC. In addition, this gene

¹Quantitative Biomedical Research Center, Department of Clinical Sciences, University of Texas Southwestern Medical Center, Dallas, Texas. ²Department of Bioinformatics, University of Texas Southwestern Medical Center, Dallas, Texas. ³Simmons Cancer Center, University of Texas Southwestern Medical Center, Dallas, Texas. ⁴Department of Translational Molecular Pathology, University of Texas MD Anderson Cancer Center, Houston, Texas. ⁵Department of Data Science, University of Mississippi Medical Center, Jackson, Mississippi. ⁶Department of Thoracic/Head and Neck Medical Oncology, University of Texas MD Anderson Cancer Center, Houston, Texas. ⁷Department of Thoracic and Cardiovascular Surgery, University of Texas MD Anderson Cancer Center, Houston, Texas. ⁸Lung Cancer Institute, West China Hospital, Sichuan University, Chengdu, China. ⁹Department of Internal Medicine, University of Texas Southwestern Medical Center, Dallas, Texas. ¹⁰Hamon Center for Therapeutic Oncology Research, University of Texas Southwestern Medical Center, Dallas, Texas. ¹¹Department of Pharmacology, University of Texas Southwestern Medical Center, Dallas, Texas.

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

Y. Xie, W. Lu, and S. Wang contributed equally as co-first authors of this article.

Current address for H. Tang: Bristol-Myers Squibb, Princeton, New Jersey.

Corresponding Authors: Yang Xie, UT Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, TX 75205. Phone: 214-648-5178; Fax: 214-648-1663; E-mail: yang.xie@utsouthwestern.edu; and Ignacio I. Wistuba, Department of Translational Molecular Pathology, University of Texas MD Anderson Cancer Center, 1515 Holcombe Blvd., Houston, TX 77030. E-mail: iiwistuba@mdanderson.org

doi: 10.1158/1078-0432.CCR-17-2543

©2018 American Association for Cancer Research.

Translational Relevance

Response to standard chemotherapy in lung cancer varies widely among patients. Therefore, it is of substantial clinical importance to be able to predict who will benefit from adjuvant chemotherapy before starting treatment. Although a large number of cancer biomarkers have been reported, few have been translated into real clinical tools. The major bottleneck in translating biomarker discovery to improved patient outcomes is the availability of accurate clinical tests (assays) that will allow treatments to be optimized and tailored to an individual's needs. In this study, we developed and validated a clinical-grade assay to measure mRNA expression levels from formalin-fixed, paraffin-embedded (FFPE) tumor samples, together with validating a predefined risk stratification algorithm for patient response to adjuvant chemotherapy after tumor resection. We validated that the previously published 12-gene signature measured from FFPE tumor samples provided prognostic information in patients with resected lung adenocarcinoma who did not receive adjuvant chemotherapy. Importantly, this gene signature and the FFPE-based clinical assay predict that patients whose resected lung adenocarcinomas exhibit an adjuvant chemotherapy benefit gene expression pattern and who then receive adjuvant chemotherapy have a significant survival advantage, compared with patients whose tumors exhibit the benefit pattern but do not receive adjuvant chemotherapy.

signature predicts patient prognosis in patients with lung adenocarcinoma, but not in patients with squamous cell carcinoma (SCC). This signature was derived and validated using microarray technology from fresh frozen tumor samples. The crucial next step in translating this discovery to clinical practice is to develop a clinical assay for this 12-gene signature in formalin-fixed, paraffin-embedded (FFPE) samples. In this study, we developed a clinical-grade assay using the NanoString nCounter GX platform to measure the expression of the 12-gene signature from FFPE samples. Both the prognostic and predictive performance of this clinical assay together with the original algorithm were validated in FFPE samples from patients with early-stage (defined as stage I and II) NSCLC, and we demonstrated that the benefit group predicted by this gene signature showed significant improvement in survival after adjuvant chemotherapy, while the predicted nonbenefit group actually showed worse survival after adjuvant chemotherapy.

Materials and Methods

A list of 12 genes (*ATP8A1*, *AURKA*, *C1orf116*, *COL4A3*, *DOCK9*, *HOPX*, *HSD17B6*, *IFT57*, *MBIP*, *NKX2-1*, *RRM2*, and *TTC37*) and the predefined risk prediction algorithm were extracted from our previous study (14). The NanoString nCounter platform was used to develop an assay that could measure the expression levels of these 12 genes from FFPE tissue samples. The assay was developed and optimized using a cohort of 30 patients with NSCLC (with both FFPE and fresh frozen tissues), and the quality control measures were determined from this cohort. A cohort of 327 patients with early-stage (stage I and II) NSCLC was used for assay validation. The assay development and validation

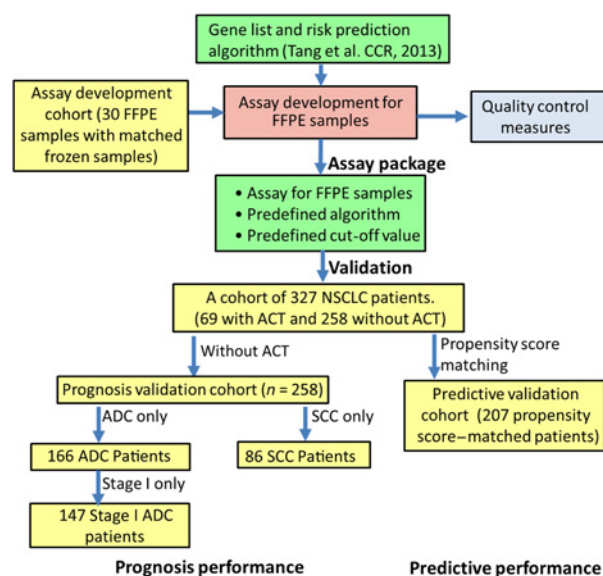


Figure 1.

Flowchart for assay development and validation. ACT, adjuvant chemotherapy; ADC, adenocarcinoma.

procedures are summarized in Fig. 1 and also detailed in the following sections. The resulting tables and figures are summarized in Supplementary Table S1.

Patient cohorts

The patients gave written consent; the study was approved by the ethics committee at the University of Texas MD Anderson Cancer Center (IRB #LAB90-020 and #LAB03-0320; Houston, TX); and the study was conducted in accordance with the Declaration of Helsinki.

Cohort for assay development (n = 30). A cohort of 30 patients with NSCLC (with matched FFPE and fresh frozen tissues) was used for the assay development. The expression levels of the 12-gene set were measured in both FFPE and fresh frozen tissues using the NanoString nCounter assay. The correlation between expression levels measured from FFPE tissue samples and those measured from corresponding fresh frozen samples was calculated for each individual gene to validate the measuring accuracy of the assay for FFPE samples.

Cohort for validating prognostic performances (n = 258). The FFPE tissue samples of 327 patients with early-stage (stage I and II) NSCLC were obtained from the University of Texas Lung Cancer Specialized Program of Research Excellence (SPoRE) Tissue Bank at the MD Anderson Cancer Center (MDACC, Houston, TX). This study was approved by the MD Anderson Institutional Review Board. Among these 327 early-stage patients, 69 (21.1%) were treated with adjuvant chemotherapy, and the remaining 258 (78.9%) were not treated with adjuvant chemotherapy. None of the patients received neoadjuvant chemotherapy. The 258 patients without adjuvant chemotherapy (166 adenocarcinoma, 86 SCC, and 6 others) were used as the validation cohort for prognostic performance of the 12-gene signature.

Table 1. Patient characteristics of the study cohorts

	Assay development cohort (n = 30)	Prognostic validation cohort (n = 258)	Predictive validation cohort (n = 207)
Histology			
Adenocarcinoma	13 (43.3%)	166 (64.3%)	138 (66.7%)
SCC	13 (43.3%)	86 (33.3%)	69 (33.3%)
Other	4 (13.3%)	6 (2.3%)	0 (0.0%)
Stage			
IA	5 (16.7%)	139 (53.9%)	71 (34.3%)
IB	6 (20.0%)	80 (31.0%)	50 (24.2%)
IIA	6 (20.0%)	23 (8.9%)	46 (22.2%)
IIB	8 (26.7%)	16 (6.2%)	40 (19.3%)
III	4 (13.3%)	0 (0.0%)	0 (0.0%)
IV	1 (3.3%)	0 (0.0%)	0 (0.0%)
Gender			
Female	14 (46.7%)	129 (50.0%)	97 (46.9%)
Male	16 (53.3%)	129 (50.0%)	110 (53.1%)
Age (year)	68.77	68.08	63.85
Follow-up time (months)	52.49	56.07	53.94
Death			
Yes	23 (76.7%)	193 (74.8%)	151 (73.0%)
No	7 (23.3%)	65 (25.2%)	56 (27.1%)
Adjuvant therapy			
Yes	9 (30.0%)	0 (0.0%)	69 (33.3%)
No	15 (50.0%)	258 (100%)	138 (66.7%)

NOTE: In this study, we developed and optimized the clinical assay using a cohort of 30 patients with NSCLC. The prognostic performance of the assay was validated in the FFPE samples of a cohort with 258 patients with NSCLC without adjuvant chemotherapy, and the predictive performance was validated in 207 propensity score-matched patients from the cohort of the 327 patients with NSCLC. The patient characteristics of the three cohorts are summarized in the table above.

Cohort for validating predictive performances (n = 207). Because this is a retrospective cohort, to minimize the confounding factors, we used a propensity score matching technique to estimate the effect of adjuvant chemotherapy by accounting for the covariates (detailed in a later section). From the original 327 patients, we derived a cohort of 207 propensity score-matched patients, among which 69 patients (33.3%) were treated with adjuvant chemotherapy and 138 patients (66.7%) were not treated with adjuvant chemotherapy, to validate the predictive performance of the assay in FFPE samples.

Detailed information on the patients in the development cohort (n = 30), the prognostic validation cohort (258 patients without adjuvant chemotherapy), and the predictive validation cohort (207 propensity score-matched patients) are summarized in Table 1.

Propensity score matching

We used a propensity score matching technique (15, 16) to match the patients with and without adjuvant chemotherapy, to adjust for potential confounding factors in patient selection for adjuvant chemotherapy. The clinical variables, including histology, gender, age, smoking history, and stage, were used in the propensity score matching, and the coefficient and *P* value of each variable in the logistic regression model for propensity score matching are summarized in Supplementary Table S2. In this study, we used a 1:2 ratio for adjuvant chemotherapy treated and untreated patients, and the final matched patients included 69 adjuvant chemotherapy treated (one patient was removed

because of missing values in covariates) and 138 patients without adjuvant chemotherapy. The propensity score matching results are summarized in Supplementary Table S3. These 207 propensity score-matched patients were used in the following analyses to validate the predictive performance of the 12-gene assay. The propensity score matching was implemented using R package *matchit* (17).

RNA extraction

An unstained section was obtained from FFPE tissue blocks at 10- μ m thickness, attached to a noncoated glass slide, and stored at -80°C in a freezer. The section was briefly baked at 65°C for 30 minutes and then deparaffinized with CitriSolv (Thermo Fisher Scientific, #22-143-975). Tumor areas were macrodissected and collected in a fresh Eppendorf microcentrifuge tube containing lysis buffer. Total RNAs were extracted using AllPrep DNA/RNA FFPE Extraction Kit (Qiagen, catalog no. 80234). RNA was quantified by NanoDrop 2000 and its quality was assessed by Agilent Bioanalyzer 1000 with Agilent RNA 6000 Nano Kit (Agilent, catalog no. 5067-1511).

NanoString nCounter gene expression assay

NanoString nCounter technology uses unique color-coded molecular barcodes that can hybridize directly to target nucleic acid molecules, such as mRNA, to discover gene expression level with no need for amplification. The nCounter gene expression assay can investigate multiple genes in a single reaction. Probe sequences were custom designed for target genes and manufactured by NanoString Technologies. This CodeSet includes 7 housekeeping genes to correct for RNA input amount and/or quality differences. Housekeeping genes were selected from publicly available databases based on stability and detectable expression levels across the tissue type of interest. The manufacturer's protocol was followed to perform an assay. A master mix of hybridization buffer and Reporter CodeSet was prepared at 10 μL each, then 5 μL of total RNA sample and 5 μL of Capture ProbeSet were added to the reaction sequentially. After incubating at 65°C for hybridization for 20 to 22 hours, the reaction products were loaded to GEN2 Prep Station for washing and immobilizing signals to cartridge. The cartridge was imaged in Digital Analyzer at 555 fields of view, whose images would be interpreted to count for specified targets by nSolver software (NanoString Technologies).

Gene expression data preprocessing

We followed the procedure by Veldman-Jones and colleagues (18) for gene expression data preprocessing. NAPPA R package (<https://cran.r-project.org/>) from the CRAN (Comprehensive R Archive Network) was used to normalize nCounter data in three steps: (i) a truncated Poisson correction was used to adjust the background signals in raw NanoString counts using internal-negative controls. (ii) The data were then normalized using internal-positive controls. (iii) A sigmoid shrunken slope normalization was used to correct the data for input amount variation using the mean expression of housekeeping genes. If a raw count was below the average of the eight internal-negative control raw counts plus 2 SDs, the transcript was designated as not detected. Transcripts below the limit of detection were not included in the data analysis. Finally, the data were \log_2 transformed.

Predict the risk score and assign risk groups

This is a prospective validation study, so the exact prediction model described in our previous study (14) was implemented. Specifically, predefined supervised principal component analysis was performed for the 12 genes. Using this predefined model, a risk score was assigned to each patient on the basis of the patient's gene expression level of the 12-gene signature.

Patients were separated into the high-risk (i.e., adjuvant chemotherapy benefit) or low-risk (i.e., adjuvant chemotherapy nonbenefit) group based on the patient's risk score. A predefined cutoff for risk score was needed to validate the prognostic and predictive performances of the developed assay. In this study, we determined the risk score cutoff based on the assay development cohort ($n = 30$), and applied this predefined cutoff into the prognostic performance validation cohort ($n = 258$) and the predictive performance validation cohort ($n = 207$). The risk score was calculated by the expression level of the 12-gene signature in FFPE samples for each patient in the assay development cohort. The predicted risk scores of the 30 patients were plotted in Supplementary Fig. S1A, which indicates a bimodal distribution of the predicted risk scores. To determine a data driven threshold value of the high- and low-risk patient groups, a model-based clustering method (implemented by R package *mclust*) was used to fit the risk score distribution with a mixture model of two normal distributions, one corresponding to the high-risk group (red) and the other corresponding to the low-risk group (green). On the basis of the fitting of the mixture model, the probability of a specific patient being in the high-risk group could be calculated from the predicted risk score (derived from the expression of the 12-gene signature) of the patient. This functional relationship is shown in Supplementary Fig. S1B.

Statistical analysis

Recurrence-free survival time (RFS) was calculated from the date of surgery until recurrence of lung cancer or death or the date of last follow-up contact. Survival curves were estimated using the Kaplan–Meier product limit method (19). Differences in the survival outcomes between the predicted high- and low-risk groups were compared using a log-rank test. A multivariate Cox proportional hazards model (20) was used to determine the association between the factors of interest (i.e., adjuvant chemotherapy, predicted risk-group, and their interaction term) and the patient survival outcomes adjusted for other clinical variables, including age, gender, smoking status, histology, and stage. Multivariate Cox proportional hazards models were also used to determine the benefit of adjuvant chemotherapy adjusted for other clinical variables in the predicted high- and low-risk groups, respectively.

The time-dependent ROC curves (21) at 24 months were calculated using R package *survivalROC*. The AUC of each time-dependent ROC curve was calculated as a measure of prediction accuracy for patient prognosis. The ROC curve for the predictive analysis was calculated using R package *tsm* (22).

The concordance index (CI), defined by $\text{Concordance} = \frac{\# \text{all concordant pairs}}{\# \text{total pairs ignoring ties}}$ (23), was calculated using the R package *survcomp*.

Results

mRNA expression measured from FFPE samples

The expression levels of the 12-gene set were measured from both FFPE and frozen tissue samples by nCounter gene expression assay in the development cohort ($n = 30$) to compare the consistency of the measurements from the two types of tissue samples. The dynamic range included the average, maximum, and minimum count of each gene, and the Pearson correlation coefficients between the measurements in FFPE and fresh frozen tissue samples are summarized in Supplementary Table S5. The Pearson correlation coefficients for different genes ranged from 0.5 to 0.8, indicating that the measurements from FFPE samples were reliable and consistent with those from fresh frozen samples.

Prognostic performance of the 12-gene signature

The high- and low-risk groups defined by the 12-gene signature from FFPE tissue samples showed significant differences in RFS (Fig. 2C) for the 258 patients with NSCLC without adjuvant chemotherapy treatment: the high-risk group had worse prognosis in RFS [HR = 1.83 (1.19–2.95); $P = 0.0114$]. Prognoses for patients with different histology types' adenocarcinoma and SCC were also analyzed. The high-risk group had significantly worse RFS prognosis for patients with adenocarcinoma [Fig. 2A; HR = 4.12 (2.07–8.22); $P = 1.25e-05$], and patients with stage I-only adenocarcinoma [Fig. 2B; HR = 3.91 (1.82–8.38); $P = 0.000161$]. The time-dependent ROCs were calculated in both, the patients with adenocarcinoma (Supplementary Fig. S2A; AUC = 0.75) and the patients with stage I adenocarcinoma (Supplementary Fig. S2B; AUC = 0.74). Furthermore, the RFS differences were still significant in patients with adenocarcinoma [HR = 4.20 (2.01–8.79); $P = 0.00014$; Table 2] and in patients with stage I adenocarcinoma [HR = 3.23 (1.26–8.30); $P = 0.015$; Supplementary Table S6] after adjusting for other clinical variables in the multivariate analysis. On the other hand, for patients with SCC ($n = 90$), the predicted high- and low-risk groups did not show any differences in RFS [Fig. 2D; HR = 1.16 (0.272–4.92); $P = 0.83$]. In summary, the prognostic value of the 12-gene signature is adenocarcinoma specific, which is consistent with the findings in our original study. Therefore, this study validated the prognostic performance of the 12-gene signature in FFPE tissue samples using a biomarker assay that can be developed into a Clinical Laboratory Improvement Amendments-certified test for use in patients with adenocarcinoma, to move it further toward clinical application.

Predictive performance of the 12-gene signature

The 207 patients with early-stage NSCLC (propensity score-matched for adjuvant chemotherapy treatment) in the validation cohort were placed by the assay into two groups: those predicted to benefit from adjuvant chemotherapy (adjuvant chemotherapy benefit) and those predicted not to benefit from adjuvant chemotherapy (nonbenefit) group, using the same risk score and cutoff criteria. Figure 3 shows the RFS curves for patients with and without adjuvant chemotherapy in the predicted adjuvant chemotherapy benefit and nonbenefit groups, respectively. In the predicted adjuvant chemotherapy benefit group (high-risk group; Fig. 3A), the patients who received adjuvant chemotherapy had longer RFS time than those who did not receive adjuvant chemotherapy, while in the predicted–adjuvant chemotherapy nonbenefit group (low-risk group; Fig. 3B), patients who received

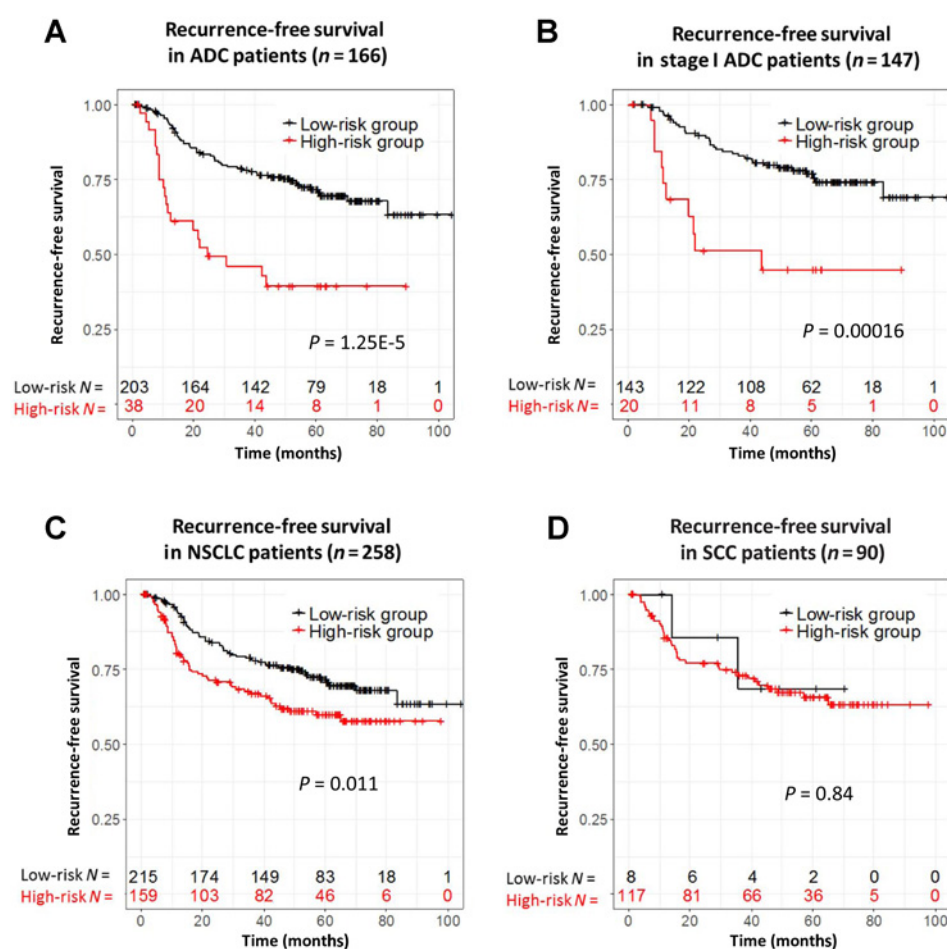


Figure 2. Prognostic performance of the 12-gene signature in patients with adenocarcinoma (ADC) without adjuvant chemotherapy (ACT) treatment. **A**, RFS in patients with adenocarcinoma ($n = 166$). **B**, RFS in patients with stage I adenocarcinoma ($n = 147$). **C**, RFS in patients with NSCLC ($n = 258$). **D**, RFS in patients with SCC ($n = 90$).

adjuvant chemotherapy actually exhibited worse survival than those who did not receive adjuvant chemotherapy. Multivariate analyses for the efficacy of adjuvant chemotherapy treatment shows that adjuvant chemotherapy had significant RFS benefit ($HR = 0.34$; $P = 0.016$) in the predicted–adjuvant chemotherapy benefit group (Fig. 3C), while adjuvant chemotherapy was not associated with RFS benefit ($HR = 1.86$; $P = 0.139$) in the predicted–adjuvant chemotherapy nonbenefit group (Fig. 3D) after adjusting for other clinical variables. To test the interaction between the predicted–adjuvant chemotherapy benefit groups and adjuvant chemotherapy treatment effects, we performed multivariate analysis adjusting for clinical variables, including histology, smoking status, age, gender, tumor size, and stage, for all 207 patients (Table 3). This analysis indicated that, after

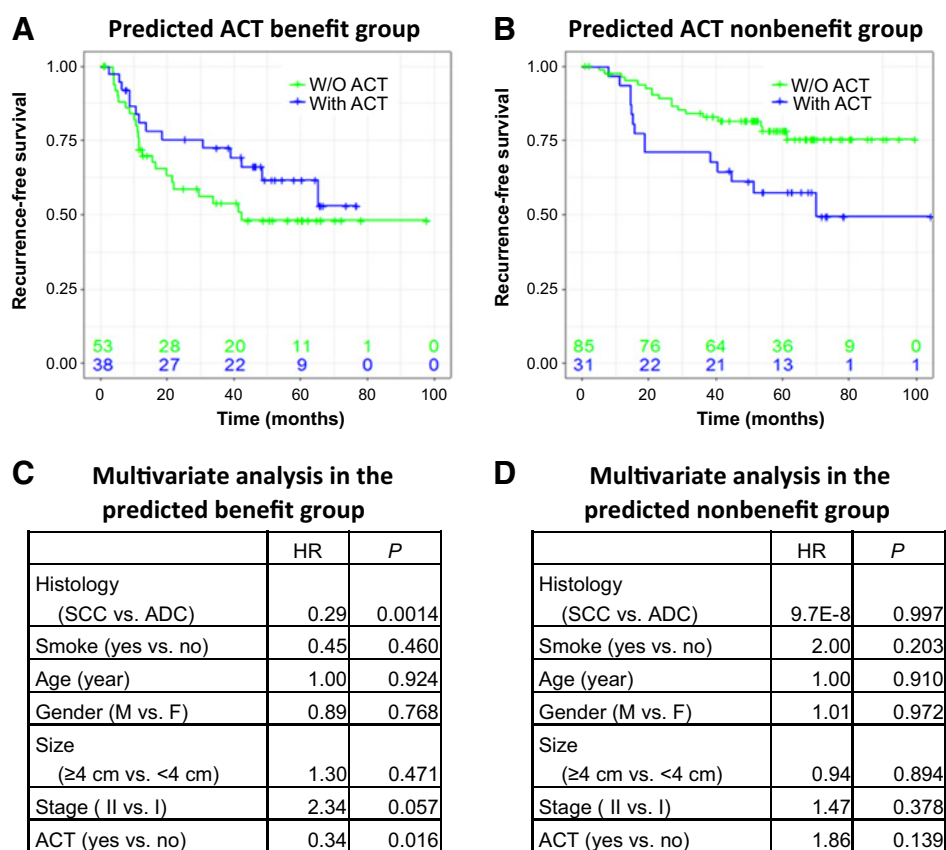
adjusting for other clinical variables, adjuvant chemotherapy has a survival benefit for patients ($HR = 0.40$; $P = 0.020$) and there is significant interaction ($P = 0.0056$) between the effect of adjuvant chemotherapy and the predicted risk groups. We calculated the ROC curve for the predictive analysis (Supplementary Fig. S3). The AUC for the predictive performance was 0.806 and the 95% confidence interval was 0.694–0.865. This demonstrated that the effects of adjuvant chemotherapy are significantly different between the predicted adjuvant chemotherapy benefit group and predicted nonbenefit group. This validated that the 12-gene signature could predict adjuvant chemotherapy survival benefit in patients with early-stage NSCLC and could be used to identify the patient subpopulation that would benefit from the treatment, thereby assisting the clinical decision for using adjuvant chemotherapy. Furthermore, the 12-gene signature could be used to identify patients in whom adjuvant chemotherapy might be harmful. Figure 3B shows that in the predicted adjuvant chemotherapy nonbenefit group, the patients with adjuvant chemotherapy have significantly worse survival outcomes ($P = 0.0154$, log-rank test comparing the two survival curves).

Table 2. Multivariate analysis for validating the prognostic performance for RFS of the 12-gene signature measured from FFPE samples after adjusting for other patient characteristics in the 166 patients with adenocarcinoma without adjuvant chemotherapy

	HR (95% CI)	P
Smoking status (yes vs. no)	9.54 (1.3–69.93)	0.026
Age (year)	1.00 (0.96–1.04)	0.88
Gender (male vs. female)	1.84 (0.99–3.41)	0.054
Size (≥ 4 cm vs. < 4 cm)	1.76 (0.62–4.97)	0.28
Stage (II vs. I)	2.30 (0.84–6.30)	0.11
Group (high-risk vs. low-risk)	4.20 (2.01–8.79)	0.00014

Discussion

Response to standard chemotherapy in lung cancer varies widely among patients. Therefore, it is of substantial clinical

**Figure 3.**

Predictive performance of the 12-gene signature. The effects of adjuvant chemotherapy (ACT) on RFS of patients in benefit and nonbenefit groups. The RFS time compared between adjuvant chemotherapy-treated patients versus nonadjuvant chemotherapy-treated patients in the predicted benefit group (A) and nonbenefit group (B). Multivariate analyses of adjuvant chemotherapy efficacy adjusted for other clinical variables in the predicted adjuvant chemotherapy benefit group (C) and predicted adjuvant chemotherapy nonbenefit group (D), respectively. In the adjuvant chemotherapy benefit group (A and C) predicted by the 12-gene signature, the patients with adjuvant chemotherapy (blue line) had better prognosis than those without adjuvant chemotherapy treatment. In the predicted adjuvant chemotherapy nonbenefit group (B and D), the patients with adjuvant chemotherapy (blue line) had worse prognosis than those without adjuvant chemotherapy treatment ($P = 0.0154$, log-rank test comparing the two survival curves). Furthermore, the multivariate analysis in the predicted adjuvant chemotherapy benefit group shows the significant benefit of adjuvant chemotherapy (HR = 0.34; $P = 0.016$) after adjusting for patient clinical variables. In the adjuvant chemotherapy nonbenefit group, the patients with adjuvant chemotherapy (blue line) had worse prognosis than those without adjuvant chemotherapy treatment. This indicates that this group of patients did not benefit from adjuvant chemotherapy treatment. The multivariate analysis in the predicted nonadjuvant chemotherapy benefit group (D) shows that adjuvant chemotherapy treatment is associated with worse prognosis (HR = 1.86; $P = 0.139$) after adjusting for other clinical variables.

importance to be able to predict who will benefit from adjuvant chemotherapy before starting treatment. Multiple studies have been conducted to identify clinical factors that are associated with

Table 3. Multivariate analysis to test the interaction between the adjuvant chemotherapy treatment effect and the risk groups predicted by the 12-gene signature measured from FFPE samples for RFS among the 207 propensity score-matched patients with NSCLC

	HR (95% CI)	P
Histology (SCC vs. adenocarcinoma)	0.29 (0.14–0.58)	0.00051
Smoking status (yes vs. no)	1.54 (0.59–4.03)	0.37
Age (year)	1.00 (0.97–1.04)	0.88
Gender (male vs. female)	0.96 (0.57–1.61)	0.87
Size (≥4 cm vs. <4 cm)	1.15 (0.66–1.99)	0.62
Stage (II vs. I)	1.73 (0.95–3.15)	0.075
Adjuvant chemotherapy (yes vs. no)	0.40 (0.18–0.87)	0.020
Group (benefit vs. nonbenefit)	0.14 (0.07–0.31)	1.1E-6
Interaction (adjuvant chemotherapy × group)	4.11 (1.51–11.18)	0.0056

chemotherapy response in NSCLC (9, 24–27). Cancer and Leukemia Group B recently demonstrated that there is no significant survival benefit of adjuvant chemotherapy in patients with stage IB NSCLC based on a randomized trial ($P = 0.12$), while a statistically significant survival advantage was observed for stage IB patients with tumors ≥4 cm (9). The Lung Adjuvant Cisplatin Evaluation study showed that the chemotherapy effect was higher in patients with better performance status, and there was no interaction between the chemotherapy effect and gender, age, histology, type of surgery, planned radiotherapy, or planned total dose of cisplatin (27). Currently, a patient's tumor–node–metastasis (TNM) stage is the main clinical variable that provides prognostic information to suggest which patients need adjuvant chemotherapy. However, the TNM information (or the specific tumor histopathologic subtype) does not predict which patients within a TNM-stage category will derive survival benefit from adjuvant chemotherapy. Therefore, identifying and validating

molecular markers with clinical assays to predict adjuvant chemotherapy response is important.

Although a large number of cancer biomarkers have been reported, few have been translated into real clinical tools. The major bottleneck in translating biomarker discovery to improved patient outcomes is the availability of accurate clinical tests (assays) that will allow treatments to be optimized and tailored to an individual's needs. In this study, we developed and tested a clinical-grade assay measuring mRNA expression level from FFPE tumor samples, together with a predefined risk stratification algorithm for patient response to adjuvant chemotherapy after tumor resection. The gene signature measured by the clinical-grade assay demonstrated promising results in predicting adjuvant chemotherapy response of patients with NSCLC.

Thus far, the dominant effort to improve the clinical outcomes of patients with lung cancer with adjuvant chemotherapy has been put toward testing the benefits of adding targeted therapies; for example, the National Cancer Institute of Canada JBR.19 trial added gefitinib to adjuvant chemotherapy for patients with resected NSCLC, and the Eastern Cooperative Oncology Group E1505 trial added bevacizumab. Although molecular targeted therapies are promising, only a small proportion of patients with early-stage NSCLC have mutations that are currently targetable for these therapies. Standard adjuvant chemotherapies are currently the primary treatment choice for patients with lung cancer. Therefore, being able to identify a subgroup of patients who would benefit from adjuvant chemotherapy would be an effective way to improve the clinical outcomes of patients with lung cancer.

This is a retrospective study with propensity score matching to minimize the effect of potential confounders for treatment effects and biomarker effects. Prospective clinical trials are needed to further validate the predictive signature for adjuvant chemotherapy. It would also be interesting to test whether this 12-gene signature could predict chemotherapy response in patients with late-stage lung cancer, and other types of cancer. Finally, multiple measurements of the same specimen on different days will be

helpful for technical validation of the performance of the clinical-grade assay for future clinical application.

Disclosure of Potential Conflicts of Interest

D.H. Johnson is a consultant/advisory board member for Genentech, Merck, Peloton Therapeutics, and Xcovery. No potential conflicts of interest were disclosed by the other authors.

Authors' Contributions

Conception and design: Y. Xie, X. Tang, H. Tang, D.H. Johnson, G. Xiao, J.D. Minna, I.I. Wistuba

Development of methodology: Y. Xie, W. Lu, X. Tang, G. Xiao, I.I. Wistuba
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): Y. Xie, W. Lu, X. Tang, C. Moran, C. Behrens, J.A. Roth, S.G. Swisher, V.A. Papadimitrakopoulou, J.D. Minna, I.I. Wistuba
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): Y. Xie, W. Lu, S. Wang, X. Tang, H. Tang, Y. Zhou, D.H. Johnson, S.G. Swisher, G. Xiao, I.I. Wistuba

Writing, review, and/or revision of the manuscript: Y. Xie, W. Lu, S. Wang, H. Tang, J.A. Roth, Q. Zhou, D.H. Johnson, S.G. Swisher, J.V. Heymach, V.A. Papadimitrakopoulou, G. Xiao, J.D. Minna, I.I. Wistuba

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): Y. Xie, X. Tang, Q. Zhou, D.H. Johnson, S.G. Swisher, J.V. Heymach, I.I. Wistuba

Study supervision: Y. Xie, X. Tang, J.A. Roth, Q. Zhou, I.I. Wistuba

Acknowledgments

This work was supported by the NIH (5R01CA152301, P50CA70907, 5P30CA142543, 1R01GM115473, and 1R01CA172211), the Cancer Prevention and Research Institute of Texas (RP120732), and the National Key Research and Development Plan of China (2016YEE0103400) and National Natural Science Foundation of China (81572288).

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.

Received August 31, 2017; revised June 12, 2018; accepted September 27, 2018; published first October 4, 2018.

References

- Jemal A, Siegel R, Ward E, Hao Y, Xu J, Murray T, et al. Cancer statistics, 2008. *CA Cancer J Clin* 2008;58:71–96.
- Jemal A, Siegel R, Ward E, Murray T, Xu J, Thun MJ. Cancer statistics, 2007. *CA Cancer J Clin* 2007;57:43–66.
- Tsuboi M, Ohira T, Saji H, Miyajima K, Kajiwara N, Uchida O, et al. The present status of postoperative adjuvant chemotherapy for completely resected non-small cell lung cancer. *Ann Thorac Cardiovasc Surg* 2007;13:73–7.
- Strauss GM, Herndon J, Maddaus MA, Johnstone DW, Johnson EA, Watson DM, et al. Randomized clinical trial of adjuvant chemotherapy with paclitaxel and carboplatin following resection in stage Ib non-small cell lung cancer (NSCLC): report of cancer and leukemia group b (CALGB) protocol 9633. *J Clin Oncol* 2004;22:7019.
- Douillard JY, Rosell R, De Lena M, Carpagnano F, Ramlau R, Gonzales-Larriba JL, et al. Adjuvant vinorelbine plus cisplatin versus observation in patients with completely resected stage IB–IIIA non-small cell lung cancer (Adjuvant Navelbine International Trialist Association [ANITA]): a randomized controlled trial. *Lancet Oncol* 2006;7:719–27.
- Kato H, Ichinose Y, Ohta M, Hata E, Tsubota N, Tada H, et al. A randomized trial of adjuvant chemotherapy with uracil-tegafur for adenocarcinoma of the lung. *N Engl J Med* 2004;350:1713–21.
- The International Adjuvant Lung Cancer Trial Collaborative Group. 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, Rigas J, Johnston M, Butts C, et al. Vinorelbine plus cisplatin versus observation in resected non-small-cell lung cancer. *N Engl J Med* 2005;352:2589–97.
- Strauss GM, Herndon JE II, Maddaus MA, Johnstone DW, Johnson EA, Harpole DH, et al. Adjuvant paclitaxel plus carboplatin compared with observation in stage Ib non-small cell lung cancer: CALGB 9633 with the Cancer and Leukemia Group B, Radiation Therapy Oncology group, and North Central Cancer Treatment Group study groups. *J Clin Oncol* 2008;26:5043–51.
- Douillard JY. [Adjuvant chemotherapy for non-small cell lung cancer— which agents for which patients?]. *Rev Mal Respir* 2005;22:8S118–23.
- Olaussen KA, Mountzios G, Soria JC. ERCC1 as a risk stratifier in platinum-based chemotherapy for non-small-cell lung cancer. *Curr Opin Pulm Med* 2007;13:284–9.
- Olaussen KA, Dunant A, Fouret P, Brambilla E, Andre F, Haddad V, et al. DNA repair by ERCC1 in non-small-cell lung cancer and cisplatin-based adjuvant chemotherapy. *N Engl J Med* 2006;355:983–91.
- Shedden K, Taylor JM, Enkemann SA, Tsao MS, Yeatman TJ, Gerald WL, et al. Gene expression-based survival prediction in lung adenocarcinoma: a multi-site, blinded validation study. *Nat Med* 2008;14:822–7.
- Tang H, Xiao G, Behrens C, Schiller J, Allen J, Chow CW, et al. A 12-gene set predicts survival benefits from adjuvant chemotherapy in non-small cell lung cancer patients. *Clin Cancer Res* 2013;19:1577–86.
- Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983;70:41–55.

16. Rosenbaum PR, Rubin DB. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *Am Stat* 1985;39:33–8.
17. Ho D, Imai K, King G, Stuart E. Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Polit Anal* 2007;15:199–236.
18. Veldman-Jones MH, Brant R, Rooney C, Geh C, Emery H, Harbron CG, et al. Evaluating robustness and sensitivity of the NanoString technologies nCounter platform to enable multiplexed gene expression analysis of clinical samples. *Cancer Res* 2015;75:2587–93.
19. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc* 1958;53:457–81.
20. Collett D. *Modelling survival data in medical research*. Boca Raton, FL: Chapman & Hall/CRC Press; 2003.
21. Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival data and a diagnostic marker. *Biometrics* 2000;56:337–44.
22. Huang Y, Gilbert PB, Janes H. Assessing treatmentselection markers using a potential outcomes framework. *Biometrics* 2012;68:687–96.
23. Schroder MS, Culhane AC, Quackenbush J, Haibe-Kains B. survcomp: an R/Bioconductor package for performance assessment and comparison of survival models. *Bioinformatics* 2011;27:3206–8.
24. Berghmans T, Paesmans M, Meert AP, Mascaux C, Lothaire P, Lafitte JJ, et al. Survival improvement in resectable non-small cell lung cancer with (neo) adjuvant chemotherapy: results of a meta-analysis of the literature. *Lung Cancer* 2005;49:13–23.
25. Hotta K, Matsuo K, Ueoka H, Kiura K, Tabata M, Tanimoto M. Role of adjuvant chemotherapy in patients with resected non-small-cell lung cancer: reappraisal with a meta-analysis of randomized controlled trials. *J Clin Oncol* 2004;22:3860–7.
26. Sedrakyan A, Van Der Meulen J, O'Byrne K, Prendiville J, Hill J, Treasure T. Postoperative chemotherapy for non-small cell lung cancer: a systematic review and meta-analysis. *J Thorac Cardiovasc Surg* 2004;128:414–9.
27. Pignon JP, Tribodet H, Scagliotti GV, Douillard JY, Shepherd FA, Stephens RJ, et al. Lung adjuvant cisplatin evaluation: a pooled analysis by the LACE Collaborative Group. *J Clin Oncol* 2008;26:3552–9.