

# Using Gene Expression Ratios to Predict Outcome Among Patients With Mesothelioma

Gavin J. Gordon, Roderick V. Jensen, Li-Li Hsiao, Steven R. Gullans, Joshua E. Blumenstock, William G. Richards, Michael T. Jaklitsch, David J. Sugarbaker, Raphael Bueno

**Background:** We have recently demonstrated that simple ratios of the expression levels of selected genes in tumor samples can be used to distinguish among types of thoracic malignancies. We examined whether this technique could predict treatment-related outcome for patients with mesothelioma. **Methods:** We used gene expression profiling data previously collected from 17 mesothelioma patients with different overall survival times to define two outcome-related groups of patients and to train an expression ratio-based outcome predictor model. A Student's *t* test was used to identify genes among the two outcome groups that had statistically significant, inversely correlated expression levels; those genes were used to form prognostic expression ratios. We used a combination of several highly accurate expression ratios and cross-validation techniques to assess the internal consistency of this predictor model, quantitative reverse transcription-polymerase chain reaction of tumor RNA to confirm the microarray data, and Kaplan-Meier survival analysis to validate the model among an independent set of 29 mesothelioma tumors. All statistical tests were two-sided. **Results:** We developed an expression ratio-based test capable of identifying 100% (17/17) of the samples used to train the model. This test remained highly accurate (88%, 15/17) after cross-validation. A four-gene expression ratio test statistically significantly ( $P = .0035$ ) predicted treatment-related patient outcome in mesothelioma independent of the histologic subtype of the tumor. **Conclusions:** Gene expression ratio-based analysis accurately predicts treatment-related outcome in mesothelioma samples. This technique could impact the clinical treatment of mesothelioma by allowing the preoperative identification of patients with widely divergent prognoses. [J Natl Cancer Inst 2003;95:598-605]

Malignant pleural mesothelioma is an asbestos-related, lethal neoplastic disease of the pleura (median survival = 4-12 months) that can be subdivided into three major histologic subtypes: epithelial, mixed, and sarcomatoid (1-4). Compared with patients having the non-epithelial subtypes, patients with the epithelial subtype show a survival benefit from a variety of treatment strategies, including aggressive multimodality therapy (5-7). Currently, patients who present to our unit with unilateral mesothelioma without extrapleural invasion undergo complete surgical resection (i.e., extrapleural pneumonectomy) followed by chemoradiation. The 5-year survival for those patients who have Brigham stage I tumors (5) with an epithelial histology is 40%. However, there are no predictive factors, prognostic molecular markers, or genetic abnormalities other than histologic subtype to preoperatively identify these (or other) long-term survivors of pleural mesothelioma. In addition, established methods to predict outcome in mesothelioma that are based on

histologic appearance of the tumors are somewhat subjective, prone to human error, and ineffective for small patient cohorts or, in extreme cases, for individual patients (3,8,9).

Gene expression profiling using oligonucleotide microarrays holds promise for improving strategies for tumor classification as well as for predicting response to therapy and survival among cancer patients (10-16). Nevertheless, no clear consensus exists regarding which computational tools are optimal for the analysis of large gene expression profiling datasets, particularly when they are used to predict outcome. As a result, microarray-based research has not yet had a substantial impact on the clinical treatment of disease. Recently, we have shown that simple ratios of gene expression levels, using as few as four to six genes, are highly accurate in the diagnosis of cancer, and we hypothesized that this technique would be equally useful in additional clinical applications (17). To explore this possibility further, we examined whether gene expression profiling data (17) obtained from mesothelioma samples from patients with widely divergent survival times could be used to create an expression ratio-based test capable of predicting outcome among patients diagnosed with mesothelioma in a manner that is independent of the histologic subtype of the tumor.

## METHODS

### Mesothelioma Tumor Samples

Discarded mesothelioma surgical specimens ( $n = 60$ ) were freshly collected from patients who underwent extrapleural pneumonectomy for mesothelioma without preoperative treatment at Brigham and Women's Hospital (Boston, MA) between 1993 and 1999; all specimens were flash frozen (6). These 60 specimens were chosen from among the 150 available specimens because they contained greater than 50% tumor cell nuclei. The specimens were linked to clinical and outcome data and were provided without patient identifying information. Thirty-one of the specimens were previously profiled using microarrays (17); 17 of those samples were used to train an outcome predictor model (i.e., the training set) and were selected on the basis of linked clinical data in order to construct a sample cohort

*Affiliations of authors:* G. J. Gordon, W. G. Richards, M. T. Jaklitsch, D. J. Sugarbaker, R. Bueno, Division of Thoracic Surgery, Brigham and Women's Hospital, Harvard Medical School, Boston, MA; R. V. Jensen, J. E. Blumenstock, Department of Physics, Wesleyan University, Middletown, CT; L.-L. Hsiao (Department of Medicine), S. R. Gullans (Department of Neurology), Brigham and Women's Hospital, Harvard Medical School, Cambridge, MA.

*Correspondence to:* Raphael Bueno, M.D., Brigham and Women's Hospital, Division of Thoracic Surgery, 75 Francis St., Boston, MA 02115 (e-mail: rbueno@partners.org).

See "Notes" following "References."

*Journal of the National Cancer Institute*, Vol. 95, No. 8, © Oxford University Press 2003, all rights reserved.

with a wide range of survival times. The remaining 29 samples (i.e., the test set) were used only for quantitative reverse transcription–polymerase chain reaction (RT–PCR) analysis to validate the model (i.e., they had not been previously used in any other analysis). Studies using human tissues were approved by and conducted in accordance with the policies of the Institutional Review Board at Brigham and Women’s Hospital.

### Real-Time Quantitative RT–PCR

Real-time quantitative RT–PCR was performed using an SYBR–Green fluorometric-based detection system and equipment and reagents purchased from Applied Biosystems (Foster City, CA), according to protocols provided by the manufacturer as previously described (18). Total RNA (2  $\mu$ g) was isolated from each of the 29 tumors in the test set with the use of TRIzol reagent (Invitrogen Life Technologies, Carlsbad, CA), reverse-transcribed into complementary DNA (cDNA) with the use of Taq–Man Reverse Transcription reagents (Applied Biosystems), and quantified using all controls recommended by the manufacturer. Primer sequences (synthesized by Invitrogen Life Technologies) used for RT–PCR were as follows (forward and reverse, respectively): L6 (5′-TTCCATTCCACAATGTGCTT-3′ and 5′-GGCCAGTGGAACTACACCTT-3′), KIAA0977 (5′-AACCGAAGCCTAACCTGAGA-3′ and 5′-GTCATTTTGGGAGCAGGTTT-3′), GDIA1 (5′-AGAAGCAGTCGTTTGTGCTG-3′ and 5′-TGTACTTCATGCCGGACACT-3′), and CTHBP (5′-ATCTGAAGTTTGGGGTTCGAG-3′ and 5′-TCTCTCCAGGACCTTCCTA-3′). PCR amplification of cDNA was performed using an Applied Biosystems 5700 Sequence Detector and default thermal cycling parameters, as previously described (18). No-template (i.e., negative) controls that contained water instead of template were run in multiple wells on every reaction plate. An automatically calculated melting point dissociation curve generated after every assay was examined to ensure the presence of a single PCR species and the lack of primer-dimer formation in each well. The comparative CT equation (Applied Biosystems) describes the exponential nature of PCR-based amplification and was used, with minor modifications, to obtain quantitative values for gene expression ratios in all samples. The “CT” term stands for the fractional PCR cycle at which the quantity of the amplified product reaches a predetermined threshold. The comparative CT equation states that the expression level of a gene in a given sample, normalized within the sample to an endogenous reference gene and relative to the expression level of the same gene in another sample (i.e. an arbitrarily chosen calibrator sample), can be represented as  $2^{-\Delta\Delta CT}$  where  $\Delta\Delta CT = [\Delta CT_{(\text{sample } x)}] - [\Delta CT_{(\text{calibrator sample})}]$  and  $\Delta CT = [CT_{(\text{target gene})}] - [CT_{(\text{reference gene})}]$ . Calculation of an expression ratio using data from two rationally selected genes in any single sample obviates the need for a calibrator sample and a reference gene for standardization when using different amounts of starting template. Therefore, to form expression ratios of two genes in a single sample, we simply presented the expression level of one gene relative to the expression level of the other gene. In this case, the  $\Delta\Delta CT$  value in the comparative CT equation was expressed as  $[CT_{(\text{gene } 1)} - CT_{(\text{gene } 2)}]$ .

### Data and Statistical Analysis

Expression profiling raw data and supplemental information are available at our Web site, <http://www.chestsurg.org>, under “Publications.” To create an expression ratio-based outcome

predictor model in mesothelioma, we used previously collected microarray data (17) to identify genes whose expression levels could be used to discriminate among tumors that came from patients with considerably different survival times. First, we ranked all 31 tumors according to patient survival (irrespective of the histologic subtype of the tumor) and determined the survival times corresponding to the 25<sup>th</sup> and 75<sup>th</sup> percentile of the cohort (6 months and 17 months, respectively). Then, using these times as a cutoff, we compared gene expression patterns between two groups of mesothelioma samples (i.e., the training set,  $n = 17$ ): those that were obtained from patients who had survived for at least 17 months (i.e., good outcome,  $n = 8$ ) and those that were obtained from patients who had survived for 6 months or less (i.e., poor outcome,  $n = 9$ ). The most accurate expression ratio-based predictor model developed in the training set was subsequently tested in an independent cohort of samples (i.e. the test set,  $n = 29$ ). A two-sided Student’s (parametric)  $t$  test was used for pairwise comparisons of average gene expression levels among multiple groups, and the significance analysis of microarrays (SAM) algorithm (19) was used to estimate the false-discovery rate. To find discriminating genes, we searched all of the genes represented on the Affymetrix U95A 12000-gene microarray (17) to identify those whose average expression levels differed statistically significantly and by at least twofold between good-outcome and poor-outcome training set tumors. To minimize the effects of background noise, we further refined this list of distinguishing genes by requiring that the mean expression levels (i.e., Affymetrix “average difference” values) be greater than 500 in at least one of the two sample sets, similar to filtering criteria used in previous studies (17). Kaplan–Meier curves were used to estimate patient survival (defined as the time, in months, from surgery until death) among each group of mesothelioma patients. The log-rank test was used to statistically assess differences among multiple survival curves. A Cox proportional hazards regression model was used for multivariate analysis to identify coefficients that best described the effect of a given variable on censored survival data. (The data conformed to Cox proportional hazards assumptions.) We used the Efron method [described in (20)] to handle observations with identical survival times. Individual  $P$  values reported for multivariate analysis were calculated by considering the Wald statistics of the individual parameters in the combined model. Hazard ratios (HRs) and 95% confidence intervals (CIs) are expressed as the exponentiated coefficient values and are interpretable as multiplicative effects on the hazard. The likelihood ratio test, the Wald test, and the score (i.e., log-rank) test were used to test the null hypothesis that all of the coefficients are zero. The “leave-one-out” method of cross-validation (16) was used to assess internal consistency of the predictor model. Classification accuracy in the test set was determined by using Fisher’s exact test (i.e., a  $2 \times 2$  contingency table). All differences were considered statistically significant if  $P$  was less than .05. Data from three highly accurate gene expression ratios were combined by calculating the geometric mean,  $(R_1 R_2 R_3)^{1/3}$ , where  $R_n$  represents a single ratio value. This approach is the mathematical equivalent of taking the average of  $[\log_2(R_1), \log_2(R_2), \log_2(R_3)]$ , which gives equal weight to fold-changes in ratios that are of identical magnitude but are in opposite directions. All calculations and statistical comparisons were generated with the use of S-PLUS software (20). All statistical tests were two-sided.

## RESULTS

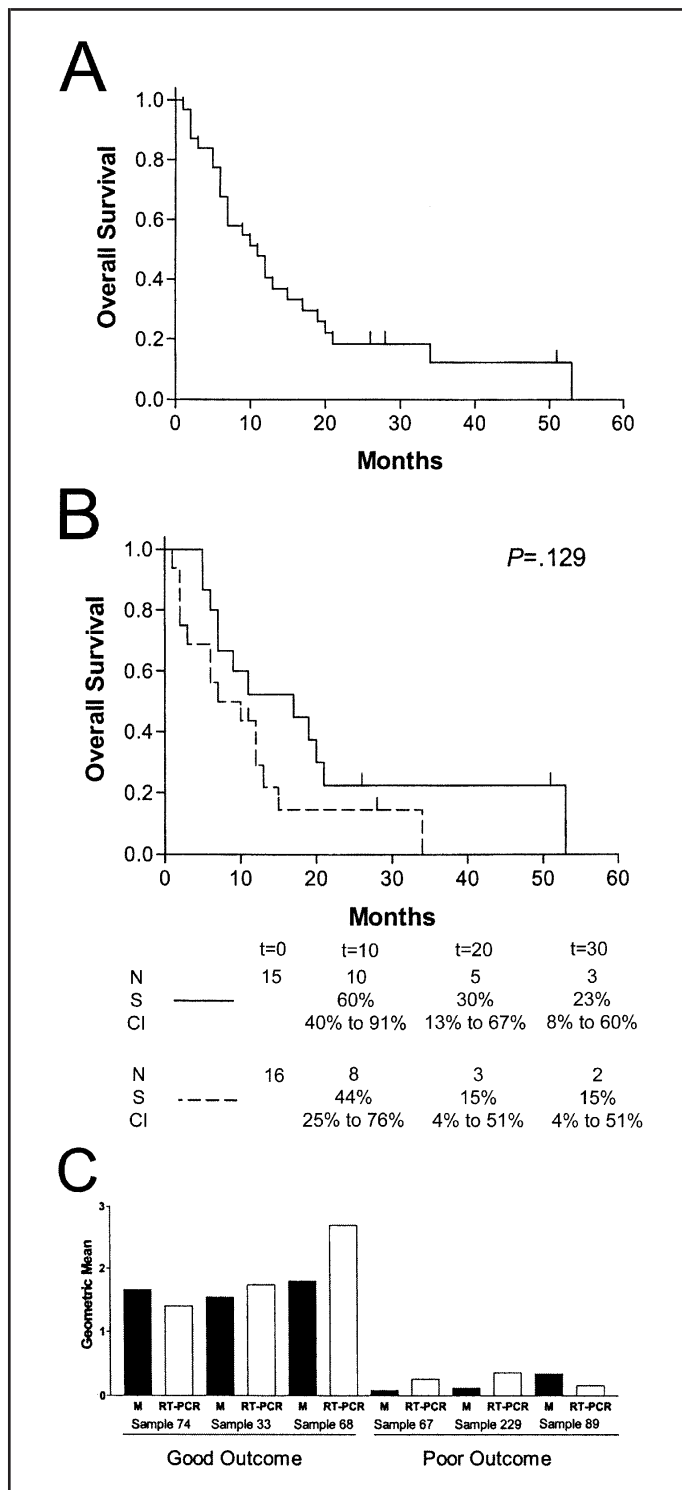
### Identification of Prognostic Molecular Markers in Mesothelioma Tumor Samples

We previously analyzed gene expression levels among a representative cohort of 31 mesothelioma tumors that were obtained from patients at pneumonectomy (17). In that study, we used differences in global gene expression patterns between the 31 mesothelioma samples and a large number of lung adenocarcinoma samples to develop a ratio-based predictor model to distinguish between these two types of neoplasms. In the current study, we used these data to define molecular markers in mesothelioma that are associated with tumors from patients with widely divergent survival times. The estimated median survival of patients from whom those tumors were obtained (11 months; Fig. 1, A) and the histologic distribution of the tumors were similar to what we have observed among the mesothelioma patients in our practice (6). The histologic subtype of the tumor was not associated with patient survival ( $P = .129$ , log-rank test; Fig. 1, B), even though the estimated median survival of patients who had tumors of the epithelial subtype (17 months) was longer than that of patients who had tumors of the non-epithelial subtypes (8.5 months).

Clinical data for all tumor samples are presented in Table 1. We identified a total of 46 putative prognostic genes in the analysis of the training set of samples, with an estimated false-discovery rate of 10%–20%. Table 2 lists, for each outcome group of tumors in the training set, the 10 overexpressed genes that had the lowest  $P$  values.

### Using Gene Expression Ratios to Predict Outcome

We chose the four genes that were the most statistically significantly overexpressed in each outcome group of tumors (Table 2) to examine whether their expression ratios could accurately classify the 17 samples used to train the model with respect to group membership. We calculated a total of 16 possible expression ratios per sample by dividing the expression value of each of the four genes (i.e., those encoding selenium-binding protein [SBP], KIAA0977 protein, an expressed sequence tag [EST] similar to the L6 tumor antigen, and leukocyte antigen-related protein [LAR]) that were expressed at relatively higher levels in good-outcome samples than in poor-outcome samples by the expression value of each of the four genes (i.e., those encoding cytosolic thyroid hormone-binding protein [CTHBP], calgizzarin, insulin-like growth factor-binding protein-3 [IGFBP-3], and guanosine diphosphate-dissociation inhibitor 1 [GDIA1]) that were expressed at relatively higher levels in poor-outcome samples than in good-outcome samples. Samples with ratio values of greater than 1 were predicted to be good outcome and those with ratio values of less than 1 were predicted to be poor outcome. The individual gene pair ratios that predicted the group membership of the training set samples with the highest accuracy were chosen for further study. Five such ratios (i.e., KIAA0977 protein/IGFBP-3, KIAA0977 protein/GDIA1, L6-related EST/CTHBP, L6-related EST/GDIA1, and LAR/GDIA1) each independently classified 15 (88%) of the 17 samples used to train the model. To incorporate the predictive accuracy of multiple ratios, we calculated the geometric mean (see "Methods" section) for all possible three-ratio combinations (formed using these five ratios) and found that we could classify the training samples with an accuracy that met or exceeded that of any of the single gene-pair ratios (average accuracy = 94%,



**Fig. 1.** Kaplan–Meier survival predictions for mesothelioma patients and verification of microarray data. **A)** Overall survival for all 31 patients from which the training set of tumor samples was chosen. The estimated median survival time for entire cohort was 11 months. **B)** Overall survival based on the histologic subtype of the tumor. The estimated median survival of epithelial subtype tumors (solid line) was 17 months, and the estimated median survival of non-epithelial subtype tumors (dashed line) was 8.5 months. N and S indicate the number of patients at risk and the Kaplan–Meier estimate of overall survival, respectively, at the indicated time points. CI = 95% confidence interval for the Kaplan–Meier survival estimate. **C)** Comparison of geometric mean values for expression ratio data obtained for six randomly chosen samples (three from the good-outcome group and three from the poor-outcome group) by using quantitative reverse transcription–polymerase chain reaction (RT–PCR; this study) and microarrays [M; (17)].



**Table 1.** Clinical characteristics of malignant pleural mesothelioma tumors\*

Training set							Test set						
Tumor sample No.	Patient age at diagnosis, y	Sex	Tumor histology	Tumor stage†	Patient survival, mo.	Patient status‡	Tumor sample No.	Patient age at diagnosis, y	Sex	Tumor histology	Tumor stage†	Patient survival, mo.	Patient status‡
72	46	M	mixed	2	53	3	169	46	M	ept	2	7	3
74	40	F	ept	1	51	2	146	67	M	ept	2	7	3
90	48	M	ept	2	28	2	219	39	M	ept	2	6	1
2	44	F	ept	2	26	2	104	40	M	ept	2	5	3
68	61	M	ept	2	21	3	110	64	M	ept	2	5	3
33	60	F	ept	2	20	3	112	31	M	ept	2	55	3
109	62	M	ept	2	19	3	165	51	M	ept	2	27	2
76	67	M	ept	1	17	3	5	51	M	ept	2	8	3
130	55	M	mixed	2	6	3	148	51	M	ept	2	17	3
166	66	M	sarc	2	6	3	96	40	M	ept	2	1	3
67	49	F	ept	2	6	3	134	56	M	ept	2	1	4
229	33	F	ept	2	5	3	216	43	F	ept	2	8	1
6	39	M	ept	2	5	3	208	63	F	ept	2	7	1
89	55	M	mixed	2	3	3	224	68	F	ept	2	6	1
133	69	M	mixed	2	2	3	225	35	F	ept	2	42	2
114	51	M	mixed	2	2	3	163	68	F	ept	2	25	1
159	62	M	sarc	2	2	3	235	46	M	mixed	2	24	3
							206	45	M	mixed	2	45	2
							107	69	M	mixed	2	16	3
							302	55	M	mixed	2	13	3
							161	59	M	mixed	2	12	3
							220	71	M	mixed	2	12	3
							217	57	M	mixed	1	5	1
							150	58	M	mixed	2	3.6	3
							44	57	M	mixed	2	2	4
							222	57	M	mixed	2	1	U
							154	56	F	mixed	2	9	3
							70	57	M	sarc	2	8	3
							228	73	M	sarc	2	4	3

\*M = male; F = female; ept = epithelial; sarc = sarcomatoid.

†Tumor stage was determined by using the Brigham staging system (5): 1 = negative lymph nodes and margins, 2 = positive intrapleural lymph nodes and/or margins, 3 = positive ipsilateral mediastinal lymph nodes and/or diaphragmatic and/or pericardial invasion.

‡Determined at the time of anonymization when tissues were obtained from the divisional tumor bank. 1 = alive without mesothelioma, 2 = alive with mesothelioma, 3 = dead from mesothelioma, 4 = dead from other causes, and U = unknown.

range = 88%–100%). For further analysis, we chose one of the two three-ratio combinations that correctly classified 100% (17/17) of the training samples. A total of four genes (those encoding KIAA0977 protein, GDIA1, the L6-related EST, and CTHBP) were used in the following combinations in this three-ratio test: KIAA0977/GDIA1, L6-related EST/CTHBP, and L6-related EST/GDIA1.

### Verification of Microarray Data

We used quantitative RT-PCR to determine the relative expression levels of all four prognostic genes (L6, GDIA1, CTHBP, and that encoding the KIAA0977 protein) among three randomly chosen samples from each outcome group in the training set: samples 74, 33, and 68 from the good-outcome group and samples 89, 229, and 67 from the poor-outcome group. We then used the values obtained from RT-PCR quantitation to calculate the three individual expression ratios previously used to predict outcome (i.e., KIAA0977/GDIA1, L6/CTHBP, and L6/GDIA1) for those six samples. Finally, we calculated the geometric mean of these three ratios and compared the magnitude and direction (i.e., >1 or <1) of that value for each of the six samples to that obtained previously by using microarray analysis. We found that outcome group classifications using the three-ratio geometric means calculated with data obtained from both

types of analyses were in perfect agreement for all six samples (Fig. 1, C).

### Validation of the Model

We used a “leave-one-out” cross-validation technique (16) to assess the internal variation of the three-ratio predictor model. This technique reduces the potential for over-fitting of the model due to an information leak because the sample that is left out of each distinct training set is not involved in the selection of predictor genes (16). We analyzed 17 different training sets by withholding one of the 17 samples to construct a new expression ratio-based classifier exactly as described above and then predicting the class (either good or poor outcome) of the withheld sample. For each training set, we first identified predictor genes using the original filtering criteria. As before, the four genes that were the most statistically significantly overexpressed in each outcome group (eight genes total) were used to calculate a total of 16 possible expression ratios whose classification accuracies were assessed in the new training set. We then used the geometric mean value for the three most accurate ratios in the new training set to classify the remaining sample. This process was repeated sequentially for all 17 samples. We found that 15 (88%) of the 17 samples were correctly identified by this analysis. In all 17 training sets, two of the four genes used in the original model

**Table 2.** Prognostic genes in mesothelioma\*

Description	GenBank accession No.	Expression ratio <sup>†</sup>	<i>P</i> value <sup>‡</sup>
Expressed at relatively higher levels in good-outcome tumors than in poor-outcome tumors			
Selenium-binding protein	U29091	2.8	.0033
KIAA0977 protein	AB023194	2.1	.0065
EST (similar to L6 tumor antigen)	A1445461	3.0	.0073
Leukocyte antigen-related protein	Y00815	2.0	.0077
Hyaluronan synthase	D84424	6.0	.0094
Complement control protein factor I	Y00318	3.6	.0103
EST	AL049963	3.7	.0103
Histone H2B	AJ223352	3.5	.0142
Semaphorin E	AB000220	2.3	.0181
Cytochrome b5	L39945	2.5	.0182
L6 tumor antigen	M90657	2.8	.0256
KIAA0303 protein	AB002301	2.1	.0257
Expressed at relatively higher levels in poor-outcome tumors than in good-outcome tumors			
Cytosolic thyroid hormone-binding protein	M26252	0.38	.0013
Calgizzarin	D38583	0.43	.0041
Insulin-like growth factor-binding protein-3 (IGFBP-3) <sup>§</sup>	M35878	0.35	.0046
GDP-dissociation Inhibitor 1	X69550	0.47	.0063
Transgelin	M95787	0.33	.0068
KIAA0991 protein	AB023208	0.49	.0069
Zyxin	X95735	0.43	.0105
EST	AA976838	0.40	.0131
Carboxyl terminal LIM domain protein	U90878	0.49	.0132
IGFBP-3 <sup>§</sup>	M35878	0.30	.0135
Plectin	U53204	0.39	.0169
Non-muscle alpha-actinin	M95178	0.40	.0215

\*EST = expressed sequence tag; GDP = guanosine diphosphate; LIM = acronym for Lin-11, Isl-1, and Mec-3.

<sup>†</sup>Average expression level in good outcome samples/average expression level in poor outcome samples.

<sup>‡</sup>*P* values were obtained using a two-tailed Student's *t* test to identify genes whose average expression levels were statistically significantly different between groups of tumor samples from patients with good and poor outcome, respectively, as detailed in the "Methods" section.

<sup>§</sup>IGFBP-3 is listed twice in the lower portion of the table because this gene is represented by multiple Affymetrix probe sets. Each microarray probe set is complementary to a unique DNA sequence and, because the probe sets are composed of relatively short sequences, a certain number of them are capable of specific hybridization to different regions of the same gene. When this occurs, it provides a direct, internal measure of confidence in relevant gene expression levels. In addition to the duplication of IGFBP-3, note that both the L6 gene and an EST with high sequence homology to L6 are represented in the table.

(i.e., those encoding KIAA0977 protein and the L6-related EST) were present among the final list of eight genes. The final three-ratio test used to classify each left-out sample included both of these genes in one training set, one of these genes in 12 training sets, and neither of these genes in four training sets.

### Verification of Expression Level Ratios as Outcome Predictors

Finally, we tested the ability of expression ratios to predict patient outcome among a new cohort of mesothelioma tumor samples that had not been subjected to microarray analysis ( $n = 29$ , the test set; Table 1). The estimated median survival of patients from whom those tumors were obtained (12 months; Fig. 2, A) and the histologic distribution of the test set tumors were similar to what we observe among mesothelioma patients in our practice (6). The histologic subtype of the tumors in the new cohort of samples was not associated with patient survival ( $P = .345$ , log-rank test; Fig. 2, B). We used quantitative RT-PCR to determine the relative expression levels of the four predictor genes among the test set samples and calculated the geometric mean of three prognostic expression ratios: KIAA0977/GDIA1, L6/CTHBP, and L6/GDIA1. Samples with geometric means of greater than 1 and less than 1 were assigned to good-outcome and poor-outcome groups, respectively; 11 samples were assigned to the good-outcome group, and 18 samples were assigned to the poor-outcome group. The number of test set samples "correctly" classified was estimated by using the median survival (12 months) of the entire cohort as a cutoff to form

two groups: the relatively good-outcome group ( $>12$ -month survival) and the relatively poor-outcome group ( $\leq 12$ -month survival). When we considered only the 17 samples that came from patients who had died from malignant pleural mesothelioma (status 3; Table 1), we found that the exact same number of test set samples were classified correctly in this analysis (88%, 15/17;  $P = .0099$ , Fisher's exact test) as were classified correctly in our analysis of the training set samples. To include all samples in an assessment of the model, we performed Kaplan-Meier survival analysis using expression ratio predictions made for the test set of samples. The estimated median survival for the good-outcome group (36 months) was more than fivefold higher than that for the poor-outcome group (7 months). In addition, we found that the three-ratio geometric mean model statistically significantly predicted outcome in the test set of samples ( $P = .0035$ , log-rank test; Fig. 2, C). Because it has been demonstrated in very large sample cohorts that patients whose tumors have epithelial histologies generally enjoy statistically significantly longer disease-free survival than patients whose tumors have non-epithelial histologies (21), we used multivariate analysis to examine whether the results we obtained using expression ratios were independent of the histologic subtype of the tumor. By fitting a Cox proportional hazards regression model, we found that the three-ratio geometric mean value statistically significantly predicted outcome (HR = 4.6, 95% CI = 1.5 to 14.8;  $P = .0094$ ), whereas the histologic subtype of the tumor did not (HR = 1.2, 95% CI = 0.45 to 3.1;  $P = .75$ ). Furthermore, the results of a likelihood ratio test ( $P = .011$ ), a

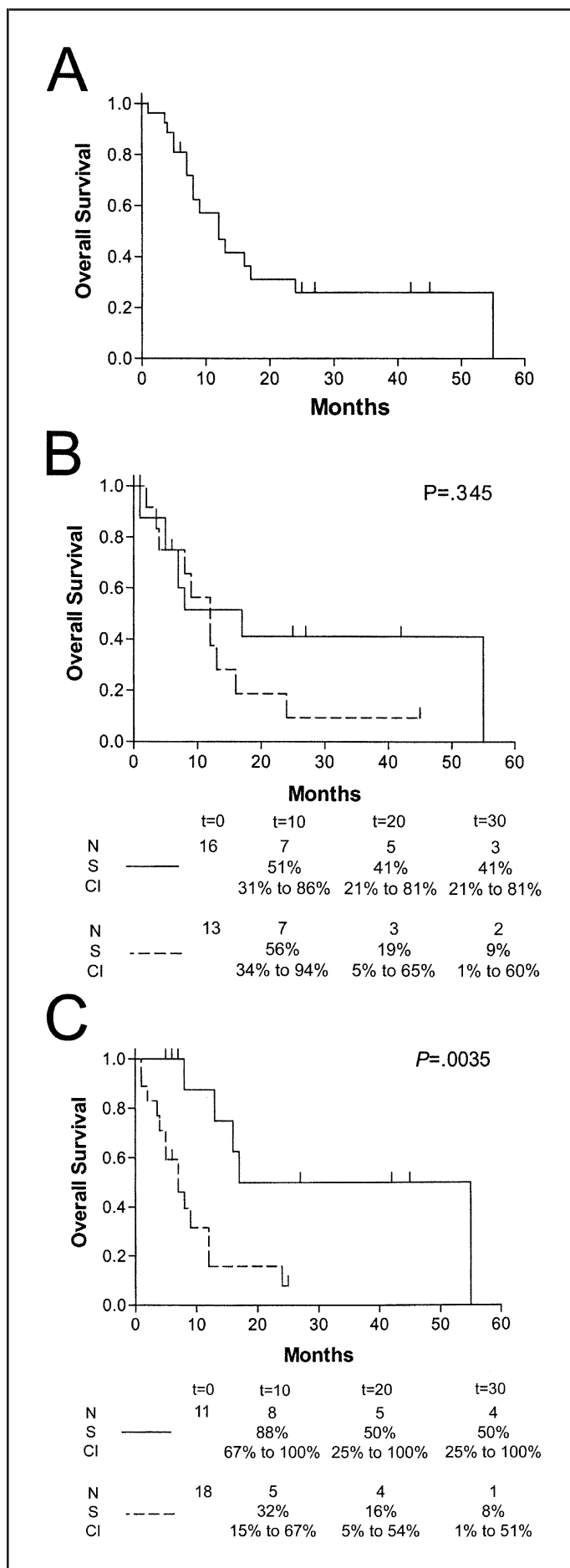
Wald test ( $P = .025$ ), and a score (i.e., log-rank) test ( $P = .013$ ) were all in close agreement, leading to the rejection of the null hypothesis, i.e., that at least one of the regression coefficients is not zero. These results demonstrate that expression ratios can predict outcome in mesothelioma independently of the histologic subtype of the tumor in an independent set of samples, indicating that the gene expression ratio method is a better prognostic tool than histology.

## DISCUSSION

Current methods of prognosis in mesothelioma include determining tumor stage and histology at the time of surgery. However, these techniques are not completely reliable, and accurate staging usually requires extensive surgery (3,8,9). Recently, we discovered that simple ratios of gene expression levels can be used to accurately diagnose cancer (17) while successfully avoiding many of the shortcomings that preclude the use of other analytical microarray techniques in wider clinical applications (10,22). In this study, we describe a technique that used expression data from four genes to independently predict outcome in mesothelioma patients who had undergone extrapleural pneumonectomy followed by standard chemoradiation therapy. Although this analysis used expression data from only four genes, the expression ratio technique can easily incorporate data from larger numbers of genes when required to achieve acceptable levels of accuracy. To our knowledge, ours is the first study to use gene expression profiling techniques to identify treatment-related prognostic markers for a human cancer that can be used to develop an outcome predictor model and to validate the model in an independent cohort by using a simple data acquisition platform such as RT-PCR. Other investigators have tested outcome predictor models in independent samples (16,23,24), but those studies continue to be restricted in their clinical applicability because they rely on data gathered from relatively large numbers of genes, costly data acquisition platforms (i.e., microarrays), and sophisticated algorithms and/or software and are unable to analyze a sample independently and without reference to other samples.

The prognostic tool described herein could dramatically influence the current clinical treatment of mesothelioma by allowing the identification of those patients who are unlikely to respond to conventional treatment modalities, thus sparing them from radical surgery. It is currently our practice to obtain a tissue diagnosis before recommending therapy for patients with mesothelioma, but the absence of suitable prognostic molecular markers makes it difficult to assign optimal treatments or to investigate new modalities solely on the basis of tumor histology. The

**Fig. 2.** Independent validation of the four-gene expression ratio model. **A)** Overall survival for 29 mesothelioma patients whose tumors composed the test set. The estimated median survival for this cohort was 12 months. **B)** Overall survival of patients whose tumors composed the test set according to the histologic subtype of the tumor. The median survival for patients with epithelial subtype tumors (**solid line**) was 17 months, and the median survival for patients with non-epithelial subtype tumors (**dashed line**) was 12 months. **C)** Overall survival in the test set of samples for good-outcome (**solid line**; median survival = 36 months) and poor-outcome (**dashed line**; median survival = 7 months) groups as defined by the four-gene expression ratio model that used only reverse transcription-polymerase chain reaction for data acquisition. N and S indicate the number of patients at risk and the Kaplan-Meier estimate of overall survival, respectively, at the indicated time points. CI = 95% confidence interval for the Kaplan-Meier survival estimate.



results of this work, if confirmed prospectively in a larger patient population, should prove helpful in the development of meaningful clinical trials for patients with mesothelioma. We hypothesize that patients whose tumors are analyzed using gene expression ratios and are predicted to have relatively poor outcomes are excellent candidates for neo-adjuvant chemotherapy protocols because they are unlikely to benefit from surgery followed by chemotherapy and radiation (i.e., current standard treatment for mesothelioma), whereas patients predicted to have relatively good outcomes are more likely to enjoy long-term survival after conventional surgery and adjuvant chemotherapy.

The use of gene expression ratios to predict outcome in cancer patients overcomes several major obstacles that hinder the clinical use of microarray data. Unlike other widely accepted supervised learning techniques with similar predictive accuracy (10,16,22–24), the expression ratio method generates a simple numeric measure that can be used to predict clinical outcome by using a single biopsy specimen. The gene expression ratio method, by virtue of the fact that it generates a ratio, 1) negates the need for a third reference gene when determining expression levels, 2) is independent of the platform used for data acquisition, 3) requires only small quantities of RNA (as little as 10 pg when using RT-PCR), 4) does not explicitly require the coupling of transcription to translation for chosen genes or rely on subjective measures of expression (as is the case for immunohistochemical analysis), 5) permits analysis of individual samples without reference to additional training samples whose data were acquired on the same platform, and 6) can use any reliable method to quantify expression levels, including quantitative RT-PCR, cDNA and oligonucleotide microarrays, serial analysis of gene expression and, perhaps, enzyme-linked immunosorbent assays for encoded proteins. For these reasons, the gene expression ratio method is more likely to find immediate use in clinical settings because it confers several advantages that are lacking in other equally accurate techniques, such as standard linear discriminant analysis. In fact, the expression ratio technique can be thought of as a special case of linear discriminant analysis wherein the threshold, while perhaps not optimal, remains constant across experiments but still results in highly accurate classification.

We believe that attempts to bridge the gap between expression profiling studies in cancer and meaningful clinical applications should follow the general spirit of Occam's razor, a principle according to which "among a set of otherwise equal models, choose the simplest." Although other microarray-based predictor models in cancer may use relatively small numbers of genes to accurately predict outcome (16,21,24,25), those approaches continue to be limited in their clinical applicability. Furthermore, it has yet to be determined whether those approaches can use relatively low-cost and widely available data acquisition platforms such as RT-PCR and still allow statistically significant survival predictions. In this study we have shown that expression ratios can be useful in predicting prognosis in mesothelioma. In other clinical scenarios, the differences in gene expression patterns between groups to be distinguished may be more subtle, thus necessitating the modification of the filtering criteria used to select potential predictor genes. In addition, a rationally chosen threshold value other than 1 may be indicated when the baseline levels of expression of the predictor genes are very different. Nevertheless, it is likely that the ex-

pression ratio technique will find additional uses in the clinical management of other cancers and diseases. For example, using previously published data, we created ratio-based tests that use small numbers of genes that can be used to diagnose localized prostate cancer and predict clinical outcome in breast cancer (Gordon GJ, Loughlin KR, Powell MH, Sugarbaker DJ, Bueno R: unpublished data) and to predict clinical outcome in lung cancer (Gordon GJ, Richards WG, Sugarbaker DJ, Jaklitsch MJ, Bueno R: unpublished data).

## REFERENCES

- Pass H. Malignant pleural mesothelioma: surgical roles and novel therapies. *Clin Lung Cancer* 2001;3:102–17.
- Aisner J. Diagnosis, staging, and natural history of pleural mesothelioma. In: Aisner J, Arriagada R, Green MR, Martini N, Perry MC, editors. *Comprehensive textbook of thoracic oncology*. Baltimore (MD): Williams and Wilkins; 1996. p. 779–85.
- Ong ST, Vogelsang NJ. Current therapeutic approaches to unresectable (primary and recurrent) disease. In: Aisner J, Arriagada R, Green MR, Martini N, Perry MC, editors. *Comprehensive textbook of thoracic oncology*. Baltimore (MD): Williams and Wilkins; 1996. p. 799–814.
- Peto J, Hodgson JT, Matthews FE, Jones JR. Continuing increase in mesothelioma mortality in Britain. *Lancet* 1995;345:535–9.
- Sugarbaker DJ, Flores RM, Jaklitsch MT, Richards WG, Strauss GM, Corson JM, et al. Resection margins, extrapleural nodal status, and cell type determine postoperative long-term survival in trimodality therapy of malignant pleural mesothelioma: results in 183 patients. *J Thorac Cardiovasc Surg* 1999;117:54–65.
- Sugarbaker DJ, Garcia JP, Richards WG, Harpole DH Jr, Healy-Baldini E, DeCamp MM Jr, et al. Extrapleural pneumonectomy in the multimodality therapy of malignant pleural mesothelioma. Results in 120 consecutive patients. *Ann Surg* 1996;224:288–94.
- Sugarbaker D, Strauss GM, Lynch TJ, Richards W, Mentzer SJ, Lee TH, et al. Node status has prognostic significance in the multimodality therapy of diffuse, malignant mesothelioma. *J Clin Oncol* 1993;11:1172–8.
- Corson JM, Renshaw AA. Pathology of mesothelioma. In: Aisner J, Arriagada R, Green MR, Martini N, Perry MC, editors. *Comprehensive textbook of thoracic oncology*. Baltimore (MD): Williams and Wilkins; 1996. p. 757–78.
- Ordenez NG. The value of antibodies 44–36A, SM3, HBME-1, and thrombomodulin in differentiating epithelial pleural mesothelioma from lung adenocarcinoma. *Am J Surg Pathol* 1997;21:1399–1408.
- Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 1999;286:531–7.
- Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. *Nature* 2000;406:747–52.
- Hedenfalk I, Duggan D, Chen Y, Radmacher M, Bittner M, Simon R, et al. Gene expression profiles in hereditary breast cancer. *N Engl J Med* 2001;344:539–48.
- Khan J, Wei JS, Ringner M, Saal LH, Ladanyi M, Westermann F, et al. Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks. *Nat Med* 2001;7:673–9.
- Welsh JB, Sapinoso LM, Su AI, Kern SG, Wang-Rodriguez J, Moskaluk CA, et al. Analysis of gene expression identifies candidate markers and pharmacological targets in prostate cancer. *Cancer Res* 2001;61:5974–8.
- Dhanasekaran SM, Barrette TR, Ghosh D, Shah R, Varambally S, Kurachi K, et al. Delineation of prognostic biomarkers in prostate cancer. *Nature* 2001;412:822–6.
- van't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002;415:530–6.
- Gordon GJ, Jensen RV, Hsiao LL, Gullans SR, Blumenstock JE, Ramaswami S, et al. Translation of microarray data into clinically relevant cancer diagnostic tests using gene expression ratios in lung cancer and mesothelioma. *Cancer Res* 2002;62:4963–7.
- Shi L, Ho J, Norling LA, Roy M, Xu Y. A real time quantitative PCR-based method for the detection and quantification of simian virus 40. *Biologicals* 1999;27:241–52.



- (19) Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* 2001; 98:5116–21.
- (20) Venables WN, Ripley BD. *Modern applied statistics with S-Plus*. New York (NY): Springer; 1997.
- (21) Sugarbaker DJ, Liptay MJ. Therapeutic approaches in malignant mesothelioma. In: Aisner J, Arriagada R, Green MR, Martini N, Perry MC, editors. *Comprehensive Textbook of Thoracic Oncology*. Baltimore (MD): Williams and Wilkins; 1996. p. 786–98.
- (22) Shipp MA, Ross KA, Tamayo P, Weng AP, Kutok JL, Aguiar RC, et al. Diffuse large B-cell lymphoma outcome prediction by gene expression profiling and supervised machine learning. *Nat Med* 2002;8:68–74.
- (23) Rosenwald A, Wright G, Chan WC, Connors JM, Campo E, Fisher RI, et al. The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *N Engl J Med* 2002;346:1937–47.
- (24) Beer DG, Kardia SL, Huang CC, Giordana TJ, Levin AM, Misek DE, et al. Gene-expression profiles predict survival of patients with lung adenocarcinoma. *Nat Med* 2002;8:816–24.
- (25) Pomeroy SL, Tamayo P, Gaasenbeek M, Sturla LM, Angelo M, McLaughlin ME, et al. Prediction of central nervous system embryonal tumor outcome based on gene expression. *Nature* 2002;415:436–42.

## NOTES

Partly funded by grants to R. Bueno from the Brigham Surgical Group Foundation and The Milton Fund of Harvard Medical School, and by Public Health Service grant DK58849 (to S. R. Gullans) from the National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Department of Health and Human Services. R. Bueno is a recipient of the Mesothelioma Applied Research Foundation (MARF) 2001 grant.

Manuscript received July 23, 2002; revised February 10, 2003; accepted February 21, 2003.