Four-Gene Expression Ratio Test for Survival in Patients Undergoing Surgery for Mesothelioma

Gavin J. Gordon, Lingsheng Dong, Beow Y. Yeap, William G. Richards, Jonathan N. Glickman, Heather Edenfield, Madhubalan Mani, Richard Colquitt, Gautam Maulik, Branden Van Oss, David J. Sugarbaker, Raphael Bueno

Background Malignant pleural mesothelioma has few effective treatments, one being cytoreductive surgery. We previously developed a gene ratio test to predict outcome of malignant pleural mesothelioma patients undergoing surgery. In this study, we investigated the predictive value and technical assay performance of this test in patients with malignant pleural mesothelioma.

Methods Clinical data were obtained prospectively from 120 consecutive patients with malignant pleural mesothelioma who were scheduled for debulking surgery at one institution. Specimens were obtained at surgery or by pleural biopsy examination. Expression data for four genes were collected from tumor specimens, and three ratios of gene expression (TM4SF1/PKM2, TM4SF1/ARHGDIA, and COBLL1/ARHGDIA) were determined by quantitative reverse transcriptase–polymerase chain reaction. Patients were assigned to good or poor outcome groups by the gene ratio test. Survival was estimated by the Kaplan–Meier method and the log-rank test in univariate analyses. A multivariable Cox proportional hazards model was used to control for prognostic factors. Technical robustness was determined by using up to 30 specimens per patient, two biopsy techniques, and two performance sites. All statistical tests were two-sided.

Results The test predicted overall survival ($P < .001$) and cancer-specific survival ($P = .007$) in univariate analysis and overall survival in multivariable analysis (hazard ratio for death = 2.09, 95% confidence interval [CI] = 1.27 to 3.45, $P = .004$). The test was reproducible within patients and repeatable between two determinations for specimens with widely varying tumor cell contents. Repeatability between two determinations was 88.5% (95% CI = 84.0% to 92.2%) or, when technically unacceptable test values were excluded, 91.9% (95% CI = 87.4% to 95.1%). Reproducibility between two determinations was 96.1% (95% CI = 86.5% to 99.5%). Combining the gene ratio test and other prognostic factors allowed prospective discrimination between patients at high risk (median survival = 6.9 months, 95% CI = 2.6 to 8.9 months; 3-year survival = 0%) and low risk (median survival = 31.9 months, 95% CI = 21.9 to 41.7 months; 3-year survival = 42%).

Conclusion The gene ratio test for survival of patients with malignant pleural mesothelioma has robust predictive value and technical assay performance.

J Natl Cancer Inst 2009;101:678–686

Malignant pleural mesothelioma is an aggressive, asbestos-induced malignancy diagnosed in approximately 3000 new patients each year in the United States. Chemotherapy and radiotherapy alone or in combination are usually ineffective (1), and the best randomized trial data show an improvement in median survival only from 9 to 12 months for patients with malignant pleural mesothelioma who are treated with chemotherapy (2). Patients with localized malignant pleural mesothelioma tumors and sufficient cardiopulmonary reserve may benefit from surgical resection, such as extrapleural pneumonectomy, followed by combination of adjuvant chemotherapy and radiation therapy. Lymph node status, stage, resection margins, and histological subtype have been associated with survival (3,4). However, many of these parameters are only available for analysis after complete surgical resection, which is too late to identify the important minority of patients who do not benefit from surgical resection and die from disease within 12 months of surgery. Therefore, availability of a preoperative predictive test

Affiliations of authors: Department of Surgery, Division of Thoracic Surgery, Harvard Medical School (GJG, LD, WGR, HE, MM, RC, GM, BVO, DJS, RB) and Department of Pathology (JNG), Brigham and Women's Hospital, Boston, MA; Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, MA (BYY).

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

See "Funding" and "Notes" following "References."

DOI: 10.1093/jnci/djp061

that identifies patients likely to benefit from surgery would likely improve patient care.

Cancer researchers have long sought to improve upon existing cancer diagnosis and prognosis methods by analyzing gene expression with tools such as high-density microarrays that simultaneously measure the expression levels of thousands of genes in a given specimen. This technology has been used to classify and dissect the pathobiology of human cancers (5–8), and many global cancer gene expression signatures and patterns have been identified that are associated with cancer diagnosis, stage, subtype, and prognosis (9,10). Nevertheless, only a few candidate predictive gene panels have been validated in independent cohorts, and among those that have been validated, only two are in clinical use, both for predicting recurrence in patients with resected breast cancer (11–13). A number of technical barriers serve to impede the translation of such prospective clinical tests from discovery to clinical implementation (14), including between-platform variability, complex bioinformatics-based algorithms that are not easily reproducible, lack of sufficient adequate-quality frozen specimens for test validation, and the necessity of lengthy patient follow-up to validate the test for cancer prognosis. We designed this study to test the validity of a gene expression ratio–based predictive test that addresses many of these problems in a model of human cancer.

We developed the four-gene expression ratio test to translate comprehensive expression profiling data into simple clinical tests that are based on the expression levels of a relatively small number of genes (15–20). This algorithm identifies genes that are differentially expressed between two clinically distinct conditions and calculates ratios of gene expression for gene pairs that can predict the condition alone or in combination. In contrast to traditional approaches, both genes in such a given gene pair ratio are informative, and by its use of a unitless ratio, this technique has the major advantages of being independent of microarray platform and easy to use with individual specimens. Previously, we compared (16,19) the differences in gene expression in tissue samples of malignant pleural mesothelioma stratified by patient outcome after surgical therapy (ie, survival) and developed a gene ratio test that was calculated from the expression levels of four genes to predict patient outcome. We demonstrated (16,19) that this test could differentiate between patients on the basis of postsurgical outcome in two independent retrospective cohorts. In this study, we used results of a prospective clinical trial to evaluate the ability of this test to predict overall survival and cancer-specific survival in patients undergoing surgery (ie, extrapleural pneumonectomy) for malignant pleural mesothelioma. We also investigated the predictive value and technical assay performance of this test in the same prospectively consented cohort of patients with malignant pleural mesothelioma.

CONTEXT AND CAVEATS

Prior knowledge

Malignant pleural mesothelioma has few effective treatments, except for Chemotherapy and cytoreductive surgery.

Study design

Clinical data were obtained before surgery from patients with malignant pleural mesothelioma. Tumor specimens were obtained at surgery or by pleural biopsy examination, and gene expression data were obtained for four genes. A gene ratio test was used to assign patients to good or poor prognosis groups, and their survival was analyzed. Robustness was determined by using many specimens per patient, two biopsy techniques, and two performance sites.

Contribution

The gene ratio test for survival of patients with malignant pleural mesothelioma has robust predictive value and technical assay performance.

Implications

The gene ratio test should be further evaluated with patient specimens collected before clinical intervention to determine whether its results can be incorporated into decision making for patient treatment.

Limitations

Patients had to agree to undergo aggressive surgery at enrollment and patients who did not were excluded. Consequently, the results may not apply to patients who are older or less fit.

From the Editors

Patients and Methods

Patients and Specimens

A total of 120 consecutive patients who were scheduled for debulking surgery for malignant pleural mesothelioma at one institution (Brigham and Women's Hospital) provided written informed consent for this institution review board–approved study to test their surgical specimens for gene expression and link that data to their clinical outcomes (2001–2006). None of the patients had received preoperative chemotherapy or radiation therapy. All patients were considered surgical candidates and were preoperatively staged by standard clinical and radiological criteria (1). These patients constituted an independent prospectively enrolled cohort whose tumors had never before been subjected to this type of molecular analysis and who were distinct from patients in our previous studies (16,19), whose data were used to develop the gene ratio test for malignant pleural mesothelioma prognosis.

Clinical data, including diagnosis, age, sex, survival, and other demographic and outcome data, were collected by independent clinical research staff who were blinded to the test results. The clinical and gene expression datasets, including gene ratio test predictions, were merged at the final stage of data analysis by the statistician who was not involved in any of the clinical or gene expression data collection.

Tumor Tissue and Reverse Transcriptase–Polymerase Chain Reaction Analysis

Specimens were removed by extrapleural pneumonectomy. After the entire specimen was surgically removed, it was immediately examined by a pathologist who grossly divided it into sub-specimens that were stored fresh frozen as described in the tumor bank protocol (21).

Portions of each tumor sub-specimen were used to make slides for staining with hematoxylin–eosin, which were reviewed by a

pathologist who counted tumor cell nuclei, with results being expressed as a percentage of all nuclei. A portion from each sub-specimen (each allotted a unique identifier) was homogenized in Trizol (Invitrogen, Carlsbad, CA) to prepare total RNA by use of the manufacturer's recommended protocol. Total RNA (2 μg) from each sub-specimen was used to prepare cDNA and quantify expression by quantitative reverse transcriptase–polymerase chain reaction (RT-PCR) of mRNA for TM4SF1 (L6 tumor antigen), ARHGDIA (GDIA1), COBLL1 (KIAA0977), and PKM2 (CTHBP1), exactly as described previously (16) (including PCR primer concentrations and cycling parameters), with a single modification being that the following PCR primers were used, which offered increased specificity (R. Bueno, Division of Thoracic Surgery, Brigham and Women's Hospital, unpublished observations) over previously reported (16) primers (forward and reverse): TM4SF1 (5′-GCACATTGTGGAATGGAATG-3′ and 5′-TCTGTCCTGGGTTGGTTCTT-3′), ARHGDIA (5′-CAACGTCGTGGTGACTGG-3′ and 5′-TCGGTTAACCCG GAAAGAG-3′), COBLL1 (5′-GATGCGACAGAGTTTGCT GA-3′ and 5′-GGTGTGGCAGGGTAACATTT-3′), PKM2 (5′-CTCGGGCTGAAGGCAGT-3′ and 5′-AATTGCAAGTGGT AGATGGCA-3′).

Gene Ratio Test Predictions

The relative expression levels of genes analyzed by quantitative RT-PCR were used to calculate the geometric mean of the three gene pair ratios, TM4SF1/PKM2, TM4SF1/ARHGDIA, and COBLL1/ARHGDIA, to assign a good (ie, geometric mean >1) or poor (ie, geometric mean <1) prognosis, exactly as described earlier (16). The cutoff threshold of "1" and all test-related calculations were prespecified (16) before collecting gene expression measurements in the current set of specimens, and the test was used without any modification whatsoever. In a related analysis of the test properties, we empirically determined the range of geometric mean values in the immediate vicinity of the threshold that produced uninterpretable results due to technical artifacts resulting from variability associated with repeated RT-PCR gene expression measurements obtained on different runs.

Analysis of Test Properties

We determined the test properties (ie, minimal tumor cell requirements and within-assay, between-assay, and within-patient reproducibility) by use of surgical specimens from 51 consecutive patients (a subset of the 120 patients). Briefly, each surgical specimen was divided in up to five separate sub-specimens in which the pleural surface appeared to be grossly involved. A scalpel was then used to excise three additional small portions from each sub-specimen to produce a set of up to 15 samples. These samples are referred to as the tissue samples hereafter. To determine if the gene ratio test is potentially amenable to the analysis of material acquired through minimally invasive procedures, surgical biopsy forceps were also used to obtain three additional samples from each sub-specimen for a total of another 15 samples. These samples are referred to as the pleural biopsy samples hereafter. In total, up to 30 samples from each patient's tumor were analyzed. All 30 specimens from each patient were immediately frozen in OCT compound, and the relative location of each specimen was recorded. Each specimen was labeled with a randomly generated code that was not linked directly to any patient identifier, so that the laboratory personnel were blinded to the clinical data. The clinical research staff was blinded to the predictive test results, and the statistician, who was blinded to the collection of clinical and gene expression data, merged the datasets at the final stage of data analysis.

Portions of each of the 30 samples per patient were reviewed by a pathologist to determine the percentage of all nuclei in the sample that were tumor nuclei in at least five high-power fields per slide by examining at least 100 nuclei per slide. Two adjacent portions from each specimen were homogenized in Trizol (Invitrogen) to prepare total RNA for duplicate gene ratio testing, exactly as described above.

To control for laboratory or performance site–dependent variability, each of the 30 sub-specimens (15 tissue samples and 15 pleural biopsy samples) from all 51 patients were subjected to an identical protocol and duplicate gene expression analysis with quantitative RT-PCR as described above, at a second location. This facility (Xceed Molecular, Toronto, ON, Canada) was a Clinical Laboratory Improvement Amendments–certified commercial facility, and the assays were performed in a contracted fee-for-service manner in this laboratory. Both laboratories, Xceed Molecular and Brigham and Women's Hospital, used exactly the same model of PCR machines. For all study samples, the gene ratio test was repeated twice.

Statistical Analysis

Patients were followed until death, and their status was obtained by contacting their referring physicians and/or families. Survival time was measured from the date of debulking surgery until October 1, 2007. Patients who were still alive at the time of the last follow-up appointment were censored as of that date. Two patients who were lost to long-term follow-up were also censored at the date of last contact. Overall survival was estimated by the Kaplan–Meier method, and the log-rank test was used to compare the difference between patient subgroups. Cancer-specific survival was based on the cumulative incidence estimate, with death due to causes other than malignant pleural mesothelioma treated as a competing risk. Patients were assumed to have died from malignant pleural mesothelioma if the primary cause of death could not be determined. A point-wise confidence interval (CI) that was based on the asymptotic variance was computed by use of log transformation (22), and the Gray test was used to compare the cancer-specific survival difference between patient subgroups (23). The Cox proportional hazards model was used to assess the predictive utility of the gene ratio test on overall survival while controlling for the simultaneous effects of other prognostic factors. The proportional hazards assumption of covariates in the Cox model was assessed by testing for a nonzero slope of the scaled Schoenfeld residuals (24). Repeatability of test results between two independent determinations was defined as the proportion of specimens assigned to the same risk group in both runs. When the gene ratio test was performed on multiple specimens obtained from an individual patient, the risk prediction was determined by a majority rule. Concordance was defined as the proportion of patients with consistent test results among all the specimens within each individual and reported as an average of observed rates in the two

independent runs. Reproducibility within patients was the proportion of patients with the same risk prediction from two independent determinations of the gene ratio test. The 95% confidence interval for rates of test performance was based on the exact binomial distribution. The statistical analysis was computed with SAS version 9.1 (SAS Institute, Cary, NC) and R version 2.5.1 (http://www.r-project.org), including the cmprsk package (http://bio-www.dfci.harvard.edu/~gray) used for the competing risks analysis. All statistical tests were two-sided.

Results

One hundred twenty patients were prospectively enrolled in the study between February 1, 2001, and December 31, 2006. Most patients were male, and slightly more than half had epithelial tumors (Table 1). The distribution of sex and histology in this group mirrors that of patients with malignant pleural mesothelioma in our practice and in practices at other centers (4,25–28). All patients underwent extrapleural pneumonectomy, with most also being treated with intraoperative instillation of heated cisplatin into the pleural and peritoneal cavities after tumor resection (29,30). The median follow-up at the time of analysis was 15 months among the 38 patients still alive, and the minimum follow-up was 4 months after surgery. In the analysis, 32 patients were alive without evidence of disease, four patients were alive with recurrent disease, 65 patients were dead of disease, 16 patients were dead of other causes, one patient died without a known cause, and two patients who were originally from outside the United States were lost to follow-up.

Survival Analysis

The median overall survival was 12.9 months (95% CI = 11.1 to 16.8 months) among the 120 patients (Figure 1, A). Overall survival was not associated with the use of intraoperative heated cisplatin (Table 2), which supports pooling data from the whole cohort for analysis, although the study was not designed to assess the effect of chemotherapy. The gene ratio test, which was performed and applied exactly as previously described (16), was used to assign the patients to two groups, with 70 (58%) of the 120 patients being assigned to the good outcome group and 50 (42%) being assigned to the poor outcome group. Overall survival between the two groups was statistically significantly ($P < .001$) different, with the median for the good outcome group being 16.8 months (95% CI = 12.4 to 25.8 months) and that for the poor outcome group being 9.5 months (95% CI = 7.2 to 13.6 months) (Figure 1, B). The gene ratio test also observed a statistically significant difference in cancer-specific survival ($P = .007$), with the median cancer-specific survival for the good outcome group being 21.9 months (95% CI = 16.7 to 40.7 months) and that for the poor outcome group being 15.9 months (95% CI = 8.6 to 21.0 months). The lymph node status and histological subtype, two well-established prognostic factors for malignant pleural mesothelioma, were also strongly related to outcome in the univariate analysis (Table 2).

We used a multivariable model to analyze patient survival data that was adjusted for the effects of histological subtype, tumor stage, and lymph node status as covariates to assess further the robustness of the gene ratio test (Table 3). The prognostic contributions of both lymph node status (hazard ratio [HR] = 1.97, 95% CI = 1.15 to 3.38, $P = .013$) and histological subtype (HR = 1.88, 95% CI = 1.14 to 3.10, $P = .013$) also remained statistically significant in the multivariable model, indicating that the gene ratio test (HR for death = 2.09, 95% CI = 1.27 to 3.45, $P = .004$) appears to provide additional predictive information to current pathological staging methods. Furthermore, the hazard ratio for the gene ratio test was in the same range (HR = 2.0–4.6) as that found in our initial studies (16,19), and those studies used several different platforms for gene expression analysis. These results support the robustness of the gene ratio test.

Table 1. Patient and treatment characteristics

Characteristics	Value
Total patients evaluable for analysis, No.	120
No. of patients alive at last follow-up	38
Median follow-up from surgery (range), mo	15 (4–58)
Median age (range), y	60 (27–77)
Sex, No. (%)	
Male	92 (77)
Female	28 (23)
Histological subtype, No. (%)	
Epithelial	68 (57)
Mixed	46 (38)
Sarcomatoid	6 (5)
Intraoperative therapy, No. (%)	
Heated cisplatin	92 (77)
No chemotherapy	28 (23)
Tumor stage, No. (%)	
T1	4 (3)
T2	36 (30)
T3	52 (43)
T4	28 (23)
Lymph node stage, No. (%)	
N0	53 (44)
N1	18 (15)
N2	49 (41)

Test Properties

The technical assay performance of the gene ratio test was examined in a set of 51 consecutive patients. Tumors from all patients were informative and contributed to the analysis. A total of 253 tissue specimens (five from each of 49 patients and four from each of two patients) were analyzed by quantitative RT-PCR to determine the properties of the gene ratio test, including run-to-run repeatability in tissue specimens, tumor cell requirements, and reproducibility within patients. The repeatability of test results between two independent determinations was 88.5% (95% CI = 84.0% to 92.2%). We further examined whether repeatability could be improved by excluding gene ratio combined scores (ie, geometric mean of the three gene pair ratios) with numerical values close to 1 as technically unacceptable (ie, within a margin of error that rendered test results uninterpretable). We empirically determined that excluding values of log(gene ratio combination) from −0.1 to +0.1 improved repeatability to 91.9% (95% CI = 87.4% to 95.1%), although it decreased the sample size to 221 specimens (excluding 32 or 12.6% of the specimens) because of the exclusion of uninterpretable results.

Tumor content in 252 specimens ranged from 0% to 90%, with a median of 43%; it was unknown in one specimen. Of these 252

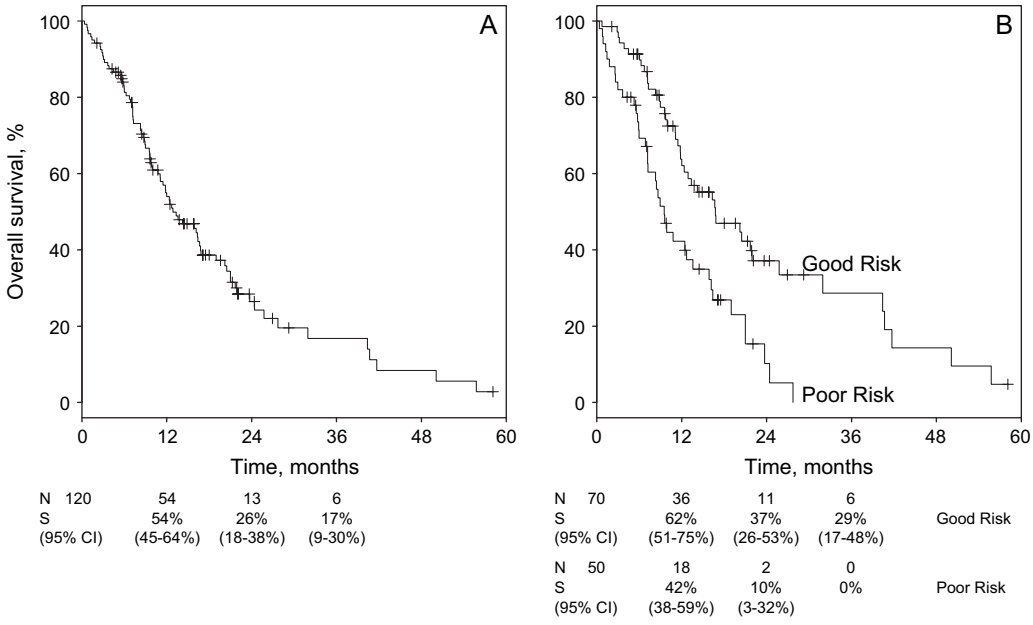

Figure 1. Overall survival of patients with malignant pleural mesothelioma and predictions from the gene ratio test. **A**) Entire cohort. **B**) Subgroups of patients as classified by the gene ratio test. **Hash marks** = censored patients; N = number of patients at risk; *S* = Kaplan–Meier survival point estimate as percent; 95% CI = 95% confidence interval for the survival estimate.

specimens, tumor content in 22 (9%) was less than 10% tumor; in 20 (8%), it was 10%; and in 144 (57%), it was at least 40%. Repeatability of the gene ratio test did not appear to be affected by or related to low levels of tumor content in tissue specimens. Only five of 252 specimens had no microscopically detectable tumor cells in either of the two slides stained with hematoxylin–eosin. These five specimens from three patients represented a minority of all the specimens available for analysis.

Within-patient variability was determined by computing the concordance of results from the gene ratio test among the four or five specimens obtained from each of the 51 patients. Among 79% of the 51 patients on average between the two independent determinations, the gene ratio test gave consistent results for at least four of the within-patient specimens tested. In particular, 51% of the 51 patients exhibited full concordance of test results among all of their five or four within-patient specimens tested. The gene ratio predictions for individual patients from two independent determinations were reproducible in 96.1% (95% CI = 86.5% to 99.5%) of the patients. When 32 specimens with values in the range of log(gene ratio combination) from −0.1 to +0.1 were omitted from the analysis, the reproducibility was 100% among the 44 patients who could definitively be assigned to a prognosis group in both determinations. Exclusion of the 32 specimens from this analysis resulted in an equivocal gene ratio–based prognosis for seven (14%) patients.

Table 2. Overall survival: univariate analysis*

Group assignment by prognostic factors and gene ratio test	Overall survival	
	Median (95% CI), mo	*P* value†
Tumor stage		.118
T1–T2	21.0 (12.7 to 27.7)	
T3–T4	11.8 (9.5 to 16.4)	
Lymph node status		.009
N0	21.7 (12.7 to 25.8)	
N1–N2	11.1 (8.6 to 14.2)	
Histological subtype		<.001
Epithelial	16.8 (13.6 to 24.4)	
Mixed or sarcomatoid	9.5 (7.2 to 12.4)	
Predictive test		<.001
Good risk	16.8 (12.4 to 25.8)	
Poor risk	9.5 (7.2 to 13.6)	

* The median of overall survival was 12.9 months (95% CI = 11.1 to 16.8 months) among the 120 patients. CI = confidence interval.

† Two-sided log-rank test.

Pleural Biopsy Specimens

We determined properties for the gene ratio test (ie, technical assay performance) by use of the 242 pleural biopsy specimens (five specimens from each of 40 patients, four from each of nine patients, and three from each of two patients) that were analyzed in two independent quantitative RT-PCRs. The between-run repeatability of pleural biopsy results was 93.4% (95% CI = 89.5% to 96.2%). Among 84% of the 51 patients, the gene ratio test gave consistent results for all specimens or for all but one of the within-patient specimens tested, including 59% of the 51 patients who exhibited full concordance of test results among all within-patient biopsy specimens tested. For pleural biopsy specimens, within-patient reproducibility of prognosis determined by the gene ratio test between two independent determinations was 94.1% (95% CI = 83.8% to 98.8%). Both the repeatability between two independent determinations and within-patient reproducibility for pleural biopsy specimens were as great or greater than those obtained with tumor specimen portions discussed in the previous section. We concluded that for predictive analysis, it is adequate to sample five minimally invasive pleural biopsy specimens from visible tumors with a tumor cell content of at least 10%.

A Model Combining Molecular and Pathological Predictive Parameters

We next investigated associations of the gene ratio test and pathological predictive parameters (ie, histology and lymph node status) with patient survival (Table 3). For this analysis, we assigned a

Table 3. Overall survival: multivariable model*

Model covariate	Adverse effect	Comparison group	*P* value†	HR (95% CI)
Predictive test	Poor risk	Good risk	.004	2.09 (1.27 to 3.45)
Histological subtype	Mixed or sarcomatoid	Epithelial	.013	1.88 (1.14 to 3.10)
Lymph node status	N1–N2	N0	.013	1.97 (1.15 to 3.38)
Tumor stage	T3–T4	T1–T2	.443	1.24 (0.71 to 2.17)

* HR = hazard ratio for death; CI = confidence interval.

† Two-sided Wald test.

value of 0 or 1 to each of the three predictive parameters: histology, lymph node status, and the predictive gene ratio test result. Specifically, we assigned a value of 1 to mixed or sarcomatoid histology, presence of cancer in a lymph node, and a poor gene ratio test result, and a value of 0 to epithelial histology, absence of cancer in a lymph node, and a good gene ratio test result. This approach segregated the patients into four distinct survival groups (Table 4). Median overall survival was 31.9 months (95% CI = 21.9 to 41.7 months) in the low-risk group, 13.6 months (95% CI = 11.5 to 20.2 months) in the low intermediate-risk group, 11.1 months (95% CI = 8.3 to 16.2 months) in the high intermediate-risk group, and 6.9 months (95% CI = 2.6 to 8.9 months) in the high-risk group. Moreover, overall survival of the low intermediate-risk and high intermediate-risk groups was sufficiently similar so that they were combined into a single intermediate-risk group (Figure 2). The three-subgroup model was associated with median overall survival of 31.9 months (95% CI = 21.9 to 41.7 months) in the low-risk groups, 12.9 months (95% CI = 9.9 to 16.4 months) in the intermediate-risk groups, and 6.9 months (95% CI = 2.6 to 8.9 months) in the high-risk groups; the rates of overall survival at 3 years were 42%, 12%, and 0%, respectively.

Discussion

We have validated the use of the gene ratio test to predict postsurgical outcome for patients with malignant pleural mesothelioma by use of prospectively collected clinical data. We previously developed (16,19) the gene ratio test and showed that a binary test that was based on three ratios of four genes of mRNA levels in tumors (ie, TM4SF1/PKM2, TM4SF1/ARHGDIA, and COBLL1/ARHGDIA) could predict postsurgical outcome for patients with malignant pleural mesothelioma. We found that the test was repeatable and reproducible when the same specimen set was used in independent determinations and when determinations with the same specimen set were performed by different personnel, with different instruments, at separate laboratories, and at different times. These findings indicate that tumor specimens selected by a clinician on the basis of their appearance should be sufficient to ensure consistent test results for each patient tested. We identified a working range for improving the test properties in terms of repeatability between two independent determinations and reproducibility within patients (for poor outcome, combined score of less than 0.905 or less than −0.1 on a log scale, and for good outcome, combined score of greater than 1.105 or greater than +0.1 on a log scale) and also found that highly reproducible results could generally be obtained by using five samples per patient. We further demonstrated that pleural biopsy specimens from a visible tumor were generally adequate for analysis even when the tumor content was relatively low. We concluded that for predictive analysis, it is adequate to sample five minimally invasive pleural biopsy specimens from visible tumors with a tumor cell content of at least 10%.

The gene ratio test for malignant pleural mesothelioma has several advantages over standard gene expression–based predictive algorithms with widespread usage (7). Since the seminal paper in the field by Golub et al. (31), there have been many reports of microarray data used to classify samples; to identify predictive genes in cancer; and to develop cancer diagnosis,

Table 4. Combined predictions algorithm: patients with malignant pleural mesothelioma after surgery (ie, extrapleural pneumonectomy)*

Model and risk group	Risk factor†	Patients, No. (%)	Median OS, mo (95% CI)	OS (95% CI), %		
				1 y	2 y	3 y
Four subgroups						
Low	0	21 (18)	31.9 (21.9 to 41.7)	84 (69 to 100)	68 (47 to 97)	42 (20 to 91)
Low intermediate	1	44 (37)	13.6 (11.5 to 20.2)	57 (43 to 75)	23 (11 to 46)	18 (8 to 42)
High intermediate	2	40 (33)	11.1 (8.3 to 16.2)	49 (35 to 68)	11 (2 to 53)	0
High	3	15 (12)	6.9 (2.6 to 8.9)	18 (6 to 57)	0	0
Three subgroups‡						
Low	0	21 (18)	31.9 (21.9 to 41.7)	84 (69 to 100)	68 (47 to 97)	42 (20 to 91)
Intermediate	1–2	84 (70)	12.9 (9.9 to 16.4)	53 (43 to 66)	20 (11 to 35)	12 (6 to 30)
High	3	15 (12)	6.9 (2.6 to 8.9)	18 (6 to 57)	0	0

* OS = overall survival; CI = confidence interval.

† Total number of adverse risk factors among mixed or sarcomatoid histology, presence of cancer in a lymph node, and a poor predictive gene ratio test result.

‡ In this model, the intermediate-risk groups from the four-subgroup model were combined.

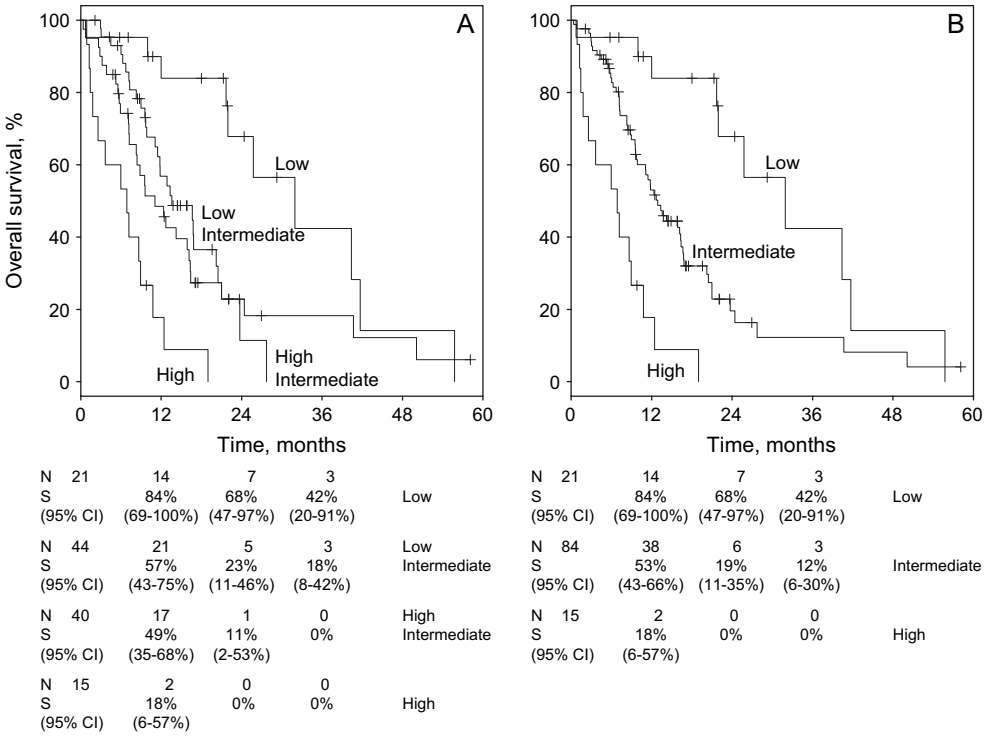

Figure 2. Overall survival of patients with malignant pleural mesothelioma after surgery from combined prediction models (as defined in Table 4). **A**) Model for four subgroups of patients. **B**) Model for three subgroups of patients. **Hash marks** = censored patients; N = number of patients at risk; S = Kaplan – Meier survival point estimate as percent; 95% CI = 95% confidence interval for the survival estimate.

prognosis, and predictive tests, most of which used microarrays to develop a predictive algorithm in a training set of specimens and validated the same platform and algorithm by use of a comparable independent validation set of specimens (7). In previous research (32), most specimen sets tested have been assembled retrospectively from tumor banks, specimen quality has rarely been assessed, and the algorithms used have typically been quite complex and required analysis of the expression of many genes. Consequently, data from these approaches have not been easy to reproduce. In addition, industry standards have not been developed that would allow the straightforward translation of these tests into clinical use.

There have been very few instances of later validation with prospectively obtained specimens (11,13) and no instance that uses alternative platforms. Only two such tests are presently available for clinical use in the United States: MammaPrint (Agendia BV) and Oncotype DX (Genomic Health). MammaPrint uses expression patterns of 70 genes to identify patients after surgery for breast cancer who are at high risk of recurrence (11,13) and was the first mRNA microarray-based assay to be approved by the Food and Drug Administration. Oncotype DX uses expression levels of 21 genes to predict the likelihood of breast cancer recurrence after surgery for breast cancer (12). Neither test addressed sample characteristics comprehensively including the minimal required tumor cell content, the effects of sample handling on test quality, nor other quality control measures. Both of these tests are presently being used to determine which patients might benefit from additional chemotherapy by identifying high risk of recurrence, but neither test was designed or validated to predict outcome after any available therapy (33).

In this study, a previously identified gene ratio test appears to predict preoperatively who will benefit from a specific cancer therapy (ie, surgery). This test can thus likely be applied informatively in a more widespread manner to the subset of patients with malignant pleural mesothelioma who are considering surgery. In current clinical practice, cancer-specific risk factors used to predict outcomes after surgery for malignant pleural mesothelioma are histological subtype, lymph node status, resection margin status, and stage. The limitation of these risk factors is that none can be determined definitively until after major surgery (4,25). A positive mediastinoscopy is certainly an indicator of poor prognosis, but this procedure has a relatively low sensitivity, and a negative finding does not rule out lymph node metastasis associated with poor outcome (26). In addition, half of patients with malignant pleural mesothelioma whose final pathology at the time of definitive surgery indicate a mixed histology tumor were found to have only epithelial histology upon initial diagnosis before surgery because of inadequate sampling highlighting the limitations of preoperative subtyping (25). Therefore, a minimally invasive test that can be performed before the major surgical intervention and that can accurately predict postsurgical outcome is likely to be clinically useful. Tissue specimens for the gene ratio test can be obtained, for example, at the time of pleuroscopy and mediastinoscopy that are performed to confirm the diagnosis and surgical staging, although we did not examine such specimens in this study. A sufficient number of biopsy specimens are usually taken during routine patient workup to provide adequate tissue for the gene ratio analysis. Patients assigned to the predicted poor outcome group, particularly when other established prognostic factors such as histology and lymph node status are also suggestive of poor outcome, could be counseled to forgo surgery, which would not benefit them, and to seek best supportive care. They could alternatively be encouraged to participate in more rationally targeted clinical trials of nonsurgical modalities.

Our study is not without limitations. For example, it may be to some degree biased by the fact that enrolled patients made an a priori decision to undergo aggressive surgical therapy and that the specimens were obtained at extrapleural pneumonectomy rather than at a preoperative staging procedure. Thus, it may not be applicable to all patients with mesothelioma, particularly to older, less fit patients. However, given that the size of pleural biopsy samples that we obtained is identical to that obtained at pleuroscopy and pleuroscopy allows direct observation and selection of tumors for biopsy, we expect that similar results will be observed.

Unfortunately, most therapies for malignant pleural mesothelioma do not work and the median survival is short (1,2). A doubling of the median survival to approximately 16 months, as we found in the good prognosis group, is highly statistically significant and clinically relevant. Patients whose gene ratio test results predict a good prognosis after surgery may more confidently select the treatment option that includes surgery. We suggest that ultimately, this strategy will lead to improved survival after surgery (because those unlikely to benefit from surgery will be excluded from surgical treatment), will minimize undue expense and surgical trauma to patients who are unlikely to benefit from surgery, and will allow improved comparisons among patients within individual predictive groups between clinical centers and within clinical trials. In this study, we examined a cohort of patients whose treatment was not modified by the results of the predictive test. Another important step will be to apply the gene ratio test to patient specimens collected before clinical intervention to validate the incorporation of results from this test into decision making for patient treatment. Our findings also point to the possible applicability of other gene expression ratio tests to other cancers because of its many technical strengths, simplicity, and robustness of this method.

References

1. Chang MY, Sugarbaker DJ. Extrapleural pneumonectomy for diffuse malignant pleural mesothelioma: techniques and complications. *Thorac Surg Clin.* 2004;14(4):523–530.
2. Vogelzang NJ, Rusthoven JJ, Symanowski J, et al. Phase III study of pemetrexed in combination with cisplatin versus cisplatin alone in patients with malignant pleural mesothelioma. *J Clin Oncol.* 2003;21(14):2636–2644.
3. Sugarbaker D, Strauss GM, Lynch TJ, et al. Node status has prognostic significance in the multimodality therapy of diffuse, malignant mesothelioma. *J Clin Oncol.* 1993;11(6):1172–1178.
4. Sugarbaker DJ, Flores RM, Jaklitsch MT, 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(1):54–65.
5. Quackenbush J. Weighing our measures of gene expression [published online ahead of print November 14, 2006]. *Mol Syst Biol.* 2006;2:63. doi:10.1038/msb4100096.
6. Quackenbush J. Computational approaches to analysis of DNA microarray data. *Methods Inf Med.* 2006;45(suppl 1):91–103.
7. Quackenbush J. Microarray analysis and tumor classification. *N Engl J Med.* 2006;354(23):2463–2472.
8. Quackenbush J. Standardizing the standards. *Mol Syst Biol.* 2006;2: 2006 0010.
9. Chen HY, Yu SL, Chen CH, et al. A five-gene signature and clinical outcome in non-small-cell lung cancer. *N Engl J Med.* 2007;356(1):11–20.
10. Chung CH, Bernard PS, Perou CM. Molecular portraits and the family tree of cancer. *Nat Genet.* 2002;32(suppl):533–540.
11. van 't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature.* 2002;415(6871):530–536.
12. Paik S, Shak S, Tang G, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med.* 2004; 351(27):2817–2826.
13. van de Vijver MJ, He YD, van't Veer LJ, et al. A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med.* 2002;347(25): 1999–2009.
14. Sotiriou C, Piccart MJ. Taking gene-expression profiling to the clinic: when will molecular signatures become relevant to patient care? *Nat Rev Cancer.* 2007;7(7):545–553.
15. Bueno R, Loughlin KR, Powell MH, Gordon GJ. A diagnostic test for prostate cancer from gene expression profiling data. *J Urol.* 2004;171 (2 Pt 1):903–906.
16. Gordon GJ, Jensen RV, Hsiao LL, et al. Using gene expression ratios to predict outcome among patients with mesothelioma. *J Natl Cancer Inst.* 2003;95(8):598–605.
17. Gordon GJ, Jensen RV, Hsiao L-L, 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(17):4963–4967.
18. Gordon GJ, Richards WG, Sugarbaker DJ, Jaklitsch MT, Bueno R. A prognostic test for adenocarcinoma of the lung from gene expression profiling data. *Cancer Epidemiol Biomarkers Prev.* 2003;12:905–910.
19. Gordon GJ, Rockwell GN, Godfrey PA, et al. Validation of genomics-based prognostic tests in malignant pleural mesothelioma. *Clin Cancer Res.* 2005;11(12):4406–4414.
20. Gordon GJ, Rockwell GN, Jensen RV, et al. Identification of novel candidate oncogenes and tumor suppressors in malignant pleural mesothelioma using large-scale transcriptional profiling. *Am J Pathol.* 2005;166(6): 1827–1840.
21. Richards W, Van Oss S, Glickman J, et al. A microaliquoting technique for precise histological annotation and optimization of cell content in frozen tissue specimens. *Biotech Histochem.* 2007;82(4&5):189–197.
22. Aalen O. Nonparametric estimation of partial transition probabilities in multiple decrement models. *Ann Stat.* 1978;6(3):534–545.
23. Gray R. A class of K-sample tests for comparing the cumulative incidence of a competing risk. *Ann Stat.* 1988;16(3):1141–1154.
24. Grambsch P, Therneau T. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika.* 1994;81(3):512–526.
25. Bueno R, Reblando J, Glickman J, Jaklitsch MT, Lukanich JM, Sugarbaker DJ. Pleural biopsy: a reliable method for determining the diagnosis but not subtype in mesothelioma. *Ann Thorac Surg.* 2004;78(5):1774–1776.
26. Stevens CW, Forster KM, Smythe WR, Rice D. Radiotherapy for mesothelioma. *Hematol Oncol Clin North Am.* 2005;19(6):1099–1115.
27. Flores RM, Pass HI, Seshan VE, et al. Extrapleural pneumonectomy versus pleurectomy/decortication in the surgical management of malignant pleural mesothelioma: results in 663 patients [published online ahead of print February 14, 2008]. *J Thorac Cardiovasc Surg.* 2008;135:620–626, 626.e1–3.
28. Flores RM, Zakowski M, Venkatraman E, et al. Prognostic factors in the treatment of malignant pleural mesothelioma at a large tertiary referral center. *J Thorac Oncol.* 2007;2(10):957–965.
29. Sugarbaker DJ. Macroscopic complete resection: the goal of primary surgery in multimodality therapy for pleural mesothelioma. *J Thorac Oncol.* 2006;1(2):175–176.
30. Richards WG, Zellos L, Bueno R, et al. Phase I to II study of pleurectomy/decortication and intraoperative intracavitary hyperthermic cisplatin lavage for mesothelioma. *J Clin Oncol.* 2006;24(10):1561–1567.
31. Golub TR, Slonim DK, Tamayo P, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science.* 1999;286(5439):531–537.
32. Quackenbush J. Computational analysis of microarray data. *Nat Rev Genet.* 2001;2(6):418–427.
33. Marchionni L, Wilson RF, Wolff AC, et al. Systematic review: gene expression profiling assays in early-stage breast cancer [published online ahead of print February 4, 2008]. *Ann Intern Med.* 2008;148(5):358–369.

Funding

National Cancer Institute (100315 and 120528 to R.B.) as well as grants from the International Mesothelioma Program at Brigham and Women's Hospital (to R.B.) and the Maurice Favell Fund at the Vancouver Foundation (to R.B.).

Notes

G. J. Gordon, L. Dong, and B. Y. Yeap contributed equally to this work.

The study sponsors played no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the paper for publication.

We thank Carl Alsup, Laura Hoffmeister, Julia Fleming, Jennifer Stuart, Kathy Kee, Harriet Mercer, Dan Wilson, and David Englert for technical help, and Dr Jeffrey B. Tatro for editorial review. We thank Ann S. Adams for her editorial work.

Manuscript received June 10, 2008; revised January 27, 2009; accepted February 19, 2009.