VOLUME 27 · NUMBER 34 · DECEMBER 1 2009

JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

Melanoma Prognostic Model Using Tissue Microarrays and Genetic Algorithms

Bonnie E. Gould Rothberg, Aaron J. Berger, Annette M. Molinaro, Antonio Subtil, Michael O. Krauthammer, Robert L. Camp, William R. Bradley, Stephan Ariyan, Harriet M. Kluger, and David L. Rimm

From the Departments of Pathology, Medicine, Dermatology, and Surgery, Yale University School of Medicine; Department of Biostatistics, Yale School of Public Health, New Haven, CT.

Submitted February 26, 2009; accepted July 13, 2009; published online ahead of print at www.jco.org on November 2, 2009.

Supported by Grant No. RO-1 CA114277 (D.L.R.) from the National Institutes of Health.

Terms in blue are defined in the glossary, found at the end of this article and online at www.jco.org.

Authors' disclosures of potential conflicts of interest and author contributions are found at the end of this article.

Corresponding author: David L. Rimm, MD, PhD, Department of Pathology, Yale University School of Medicine, 310 Cedar St, PO Box 208023, New Haven, CT 06520-8023; e-mail: david.rimm@yale.edu.

The Acknowledgment and Appendix are included in the full-text version of this article; they are available online at www.jco.org. They are not included in the PDF version (via Adobe® Reader®).

0732-183X/09/2734-5772/$20.00

DOI: 10.1200/JCO.2009.22.8239

ABSTRACT

Purpose
As a result of the questionable risk-to-benefit ratio of adjuvant therapies, stage II melanoma is currently managed by observation because available clinicopathologic parameters cannot identify the 20% to 60% of such patients likely to develop metastatic disease. Here, we propose a multimarker molecular prognostic assay that can help triage patients at increased risk of recurrence.

Methods
Protein expression for 38 candidates relevant to melanoma oncogenesis was evaluated using the automated quantitative analysis (AQUA) method for immunofluorescence-based immunohistochemistry in formalin-fixed, paraffin-embedded specimens from a cohort of 192 primary melanomas collected during 1959 to 1994. The prognostic assay was built using a genetic algorithm and validated on an independent cohort of 246 serial primary melanomas collected from 1997 to 2004.

Results
Multiple iterations of the genetic algorithm yielded a consistent five-marker solution. A favorable prognosis was predicted by ATF2 ln(non-nuclear/nuclear AQUA score ratio) of more than −0.052, $p21^{WAF1}$ nuclear compartment AQUA score of more than 12.98, $p16^{INK4A}$ ln(non-nuclear/nuclear AQUA score ratio) of ≤ −0.083, β-catenin total AQUA score of more than 38.68, and fibronectin total AQUA score of ≤ 57.93. Primary tumors that met at least four of these five conditions were considered a low-risk group, and those that met three or fewer conditions formed a high-risk group (log-rank $P < .0001$). Multivariable proportional hazards analysis adjusting for clinicopathologic parameters shows that the high-risk group has significantly reduced survival on both the discovery (hazard ratio = 2.84; 95% CI, 1.46 to 5.49; $P = .002$) and validation (hazard ratio = 2.72; 95% CI, 1.12 to 6.58; $P = .027$) cohorts.

Conclusion
This multimarker prognostic assay, an independent determinant of melanoma survival, might be beneficial in improving the selection of stage II patients for adjuvant therapy.

J Clin Oncol 27:5772-5780. © 2009 by American Society of Clinical Oncology

INTRODUCTION

Adjuvant therapy is the standard of care for many low-stage cancers that can be completely resected with tumor-free margins. However, for some other cancers, the lack of effective and safe adjuvant therapy leads to an excess of mortality directly related to the development of metastatic disease in patients assumed to have undergone a complete resection of their malignancy. One important example is cutaneous malignant melanoma, the sixth most common cancer in the United States.[1] Although more than 80% of new cases are still localized to the skin[1] where a wide local excision should be curative in the setting of a negative sentinel lymph node biopsy, the unfavorable risk-to-benefit ratio of available adjuvant regimens advocates caution when administering such agents to individuals with stage I to IIA and even stage IIB or IIC disease, where high-dose interferon alfa-2b is currently approved by the US Food and Drug Administration in the adjuvant setting.[2] Consequently, 20% of these patients will develop metastases and die of their disease within 10 years, with more than 30% 10-year mortality among patients with T3 and T4 tumors.[3] Development of a prognostic tool that could selectively triage the subset of stage II patients at high risk of recurrence for adjuvant therapy could potentially lower the burden of untreatable metastatic cancer

and enable us to selectively treat those patients who are more likely to develop distant metastatic disease.

Nine clinicopathologic prognostic markers have been identified and incorporated into clinically validated outcome risk stratification models.[3,4] However, these do not account for all of the observed variability associated with melanoma-related survival. Immunohistochemistry (IHC) is a widely accepted and well-documented method for characterizing patterns of protein expression in formalin-fixed, paraffin-embedded samples while preserving tissue and cellular architecture.[5] Although no IHC marker has become standard of care, new work may suggest the inclusion of Ki-67.[6] Our recent systematic review of melanoma IHC data shows that individual contributions of IHC markers to overall prognosis are of narrow statistical significance and thus unlikely to demonstrate broad clinical utility[7] or see wide adoption.

Here, we describe the generation of an independently significant, multimarker prognostic model for melanoma using genetic algorithms on a subset of 38 candidate proteins assessed on a cohort of 192 primary melanomas. Our model shows two prognostic groups (low risk and high risk), created from five markers, that were successfully validated as significant independent prognostic factors in a second cohort of 246 primary melanomas. These data demonstrate the potential for multimarker assays in improving melanoma prognostic assessment and warrant a prospective, randomized, controlled melanoma prognostic study. This test could be a valuable tool to help determine which sentinel node–negative stage II melanoma patients should seek adjuvant therapy or other aggressive management strategies.

METHODS

Patients and Tumor Samples

Seven hundred thirty-seven tumor samples from three nonoverlapping series of patients with cutaneous melanoma were analyzed for protein expression. The Yale Melanoma Discovery Cohort consisted of 192 white patients who underwent resection of a primary invasive cutaneous melanoma at Yale-New Haven Hospital during 1959 to 1994 for whom the surgical specimen was not exhausted during diagnosis and for which follow-up information is available. The Yale Melanoma Validation Cohort included 246 patients with serial Clark level III to V cutaneous melanoma who underwent sentinel lymph node biopsy by a single surgeon during 1997 to 2004.[8] The Yale Metastatic Series includes 299 unique subcutaneous metastases, lymph node metastases, or visceral metastases occurring in patients previously diagnosed with cutaneous melanoma and surgically removed at Yale-New Haven Hospital during 1959 to 1994 (n = 198) or during 1995 to 2002 (n = 101). For the primary melanomas, clinical data describing patient demographics, date of diagnosis, clinical course, and follow-up through August 1, 2007 were obtained after a comprehensive review of the medical record, the archives of the Connecticut Tumor Registry, and, if applicable, the State of Connecticut Vital Records. This study was approved by the Yale Human Investigations Committee.

Tissue Microarray Construction, IHC, and Automated Image Acquisition and Analysis

Formalin-fixed, paraffin-embedded blocks were retrieved from the Yale Pathology Archives, and 0.6-mm tissue microarrays (TMAs) were constructed according to the published method.[9] The discovery TMA included single cores from the 192 primary melanomas, the 299 metastases, and a series of controls. The validation TMA included two-fold redundant cores in separate blocks from the 246 patients plus a random selection of 60 individuals from the discovery series to facilitate normalization of the validation array. Fluorescence-based immunohistochemical staining was performed by using the automated quantitative analysis (AQUA) technology as previously described.[10,11] Using this method, target antigen expression is automatically determined, blinded to any a priori clinical information, as the sum of intensities from the Cy5 channel in all pixels within a compartment defined by S100 staining divided by the number of pixels within that compartment (see Appendix, online only, for details).

Statistical Analysis

Cores whose tumor mask covered less than 5% of the total histospot area were dropped from further analysis. For individuals represented by multiple cores on the TMA, AQUA scores were averaged before analysis. To normalize the AQUA scores between the discovery and validation cohorts, a regression equation was calculated for the set of 60 samples spotted on both arrays, and the mean values for the validation cohort were adjusted according to the regression equation.

To develop a multimarker prognostic model from the discovery cohort data, a genetic algorithm using standard methodology[12,13] within the X-tile software suite[14] was used (see Appendix). Bivariate and survival analyses were performed using SAS version 9.1.3 and Statview 5.0 (SAS Institute, Cary, NC), and adjustments for multiple comparisons were performed using the standard Bonferroni method.

RESULTS

Patient Characteristics

The distribution of demographic and clinicopathologic characteristics for both the discovery and validation cohorts is presented in Table 1. In addition to the longer follow-up time ($P < .0001$), the discovery cohort displayed overall thicker tumors ($P = .01$), a more balanced sex distribution ($P = .04$), a higher prevalence of ulcerated melanomas ($P = .01$), and fewer superficial spreading melanomas ($P = .04$) than the validation cohort.

Clinicopathologic Correlates of Candidate Marker Expression

Thirty-eight unique protein markers were assayed by AQUA on the discovery cohort (n = 192) and, for comparison, the metastatic series (n = 299). Exclusion of individual tumors as a result of random failure for individual histospots and attrition of samples as a result of exhaustion of the arrayed tumor core resulted in less than 100% of tumor samples available for analysis from each assay. Only the subset of 20 markers with missingness completely at random was included in subsequent analyses.

Associations between levels of protein expression and tumor progression were evaluated by the Mann-Whitney *U* test (Appendix Table A2, online only). After adjustment for multiple comparisons, levels of fibronectin, Ki-67, and $p21^{WAF1}$ and the ratios for both ATF2 and $p16^{INK4A}$ were significantly elevated, whereas Hey1, HDM2, N-cadherin, nuclear $p16^{INK4A}$, and non-nuclear ATF2 were significantly decreased among the metastases compared with the primary tumors ($P \leq .0025$).

To determine the independent crude and adjusted effects of each marker on melanoma-specific mortality, the AQUA scores or calculated ratios were divided into quartiles, and the hazard ratios and the associated *P* values were calculated using Cox proportional hazards modeling (Appendix Table A4, online only). Using these cut points, five markers, one that increased risk with increasing value ($p16^{INK4A}$ ratio, $P = .04$) and four that decreased risk with increasing value (ATF2, $P = .001$; β-catenin, $P = .04$; N-cadherin, $P = .001$; $p16^{INK4A}$, $P = .047$), were significant at $P < .05$ on univariate analysis, but only

Table 1. Demographics and Clinical Characteristics of the Yale Melanoma Discovery and Validation Cohorts

Characteristic	Discovery Cohort (n = 192)		Validation Cohort (n = 246)		*P*
	No.	%	No.	%	
Follow-up time for censored individuals, years					< .0001*
Mean	9.50		4.05		
Standard deviation	9.14		2.12		
Breslow thickness, mm					.01*
Mean	2.42		1.95		
Standard deviation	2.01		1.78		
Age at diagnosis, years					.34
Mean	57.77		59.28		
Standard deviation	15.65		16.76		
Sex					.04*
Male	96	50.0	147	59.8	
Female	96	50.0	99	40.2	
Stage at diagnosis					NA
Localized	160	84.2	246	100	
Regional spread	16	8.4	0	0	
Distant metastases	14	7.4	0	0	
Ulceration					.01*
Absent	135	70.3	198	80.5	
Present	57	29.7	48	19.5	
Tumor-infiltrating lymphocytes					.09
Nonbrisk	150	78.5	208	84.9	
Brisk	41	21.5	37	15.1	
Histologic subtype					.04*
Superficial spreading	127	66.1	132	73.7	
Nodular	30	15.6	24	13.4	
Lentigo maligna	8	4.2	4	2.2	
Acral lentiginous	11	5.7	1	0.6	
Other	16	8.3	18	10.1	
Chronically sun-exposed anatomic site†					.14
No	95	49.7	105	42.7	
Yes	96	50.3	141	57.3	
Received any nonsurgical therapy					.37
No	153	79.7	201	83.1	
Yes	39	20.3	41	16.9	
Microsatellitosis					NA
Absent	149	77.6	0	0	
Present	43	22.4	0	0	
Positive sentinel lymph node biopsy					NA
No	0	0	211	87.6	
Yes	0	0	30	12.4	

NOTE. Numbers may not sum to total because of missing values; percentages may not sum to 100% as a result of rounding.
Abbreviation: NA, not applicable.
*Significant at $P < .05$.
†Anatomic location was dichotomized as chronically sun exposed (face, scalp, neck, arms, legs, and nonacral lentiginous lesions of hands and feet) and non–chronically exposed (chest, back, abdomen, groin, and hand and foot acral lentiginous lesions).

two markers, ATF2 and N-cadherin, remained significant after adjustment for multiple comparisons ($P \leq .0025$).

Multivariable Cox proportional hazards modeling included adjustment for age at diagnosis, sex, Breslow thickness (millimeters), stage at diagnosis, presence of microsatellitosis, sun exposure to anatomic site, and receipt of systemic therapy. Two of the five markers significant on univariate analysis (β-catenin, $P = .04$; $p16^{INK4A}$, $P = .04$) retained both their significance at $P < .05$ and directionality of effect after adjustment for clinicopathologic parameters. Three additional markers that were not significant on crude analysis became significant at $P < .05$ on multivariable analysis (α-catenin, p27/KIP1, and tenascin-C).

Constructing a Genetic Algorithm–Based Multimarker Prognostic Model

Because the power of multiplexed biomarker assays is thought to be greater than that obtainable with any single marker, we sought to identify a robust prognostic indicator by combining information from all 20 available markers, regardless of whether a significant independent association with progression or prognosis was obtained, using

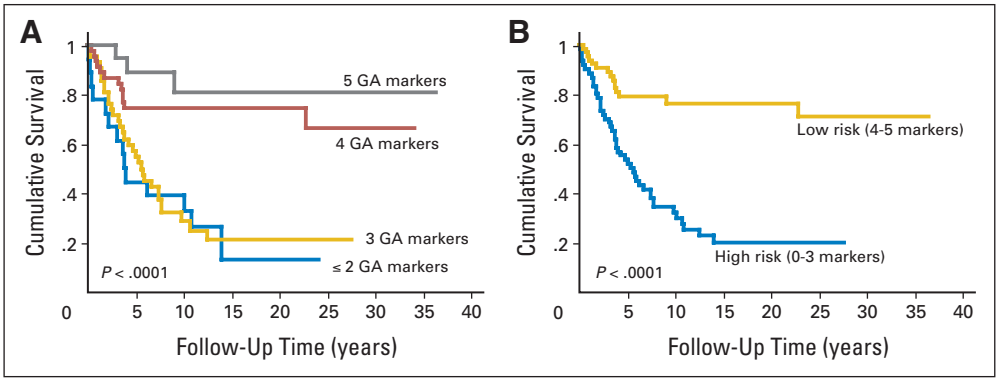

Fig 1. Kaplan-Meier estimates of melanoma-specific mortality among the 129 Yale Melanoma Discovery Cohort participants with complete data across the five markers comprising the genetic algorithm (GA) –based multimarker prognostic assay according to algorithm-derived prognostic score. (A) Survival curves drawn according to number of prognostic conditions met. (B) Survival curves for the dichotomized model describing low-risk (four to five conditions met) or high-risk (≤ three conditions met) groups.

genetic algorithms. Our selected model, obtained in each of the five independent iterations, yielded a log-rank χ^2 of 24.27 ($P = 1.5 \times 10^{-6}$) and consisted of the following five markers and associated cut points: ATF2 ratio more than −0.052, β-catenin more than 38.68, fibronectin ≤ 57.93, $p16^{INK4A}$ ratio ≤ −0.083, and $p21^{WAF1}$ more than 12.98.

The Kaplan-Meier curves for the four classes obtained from the genetic algorithm are presented in Figure 1A. On the basis of the similar survival experiences of the groups with ≤ two or three conditions and those with four or five conditions, we further simplified our model to the following two groups: a low-risk group with four or five marker conditions being met and a high-risk group with less than four marker conditions being met (Fig 1B). Crude and multivariable survival estimates were calculated for the multimarker predictor and the clinicopathologic covariates using Cox proportional hazards modeling (Table 2). In our final multivariable model, the high-risk group demonstrated a nearly three-fold increased risk of mortality ($P = .002$) compared with the low-risk group. Other variables remaining significant in the multivariable model included stage at diagnosis and receipt of nonsurgical therapy ($P \leq .01$), with Breslow thickness trending toward significance ($P = .06$).

Assessment of Multimarker Model Reproducibility in the Validation Cohort

To determine the prognostic breadth and strength of our genetic algorithm-based multimarker predictor, we performed the assay on the independent validation TMA, normalizing the two builds as described. Complete AQUA data were obtained for 226 of the 246 eligible individuals, with 76 individuals (33.6%) meeting criteria for the low-risk group and 150 (66.4%) belonging to the high-risk group.

Table 2. Crude and Multivariable-Adjusted Melanoma-Specific Mortality Hazard Ratios for the Genetic Algorithm–Based Multimarker Predictor in the Yale Melanoma Discovery Cohort

	Univariate Analysis			Multivariate Analysis		
Parameter	Hazard Ratio	95% CI	P*	Hazard Ratio	95% CI	P
Genetic algorithm-based predictor			< .0001†			.002†
Low-risk group (4 or 5 conditions met)	1.00			1.00		
High-risk group (< 4 conditions met)	3.88	2.16 to 6.94		2.84	1.46 to 5.49	
Breslow thickness, mm	1.28	1.14 to 1.43	< .0001†	1.14	0.99 to 1.31	.06
Age at diagnosis, years	1.01	0.99 to 1.03	.41	1.01	0.99 to 1.03	.39
Sex			.14			.14
Male	1.00			1.00		
Female	0.68	0.41 to 1.14		0.66	0.38 to 1.14	
Stage at diagnosis						
Localized	1.00			1.00		
Regional spread	3.54	1.72 to 7.30	.0006†	4.67	2.08 to 10.47	.0002†
Distant metastases	5.05	2.35 to 10.94	< .0001†	3.32	1.31 to 8.39	.01†
Chronically sun-exposed anatomic site			.03†			.24
No	1.00			1.00		
Yes	0.56	0.33 to 0.95		0.70	0.39 to 1.26	
Microsatellitosis			.047†			.63
Absent	1.00			1.00		
Present	1.73	1.01 to 2.96		1.16	0.64 to 2.11	
Receipt of nonsurgical therapy			.0005†			.008†
No	1.00			1.00		
Yes	2.54	1.50 to 4.30		2.31	1.25 to 4.26	

*P values were calculated according to the Wald method.
†Significant at $P \leq .05$.

Notably, our predictor was independent of both Breslow thickness ($P = .41$) and sentinel lymph node status ($P = .52$; Table 3). Our predictor trended toward, but did not achieve, significance for melanoma-specific mortality in univariate analysis (Table 4). However, multivariable modeling that adjusted for Breslow thickness, age at diagnosis, anatomic site, sentinel lymph node biopsy status, and receipt of nonsurgical therapy revealed a significantly increased melanoma-specific mortality for the high-risk group (adjusted hazard ratio = 2.72; 95% CI, 1.12 to 6.58; $P = .027$; Table 4), consistent with the possibility of negative confounding by clinicopathologic parameters in the validation set. Our predictor is independent of sentinel lymph node status, and the interaction between multimarker assignment and sentinel lymph node status was not significant ($P = .78$).

Although the multivariate analysis of the validation set is statistically significant (Table 4), the Kaplan-Meier analysis of the validation set is not (Fig 2A) most likely because of the confounding effect of nonuniform treatment. McShane et al,[15] in the Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK) guidelines, point out the value of the multivariate analysis over the log-rank assessment performed on the Kaplan-Meier data. This work is an example of the multivariate analysis adjusting for confounding to show significance, as anticipated by the REMARK criteria. However, the Kaplan-Meier plot is shown to help convey the data in a more simple form related to the envisioned utility of the test in sentinel node–negative patients. In this population, the high-risk group has only a 60% 10-year survival rate compared with a 10-year survival rate of more than 90% in the low-risk group (Fig 2B, log-rank $P = .09$).

Table 3. Bivariate Associations Between the Genetic Algorithm–Derived Prognostic Indicator and Clinicopathologic Correlates of Melanoma Outcome for the Yale Melanoma Validation Cohort

Parameter	Low-Risk Group (n = 76) No.	%	High-Risk Group (n = 150) No.	%	P
Breslow thickness, mm					.41
Mean	1.86		2.08		
Standard deviation	1.73		1.89		
Age at diagnosis, years					.08
Mean	57.17		61.14		
Standard deviation	15.66		16.57		
Sex					.12
Male	41	54.0	97	64.7	
Female	35	46.1	53	35.3	
Ulceration					.39
Absent	63	82.9	117	78.0	
Present	13	17.1	33	22.0	
Tumor-infiltrating lymphocytes					.94
Nonbrisk	64	84.2	126	84.6	
Brisk	12	15.8	23	15.4	
Histologic subtype					.43
Superficial spreading	45	76.3	76	72.4	
Nodular	7	11.9	16	15.2	
Lentigo maligna	2	3.4	1	1.0	
Acral lentiginous	1	1.7	0	0.0	
Other	4	6.8	12	11.4	
Chronically sun-exposed anatomic site					.99
No	33	43.4	65	43.3	
Yes	43	56.6	85	56.7	
Received any nonsurgical therapy					.50
No	60	80.0	123	83.7	
Yes	15	20.0	24	16.3	
Sentinel lymph node biopsy					.52
Negative	64	85.3	129	88.4	
Positive	11	14.7	17	11.6	

NOTE. Numbers may not sum to total because of missing values; percentages may not sum to 100% as a result of rounding.

DISCUSSION

Over the last few years, multimarker molecular models have been constructed to supplement available clinicopathologic parameters for refining prognosis in some tumor types. Here, we report on a multimarker melanoma prognostic assay with potential for translation into the clinic that may be especially useful for identifying the subset of stage II melanoma patients most appropriate for supplemental therapy. Presently, up to 40% of patients with stage IIA or IIB melanoma will die of their disease within 10 years of diagnosis.[3,16] Because of the poor risk-to-benefit ratio and toxicity of current adjuvant therapy regimens,[2] these are not often administered in this population. We believe there is a significant clinical need to stratify this population at the time of diagnosis into a subset of stage II patients with the highest risk for recurrence and a lower risk group. The goal of this stratification, using the test described here, would be to offer adjuvant intervention or at least aggressive follow-up screening to high-risk stage II patients. We believe this would improve the overall survival of these vulnerable patients without exposing the remaining patients to the risk of excessive toxicity; thus, this test has the potential to alter the standard of care for management of melanoma. However, such an approach would require prospective validation in the target (sentinel node–negative) population.

To our knowledge, only one other prognostic multimarker molecular classifier for primary melanoma has been described specifying a 254-gene classifier obtained from differential mRNA expression profiling on a series of 83 snap-frozen samples.[17] Although protein expression by IHC was confirmed for the 23-gene subset with commercially available antibodies, the authors only reported on their marginal univariate and multivariate prognostic relationships. Although this study is valuable, to date, the multimarker classifier has not been validated on a second population. Additional molecular classifiers of melanoma phenotype that integrate either somatic mutation[18] or gene expression information[19] have been reported but have not been evaluated for prognostic relevance. Efforts that use hierarchical clustering, which is valuable for classification, suffer from the inabilities to both calculate error associated with a clustering run and prospectively assign new patients to existing clusters without re-executing the clustering, which risks reorganizing cluster assignment. Assignment of new patients according to our genetic algorithm profile, as demonstrated in our validation strategy, only requires simultaneous AQUA analysis of selected reference standards.

Assignment to the low-risk group requires elevated levels of overall β-catenin and nuclear $p21^{WAF1}$, decreased levels of fibronectin, and distributions that favor nuclear concentration for $p16^{INK4A}$ but cytoplasmic concentration for ATF2. Each of these assignments is consistent with the previous literature for melanoma. Our data, as well as

Table 4. Crude and Multivariable-Adjusted Melanoma-Specific Mortality Hazard Ratios for the Genetic Algorithm-Based Multimarker Predictor in the Yale Melanoma Validation Cohort

Parameter	Univariate Analysis			Multivariate Analysis		
	Hazard Ratio	95% CI	*P**	Hazard Ratio	95% CI	*P*
Genetic algorithm-based predictor			.14			.027†
Low-risk group (4 or 5 conditions met)	1.00			1.00		
High-risk group (< 4 conditions met)	1.75	0.83 to 3.72		2.72	1.12 to 6.58	
Breslow thickness, mm	1.20	1.11 to 1.31	< .0001	1.14	1.01 to 1.29	.029†
Age at diagnosis, years	1.03	1.00 to 1.05	.027	1.04	1.01 to 1.07	.007†
Sex			.07			.10
Male	1.00			1.00		
Female	0.51	0.25 to 1.06		0.52	0.24 to 1.14	
Chronically sun-exposed anatomic site			.20			.11
No	1.00			1.00		
Yes	1.55	0.79 to 3.04		1.96	0.87 to 4.44	
Sentinel lymph node biopsy status			< .0001			.017†
Negative	1.00			1.00		
Positive	4.41	2.23 to 8.71		2.78	1.20 to 6.47	
Receipt of nonsurgical therapy			< .0001			.0001†
No	1.00			1.00		
Yes	7.09	3.60 to 13.96		4.65	2.11 to 10.24	

*P values were calculated according to the Wald method.
†Significant at $P \leq .05$.

data from others,[20,21] support that increased nuclear p16^{INK4A} expression significantly improves melanoma prognosis in multivariable modeling, consistent with its role in cell cycle inhibition.[22] Although specific cytoplasmic p16^{INK4A} expression has been confirmed by multiple high-resolution imaging technologies,[23,24] little is known about its functional role or prognostic implications. Our data suggest that a ratio that favors nuclear abundance contributes to improved cell cycle control. A similar rationale can be suggested for elevated nuclear p21^{WAF1}; however, neither we nor others have shown a significant effect for the marginal effects of nuclear p21^{WAF1} on univariate[25,26] or multivariate[20,27] analysis. The requirement for a higher proportion of cytoplasmic ATF2 is supported by the observation that although ATF2 possesses both nuclear export and nuclear localization signals and shuttles between both locations, nuclear heterodimerization with c-Jun and subsequent phosphorylation of both subunits by MAP kinases are required for transcriptional activation activity.[28,29] Although we did not distinguish between membranous cadherin-associated and cytoplasmic/nuclear Wnt signaling–associated β-catenin, our association between improved prognosis and elevated β-catenin is consistent with others.[30,31] Finally, our requirement for reduced fibronectin supports both tissue- and cell-based observations that increased tumor-derived expression facilitates melanoma cell invasion and metastasis.[32-34]

This work suffers from a number of limitations. Perhaps the most significant limitation is the relatively limited set of available markers eligible for our analysis. Unlike nucleic acid arrays where tens of thousands of genes can be interrogated in each experiment, we can only assess one gene product at a time (although we have the advantage of assessing hundreds of patients per experiment). Furthermore, more than half of the markers initially considered for this study were ultimately eliminated from our genetic algorithm as a result of preferential attrition of longer surviving (typically thinner) melanomas as a result of exhaustion of their tissue cores with higher cuts of the TMA. Future replication of these results on parallel blocks of the discovery TMA may both fill gaps and also provide useful information regarding heterogeneity of marker expression. Although we selected a broad range of candidate targets, the inherent limitation of the candidategene approach omitted sufficient markers

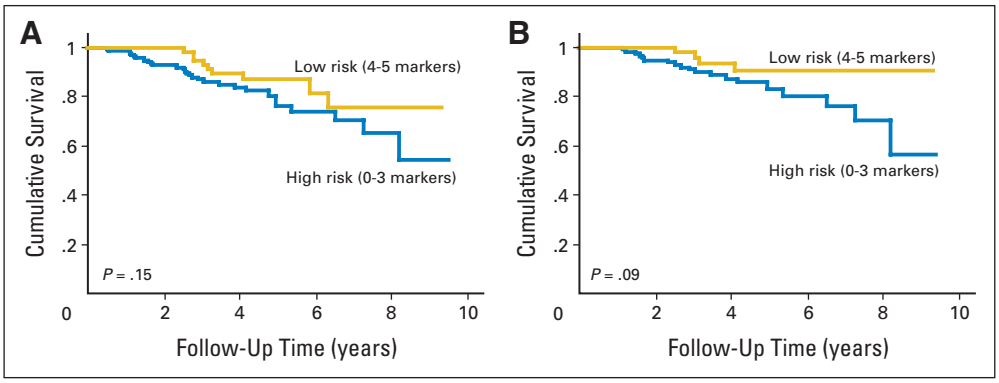

Fig 2. Kaplan-Meier estimates of melanoma-specific mortality for the dichotomized model describing favorable or unfavorable profiles among (A) all 226 participants of the Yale Melanoma Validation Cohort scored completely for the multimarker prognostic assay, and (B) the 193 members of the Yale Melanoma Validation Cohort who are sentinel lymph node negative (stage II melanoma).

from some cancer progression pathways such as evading apoptosis, sustained angiogenesis, orinsensitivity to antigrowth signals.[35] Additionally, several proteins previously shown by others to have significant independent marginal associations with melanoma outcome, such as MMP-2,[36,37] osteopontin,[38] MCAM/MUC18,[39,40] and AIB-1,[41] were not assayed (in some cases because of antibody validation failure). Another theoretical weakness of this approach is that our genetic algorithm equally weighted each protein's individual contribution. This is in contrast to a commercially available breast cancer diagnostic (Oncotype DX; Genomic Health, Redwood City, CA) where individual marker contribution is weighted according to its relative marginal contribution to the overall model.[42] The genetic algorithm approach risks bias in group assignment should the presence or absence of one specific marker disproportionately drive assignment into one of the algorithm states. However, as shown earlier, we found that this bias did not occur in our discovery phase.

Strengths of our approach include the use of equally large and robust, yet completely independent, training and validation study populations as well as choice of a computational method that supports the prospective evaluation of new patients according to its calculated criteria. Given that we were able to replicate a significant, independent association between our multimarker prognostic assay and melanoma-specific mortality after adjustment for relevant clinicopathologic covariates in our independently collected validation set, we believe these data could support the use of this test to assist management of patients with sentinel node–negative melanoma. For example, a high-risk test result in a sentinel node–negative patient might prompt that patient to choose adjuvant therapy. Although the data on the efficacy of adjuvant interferon are controversial,[43] other adjuvant therapies such as ipilimumab and vaccine therapies are currently under investigation, and these studies typically include only stage III patients. However, high-risk stage II patients identified by improved prognostic assays such as this should also be considered for these studies. Prospective validation is planned in a broader geographic constituency to determine whether this method should become part of the routine work-up for patients with malignant melanoma.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Although all authors completed the disclosure declaration, the following author(s) indicated a financial or other interest that is relevant to the subject matter under consideration in this article. Certain relationships marked with a "U" are those for which no compensation was received; those relationships marked with a "C" were compensated. For a detailed description of the disclosure categories, or for more information about ASCO's conflict of interest policy, please refer to the Author Disclosure Declaration and the Disclosures of Potential Conflicts of Interest section in Information for Contributors.

Employment or Leadership Position: None **Consultant or Advisory Role:** Annette M. Molinaro, HistoRx (C); Robert L. Camp, HistoRx (C); David L. Rimm, HistoRx (C) **Stock Ownership:** Robert L. Camp, HistoRx; David L. Rimm, HistoRx **Honoraria:** None **Research Funding:** None **Expert Testimony:** None **Other Remuneration:** None

AUTHOR CONTRIBUTIONS

Conception and design: Bonnie E. Gould Rothberg, Aaron J. Berger, Robert L. Camp, Harriet M. Kluger, David L. Rimm
Financial support: Harriet M. Kluger, David L. Rimm
Provision of study materials or patients: Aaron J. Berger, Antonio Subtil, Stephan Ariyan, David L. Rimm
Collection and assembly of data: Bonnie E. Gould Rothberg, Aaron J. Berger, Antonio Subtil, Robert L. Camp, William R. Bradley, Stephan Ariyan, Harriet M. Kluger, David L. Rimm
Data analysis and interpretation: Bonnie E. Gould Rothberg, Aaron J. Berger, Annette M. Molinaro, Michael O. Krauthammer, Robert L. Camp, William R. Bradley, Harriet M. Kluger, David L. Rimm
Manuscript writing: Bonnie E. Gould Rothberg, Aaron J. Berger, Harriet M. Kluger, David L. Rimm
Final approval of manuscript: Bonnie E. Gould Rothberg, Harriet M. Kluger, David L. Rimm

REFERENCES

1. Jemal A, Siegel R, Ward E, et al: Cancer statistics, 2008. CA Cancer J Clin 58:71-96, 2008

2. Tsao H, Atkins MB, Sober AJ: Management of cutaneous melanoma. N Engl J Med 351:998-1012, 2004

3. Balch CM, Soong SJ, Gershenwald JE, et al: Prognostic factors analysis of 17,600 melanoma patients: Validation of the American Joint Committee on Cancer melanoma staging system. J Clin Oncol 19:3622-3634, 2001

4. Gimotty PA, Elder DE, Fraker DL, et al: Identification of high-risk patients among those diagnosed with thin cutaneous melanomas. J Clin Oncol 25:1129-1134, 2007

5. Taylor C: Standardization in immunohistochemistry: The role of antigen retrieval in molecular morphology. Biotech Histochem 81:3-12, 2006

6. Gimotty PA, Van Belle P, Elder DE, et al: Biologic and prognostic significance of dermal Ki67 expression, mitoses, and tumorigenicity in thin invasive cutaneous melanoma. J Clin Oncol 23:8048-8056, 2005

7. Gould Rothberg BE, Bracken MB, Rimm DL: Tissue biomarkers for prognosis in cutaneous melanoma: A systematic review and meta-analysis. J Natl Cancer Inst 101:452-474, 2009

8. Ariyan S, Ariyan C, Farber LR, et al: Reliability of identification of 655 sentinel lymph nodes in 263 consecutive patients with malignant melanoma. J Am Coll Surg 198:924-932, 2004

9. Kononen J, Bubendorf L, Kallioniemi A, et al: Tissue microarrays for high-throughput molecular profiling of tumor specimens. Nat Med 4:844-847, 1998

10. Kreizenbeck GM, Berger AJ, Subtil A, et al: Prognostic significance of cadherin-based adhesion molecules in cutaneous malignant melanoma. Cancer Epidemiol Biomarkers Prev 17:949-958, 2008

11. Camp RL, Chung GG, Rimm DL: Automated subcellular localization and quantification of protein expression in tissue microarrays. Nat Med 8:1323-1327, 2002

12. Mitchell M: An Introduction to Genetic Algorithms. Cambridge, MA, MIT Press, 1998

13. Ooi CH, Tan P: Genetic algorithms applied to multi-class prediction for the analysis of gene expression data. Bioinformatics 19:37-44, 2003

14. Camp RL, Dolled-Filhart M, Rimm DL: X-tile: A new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. Clin Cancer Res 10:7252-7259, 2004

15. McShane LM, Altman DG, Sauerbrei W, et al: Reporting recommendations for tumor marker prognostic studies (REMARK). J Natl Cancer Inst 97:1180-1184, 2005

16. Gimotty PA, Botbyl J, Soong SJ, et al: A population-based validation of the American Joint Committee on Cancer melanoma staging system. J Clin Oncol 23:8065-8075, 2005

17. Winnepenninckx V, Lazar V, Michiels S, et al: Gene expression profiling of primary cutaneous melanoma and clinical outcome. J Natl Cancer Inst 98:472-482, 2006

18. Viros A, Fridlyand J, Bauer J, et al: Improving melanoma classification by integrating genetic and morphologic features. PLoS Med 5:e120, 2008

19. Bittner M, Meltzer P, Chen Y, et al: Molecular classification of cutaneous malignant melanoma by gene expression profiling. Nature 406:536-540, 2000

20. Alonso SR, Ortiz P, Pollan M, et al: Progression in cutaneous malignant melanoma is associated with distinct expression profiles: A tissue microarray-based study. Am J Pathol 164:193-203, 2004

21. Straume O, Sviland L, Akslen LA: Loss of nuclear p16 protein expression correlates with increased tumor cell proliferation (Ki-67) and poor

prognosis in patients with vertical growth phase melanoma. Clin Cancer Res 6:1845-1853, 2000

22. Sherr CJ, Roberts JM: CDK inhibitors: Positive and negative regulators of G1-phase progression. Genes Dev 13:1501-1512, 1999

23. Evangelou K, Bramis J, Peros I, et al: Electron microscopy evidence that cytoplasmic localization of the p16(INK4A) "nuclear" cyclin-dependent kinase inhibitor (CKI) in tumor cells is specific and not an artifact: A study in non-small cell lung carcinomas. Biotech Histochem 79:5-10, 2004

24. Keller-Melchior R, Schmidt R, Piepkorn M: Expression of the tumor suppressor gene product p16INK4 in benign and malignant melanocytic lesions. J Invest Dermatol 110:932-938, 1998

25. Sauroja I, Smeds J, Vlaykova T, et al: Analysis of G(1)/S checkpoint regulators in metastatic melanoma. Genes Chromosomes Cancer 28:404-414, 2000

26. Maelandsmo GM, Holm R, Fodstad O, et al: Cyclin kinase inhibitor p21WAF1/CIP1 in malignant melanoma: Reduced expression in metastatic lesions. Am J Pathol 149:1813-1822, 1996

27. Karjalainen JM, Eskelinen MJ, Kellokoski JK, et al: p21(WAF1/CIP1) expression in stage I cutaneous malignant melanoma: Its relationship with p53, cell proliferation and survival. Br J Cancer 79:895-902, 1999

28. Bhoumik A, Ronai Z: ATF2: A transcription factor that elicits oncogenic or tumor suppressor activities. Cell Cycle 7:2341-2345, 2008

29. Liu H, Deng X, Shyu YJ, et al: Mutual regulation of c-Jun and ATF2 by transcriptional activation and subcellular localization. Embo J 25:1058-1069, 2006

30. Bachmann IM, Straume O, Puntervoll HE, et al: Importance of P-cadherin, beta-catenin, and Wnt5a/frizzled for progression of melanocytic tumors and prognosis in cutaneous melanoma. Clin Cancer Res 11:8606-8614, 2005

31. Maelandsmo GM, Holm R, Nesland JM, et al: Reduced beta-catenin expression in the cytoplasm of advanced-stage superficial spreading malignant melanoma. Clin Cancer Res 9:3383-3388, 2003

32. Banerji A, Das S, Chatterjee A: Culture of human A375 melanoma cells in the presence of fibronectin causes expression of MMP-9 and activation of MMP-2 in culture supernatants. J Environ Pathol Toxicol Oncol 27:135-145, 2008

33. Gaggioli C, Robert G, Bertolotto C, et al: Tumor-derived fibronectin is involved in melanoma cell invasion and regulated by V600E B-Raf signaling pathway. J Invest Dermatol 127:400-410, 2007

34. Jaeger J, Koczan D, Thiesen HJ, et al: Gene expression signatures for tumor progression, tumor subtype, and tumor thickness in laser-microdissected melanoma tissues. Clin Cancer Res 13:806-815, 2007

35. Hanahan D, Weinberg RA: The hallmarks of cancer. Cell 100:57-70, 2000

36. Väisänen A, Kallioinen M, Taskinen PJ, et al: Prognostic value of MMP-2 immunoreactive protein (72 kD type IV collagenase) in primary skin melanoma. J Pathol 186:51-58, 1998

37. Väisänen AH, Kallioinen M, Turpeenniemi-Hujanen T: Comparison of the prognostic value of matrix metalloproteinases 2 and 9 in cutaneous melanoma. Hum Pathol 39:377-385, 2008

38. Rangel J, Nosrati M, Torabian S, et al: Osteopontin as a molecular prognostic marker for melanoma. Cancer 112:144-150, 2008

39. Pacifico MD, Grover R, Richman PI, et al: Development of a tissue array for primary melanoma with long-term follow-up: Discovering melanoma cell adhesion molecule as an important prognostic marker. Plast Reconstr Surg 115:367-375, 2005

40. Pearl RA, Pacifico MD, Richman PI, et al: Stratification of patients by melanoma cell adhesion molecule (MCAM) expression on the basis of risk: Implications for sentinel lymph node biopsy. J Plast Reconstr Aesthet Surg 61:265-271, 2008

41. Rangel J, Torabian S, Shaikh L, et al: Prognostic significance of nuclear receptor coactivator-3 overexpression in primary cutaneous melanoma. J Clin Oncol 24:4565-4569, 2006

42. 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 351: 2817-2826, 2004

43. Ascierto PA, Kirkwood JM: Adjuvant therapy of melanoma with interferon: Lessons of the past decade. J Transl Med 6:62, 2008

Glossary Terms

Confounding: Confounding variables are extraneous variables in a statistical model that are associated/correlated with both the independent and dependent variables but are not on the causal pathway between independent and dependent variables. When confounding variables are present, crude (unadjusted) statistical models describing the association between independent and dependent variables are biased (i.e., wrong) as the risk estimate includes the effect of the confounding variable as well (Type 1 error). As a result, to properly describe the relationship between independent and dependent variables, a multivariable model that includes both the independent variable and all relevant confounding variables as predictors must be executed.

Immunofluorescence: Refers to laboratory methods that combine the use of antibody reagents to detect the presence of specific biomolecule antigens in situ with a detection system that uses fluorescent molecules to visualize the localization of the target antigen/antibody complex.

Proportional hazards: Semiparametric approach to survival analysis developed by Cox in 1972. Unlike product-limit (Kaplan-Meier) survival analyses that are restricted to categorical predictor variables and do not produce a risk estimate, proportional hazards models can accommodate continuous and ordinal variables as well as allow for the inclusion of multiple predictor variables to compute adjusted risk estimates. Proportional hazards models are based on the fundamental premise that all individuals have the same baseline hazard that varies as a function of time [$\lambda(t)$] but that exposure to the independent variable changes the hazard by a fixed value [$h(x)$]. What is parameterized in the model is the value of this fixed effect per unit increase of the predictor variable whereas the value of $\lambda(t)$ remains uncharacterized.

Sentinel lymph node: The lymph node that is anatomically located such that it is the first site of lymph drainage from the location of the primary tumor. It is suspected and assumed that if a malignancy is going to disseminate via the lymphatic system, metastases will first be evident in the sentinel lymph node. In this manner, this lymph node is said to stand guard or sentinel over the metastatic state of the tumor. For many cancers, the sentinel lymph node is biopsied as part of the staging process and presence of macro- or micrometastases in the sentinel lymph node is a negative prognostic factor.

Genetic algorithm: Genetic algorithms are a type of iterative mathematical modeling technique used to find the optimal combinatorial state given a set of parameters of interest. Usage of the term "genetic" refers to the mechanism of the algorithm where, through the process of iteration, individual models "evolve" over time and compete with each other in a Darwinian fashion where the fittest model emerges as the solution, similar to how chromosomes evolve to create speciation. Given a set of parameters of interest, a baseline model is fit from a subset of the eligible variables, each dichotomized about a random cut point. Then, through successive iterations, the model is altered by either swapping one of the included parameters (a crossover) or by changing the dichotomization cut point for an included parameter (a mutation) and the model's fitness is reassessed. After several million iterations, the model with the best goodness of fit is selected.

Immunohistochemistry (IHC): The application of antigen-antibody interactions to histochemical techniques. Typically, a tissue section is mounted on a slide and is incubated with antibodies (polyclonal or monoclonal) specific to the antigen (primary reaction). The antigen-antibody signal is then amplified using a second antibody conjugated to a complex of peroxidase-antiperoxidase (PAP), avidin-biotin-peroxidase (ABC) or avidin-biotin alkaline phosphatase. In the presence of substrate and chromogen, the enzyme forms a colored deposit at the sites of antibody-antigen binding. Immunofluorescence is an alternate approach to visualize antigens. In this technique, the primary antigen-antibody signal is amplified using a second antibody conjugated to a fluorochrome. On UV light absorption, the fluorochrome emits its own light at a longer wavelength (fluorescence), thus allowing localization of antibody-antigen complexes.

Prognostic marker: A marker that predicts the prognosis of a patient (eg, the likelihood of relapse, progression, and/or death) independent of future treatment effects. A factor can be both prognostic and predictive.

Prognostic factor: A measurable patient characteristic that is associated with the subsequent course of disease (whether or not therapy is administered). The identification of a prognostic factor does not necessarily imply a cause-and-effect relationship. However, within a suitable outcome model, the measurement of a prognostic factor contributes to an estimate of an outcome probability (eg, the probability of disease-free survival within a given time interval).

Tissue microarray (TMA): Used to analyze the expression of genes of interest simultaneously in multiple tissue samples, tissue microarrays consist of hundreds of individual tissue samples placed on slides ranging from 2 to 3 mm in diameter. Using conventional histochemical and molecular detection techniques, tissue microarrays are powerful tools to evaluate the expression of genes of interest in tissue samples. In cancer research, tissue microarrays are used to analyze the frequency of a molecular alteration in different tumor type, to evaluate prognostic markers, and to test potential diagnostic markers.

