Gene Expression–Based Model Using Formalin-Fixed Paraffin-Embedded Biopsies Predicts Overall Survival in Advanced-Stage Classical Hodgkin Lymphoma

David W. Scott, Fong Chun Chan, Fangxin Hong, Sanja Rogic, King L. Tan, Barbara Meissner, Susana Ben-Neriah, Merrill Boyle, Robert Kridel, Adele Telenius, Bruce W. Woolcock, Pedro Farinha, Richard I. Fisher, Lisa M. Rimsza, Nancy L. Bartlett, Bruce D. Cheson, Lois E. Shepherd, Ranjana H. Advani, Joseph M. Connors, Brad S. Kahl, Leo I. Gordon, Sandra J. Horning, Christian Steidl, and Randy D. Gascoyne

See accompanying editorial on page 660 and article on page 684

Author affiliations appear at the end of this article.

Published online ahead of print at www.jco.org on November 26, 2012.

Supported in part by Public Health Service Grants No. CA21115, CA23318, CA66636, CA17145, CA11083, CA32102, CA38926, CA77202, CA21076, and CA77470 from the National Cancer Institute, National Institutes of Health, and the Department of Health and Human Services. Biospecimens were provided by the Eastern Cooperative Oncology Group Pathology Coordinating Office and Reference Laboratory. D.W.S. and K.L.T. were supported by postdoctoral fellowships of the Terry Fox Foundation Strategic Health Research Training Program in Cancer Research at Canadian Institutes of Health Research (Grant No. TGT-53912). C.S. was supported by postdoctoral fellowships of the Cancer Research Society (Steven E. Drabin Fellowship) and the Michael Smith Foundation for Health Research. R.D.G. receives funding support from the Canadian Institutes of Health Research (Grant No. 178536).

Presented in part at the 53rd Annual Meeting of the American Society of Hematology, San Diego, CA, December 10-13, 2011.

The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the National Cancer Institute.

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

Corresponding author: Randy D. Gascoyne, MD, Department of Pathology and Experimental Therapeutics, BC Cancer Agency and BC Cancer Research Centre, 675 W 10th Ave, Room 5-113, Vancouver, British Columbia, V6Z 1L3, Canada; e-mail: rgascoyn@bccancer.bc.ca.

© 2012 by American Society of Clinical Oncology

0732-183X/13/3106-692/\$20.00

DOI: 10.1200/JCO.2012.43.4589

ABSTRACT

Purpose

Our aim was to reliably identify patients with advanced-stage classical Hodgkin lymphoma (cHL) at increased risk of death by developing a robust predictor of overall survival (OS) using gene expression measured in routinely available formalin-fixed paraffin-embedded tissue (FFPET).

Methods

Expression levels of 259 genes, including those previously reported to be associated with outcome in cHL, were determined by digital expression profiling of pretreatment FFPET biopsies from 290 patients enrolled onto the E2496 Intergroup trial comparing doxorubicin, bleomycin, vinblastine, and dacarbazine (ABVD) and Stanford V regimens in locally extensive and advanced-stage cHL. A model for OS separating patients into low- and high-risk groups was produced using penalized Cox regression. The model was tested in an independent cohort of 78 patients enriched for treatment failure but otherwise similar to patients in a population-based registry of patients treated with ABVD. Weighted analysis methods generated unbiased estimates of predictor performance in the population-based registry.

Results

A 23-gene outcome predictor was generated. The model identified a population at increased risk of death in the validation cohort. There was a 29% absolute difference in 5-year OS between the high- and low-risk groups (63% v 92%, respectively; log-rank $P < .001$; hazard ratio, 6.7; 95% CI, 2.6 to 17.4). The predictor was superior to the International Prognostic Score and CD68 immunohistochemistry in multivariate analyses.

Conclusion

A gene expression–based predictor, developed in and applicable to routinely available FFPET biopsies, identifies patients with advanced-stage cHL at increased risk of death when treated with standard-intensity up-front regimens.

J Clin Oncol 31:692-700. © 2012 by American Society of Clinical Oncology

INTRODUCTION

Despite dramatic improvement in outcomes over the last half century, 10% to 15% of patients with advanced-stage classical Hodgkin lymphoma (cHL) continue to succumb to the disease.¹ Current up-front chemotherapy and radiotherapy regimens lead to different rates of relapse and progression, with more intensive regimens producing superior short-term outcomes but at the expense of greater treatment-related morbidity and mortality.^{2,3} Recent evidence suggests that planned high-dose chemotherapy and autologous transplantation for

patients whose lymphoma progresses or relapses reduces the previously apparent differences in overall survival (OS) between the up-front treatments.⁴

With improvement in outcomes from primary treatment and increased use of dose-intense salvage regimens, we lack reliable tools to identify a population of patients at significantly increased risk of death.⁵ A robust biomarker, applied at diagnosis, would ideally identify a population of patients whose low risk would allow the selection of an up-front regimen that minimizes adverse effects and long-term sequelae and a population at sufficiently high risk to justify consideration of dose-intense or

novel regimens. The tool provided by the International Prognostic Factors Project, the International Prognostic Score (IPS), was trained on freedom from progression of disease using data from patients largely treated in the 1980s.⁶ Recently, it has been demonstrated that the power of this tool to predict OS in the modern treatment era has weakened.^{5,6}

In cHL, the malignant cells typically make up less than 1% of the tumor.⁷ The remainder represents an extensive microenvironment composed of macrophages, T cells, B cells, plasma cells, mast cells, eosinophils, and fibroblasts, likely reflecting the interaction between surface proteins and secreted cytokines and chemokines produced by the malignant cells and the host immune system. We have recently demonstrated that certain characteristics of the microenvironment are associated with treatment outcomes, namely that increased number of CD68⁺ cells are associated with poor progression-free and disease-specific survival, and even as a single prognostic biomarker, CD68 immunohistochemistry outperforms the IPS.⁸

The development of gene expression–based predictors of OS in cHL has been hampered by lack of availability of suitable material from large cohorts of uniformly treated patients. This issue is largely overcome by the recent development of technologies to measure gene expression based on RNA from formalin-fixed paraffin-embedded

tissue (FFPET), a resource generated during the routine diagnostic work-up.⁹ Herein we describe the development of a gene expression–based predictor of OS in advanced-stage cHL using material from a large phase III randomized controlled trial,¹⁰ validated in an independent cohort of uniformly assessed and treated patients.

METHODS

Study Design and Patient Samples

The study design uses data from a training cohort to produce a gene expression–based predictor model and then tests the performance of the model using an independent validation cohort (Data Supplement). The training cohort was drawn from patients enrolled onto the E2496 Intergroup trial (ClinicalTrials.gov identifier: NCT00003389). Participating groups included the Eastern Cooperative Oncology Group, the Southwest Oncology Group, the Cancer and Leukemia Group B, and the National Cancer Institute of Canada Clinical Trials Group. This trial included 794 previously untreated patients, 16 years of age or older, with stage III or IV cHL or stage I or II disease with massive mediastinal lymphadenopathy. The trial compared failure-free survival (FFS) and OS between two treatment arms, namely doxorubicin, bleomycin, vinblastine, and dacarbazine (ABVD) and doxorubicin, vinblastine, mechlorethamine, vincristine, bleomycin, etoposide, and prednisone plus planned radiation (Stanford V). The FFS and OS between the two arms, at

Table 1. Demographics and Clinical Characteristics of the Patient Cohorts

Demographic or Clinical Characteristic	E2496 Intergroup Trial Cohort			British Columbia Population-Based Cohort		
	Total Cohort (N = 794)	Training Cohort (n = 290)	P*	Total Cohort (N = 539)	Validation Cohort (n = 78)	P*
Age, years			NS			NS
Median	33	30		31.5	32	
Range	16-83	18-79		16-82	16-85	
Male, %	54	55	NS	56	50	NS
Stage, %			NS			NS
I	4	2		2	0	
II	31	31		44	38	
III	38	42		31	35	
IV	26	26		23	27	
Missing	1	0		0	0	
IPS ≥ 3 (high risk), %†	34	33	NS	41	39	NS
Histologic subtype, %‡			NS			NS
Nodular sclerosis	71	79		75	81	
Mixed cellularity	11	12		6	9	
Other	2	3		4	5	
Not otherwise specified	13	6		15	5	
Missing	4	0		0	0	
EBV-positive HRS cells, %§	ND	16		ND	13	
Primary treatment, %			NS			NS
ABVD	50	50		100	100	
Stanford V	50	50		0	0	
Outcomes						
Median follow-up, years	5.3	5.3	NS	6.5	5.8	NS
5-Year failure-free survival, %	72	70	NS	77	57	< .001
5-Year overall survival, %	88	89	NS	89	76	< .001
Dead at last follow-up, No.	97	35		76	24	

Abbreviations: ABVD, doxorubicin, bleomycin, vinblastine, and dacarbazine; EBV, Epstein-Barr virus; HRS, Hodgkin Reed-Sternberg; IPS, International Prognostic Score; ND, not determined; NS, not significant; Stanford V, doxorubicin, vinblastine, mechlorethamine, vincristine, bleomycin, etoposide, and prednisone plus planned radiation.

*The P values are for comparisons between characteristics of patients in the selected cohorts and the remainder of the patients in the total cohort.

†There were insufficient data to assign patients to an IPS category for 10 patients from the validation cohort and 68 patients from the British Columbia cohort.

‡Patients with not otherwise specified histologic subtype were excluded from the comparisons of histologic subtypes between the cohorts.

§Determined by EBV-encoded RNA in situ hybridization. This failed in one patient each from the training and validation cohorts.

a median follow-up of 5.25 years, were not significantly different,¹⁰ justifying the pooling of patients for the following analyses. The training cohort represents the trial participants for whom pretreatment FFPET biopsies and consent were available. The median follow-up time for living patients was 5.3 years (range, 0.3 to 10.0 years).

The independent validation cohort represents the advanced-stage patients, uniformly treated with ABVD, from the cohort described in Steidl et al.⁸ This cohort was drawn from a population-based registry (Centre for Lymphoid Cancer database, BC Cancer Agency), enriching for treatment failure by sampling patients who had experienced treatment failure more frequently than patients who had not. The population-based registry comprised all patients, 16 years or older, with advanced-stage (systemic symptoms, massive mediastinal lymphadenopathy, and/or stage III/IV disease) cHL uniformly treated with ABVD at the BC Cancer Agency between 1994 and 2007. The median follow-up time for living patients was 5.8 years (range, 1.5 to 16.5 years).

Patients in all cohorts were HIV negative at diagnosis, and all biopsies in the training and validation cohorts were centrally reviewed by R.D.G. and classified according to the WHO 2008 classification.¹¹ The study was approved

by the University of British Columbia–BC Cancer Agency Research Ethics Board.

Gene Expression Analysis

The first 10- μ m section cut from the face of the FFPET block was discarded. Total RNA was extracted from the subsequent 10- μ m section, and gene expression levels were determined on 200 ng of RNA by means of NanoString technology (NanoString Technologies, Seattle, WA). In 12 samples (3%), up to three additional 10- μ m sections were used for RNA extraction because of the small size of the biopsy. After background subtraction, the level of gene expression was normalized using the geometric mean of reference genes *ACTB*, *CLTC*, and *RPLP0*. Reference gene selection and the development and application of quality control criteria for the NanoString data are described in the Data Supplement.

Tissue Microarray

Duplicate 1.5-mm diameter cores from each case were assembled into tissue microarrays. Epstein-Barr virus (EBV) –encoded RNA in situ hybridization was performed to determine EBV infection status of the Hodgkin

Fig 1. Gene expression associated with overall survival (OS) in locally extensive and advanced-stage classical Hodgkin lymphoma. z scores are shown for the 52 genes whose expression levels are significantly associated with OS in the training cohort by univariate Cox regression analysis. The gray dotted lines represent a z score of ± 1.96 . False discovery rates were calculated using the Benjamini-Hochberg method, controlling for 229 comparisons in the training cohort and 52 comparisons in the validation cohort, respectively. (*) Indicates false discovery rate of less than 10%.

Reed-Sternberg (HRS) cells. In the validation cohort, the CD68 immunohistochemistry results were drawn from those reported in Steidl et al.⁸

Predictive Models

Detailed descriptions of model building and model performance assessment are provided in the Data Supplement. In brief, we used the gene expression data from the training cohort to produce a parsimonious predictive model for OS using a penalized Cox model. The global performance of the model was determined by means of the concordance statistic (C-statistic)¹² and time-dependent receiver operating characteristic (ROC) curves.¹³

A threshold for the score derived from the predictive model (the predictor score) that separates patients into low- and high-risk groups was determined in X-tile software (version 3.6.1; Yale University, New Haven, CT) using the score that produced the largest χ^2 value of the log-rank test in the training cohort.

Statistical Methods

Patient baseline characteristics were compared between groups using the Fisher's exact test, χ^2 test, and *t* test. OS was measured from random assignment or initial diagnosis to death from any cause in the training and validation cohorts, respectively. FFS was similarly measured to progression, relapse, or death from any cause. OS and FFS were estimated using the Kaplan-Meier method. Univariate analyses using Cox models were performed to assess the association between the expression levels of individual genes and OS. False discovery rate calculations were performed using the method described by Benjamini and Hochberg.¹⁴

The predictive model, including the threshold, established in the training cohort was tested in the independent validation cohort. As this cohort was enriched for treatment failure, a weighted analysis approach was implemented in R (version 2.13.2; <http://www.r-project.org/>) to remove bias in estimating the relative risk, according to the method of Gray.¹⁵ A weighted log-rank test and weighted Cox proportional hazard models were implemented to test the prognostic ability of the predictor (high v low) when used alone and in combination with other established prognostic factors. Details are supplied in the Data Supplement. All other analysis was performed with SAS software (version 9.2; SAS Institute, Cary, NC). *P* < .05 was considered significant.

RESULTS

Training Cohort

Gene expression was determined for six housekeeping genes and 259 genes of interest (Data Supplement) that were selected by drawing from the literature of suggested prognostic genes^{8,16-19} including components of the microenvironment and cellular processes associated with outcomes in cHL (recently reviewed by Steidl et al⁷). Blocks of FFPET were available for 306 of the 794 patients in the E2496 trial. Data that met quality criteria were produced for 290 of these patients (95%). There were no significant differences in the patient characteristics between the training cohort and the remainder of the patients in the E2496 trial (Table 1). Among the genes of interest, 229 were expressed outside background levels in more than 20% of samples (Data Supplement). In the training cohort, the expression levels of 52 genes were significantly associated with OS in univariate Cox regression analysis (*P* < .05), with 44 being overexpressed and eight being underexpressed in patients who died (Fig 1A). At a 10% false discovery rate, 29 genes were significantly associated with OS. These results are consistent with the previously reported association with unfavorable outcome of *ALDH1A1*,¹⁷ *HSP90AA1*,¹⁸ *LYZ*,^{17,18} *RAPGEF2*,¹⁹ *STAT1*,¹⁷ *TRAF2*,⁸ and *WDR83*.⁸

A predictive model of OS for locally extensive and advanced-stage cHL was produced using data from the training cohort using a penalized Cox model. The model comprised the expression levels of 23 genes, with 20 being overexpressed and three being underexpressed in the patients who had died. The C-statistic for the model was 0.73, and the area under the time-dependent ROC curve was stable from 3 to 7 years at a level between 0.75 and 0.77 (Data Supplement).

Fig 2. Kaplan-Meier estimates of overall survival among patients with locally extensive and advanced-stage classical Hodgkin lymphoma. (A) The training cohort. (B) Weighted analysis of the independent validation cohort giving an unbiased estimate of the predictor's performance in the population-based registry cohort from which the independent validation cohort was drawn. The numbers at risk indicate the number of patients in the validation cohort contributing to the weighted analysis estimates.

Fig 3. The gene expression–based predictor for locally extensive and advanced-stage classical Hodgkin lymphoma tested in the validation cohort. The predictor is a linear equation, adding the $\log_2$ transformed normalized gene expression multiplied by the gene coefficient. (A) The relative gene expression levels of the 23 genes in the predictive model are presented in the form of a heat map. Each column represents a single patient from the validation cohort, arranged according to the predictor score, with lowest score on the left. Each row represents a gene from the model, ordered by hierarchical clustering. The coefficients for each gene, determined in the training cohort, are shown on the right. The vertical dotted blue line separates patients into low- and high-risk groups according to the threshold predictor score. (B) The score from the predictor for patients in the independent validation cohort. The patients are arranged as in panel A. The dotted blue line is placed at the threshold predictor score (0.6235) determined in the training cohort. (C) The clinical and pathology characteristics of the patients in the validation cohort. The patients are ordered as in panel A. EBER, Epstein-Barr virus–encoded RNA; HRS, Hodgkin Reed-Sternberg; IPS, International Prognostic Score; NOS, not otherwise specified.

To demonstrate the clinical utility of the model, a predictor score threshold was determined in the training cohort to separate patients into low- and high-risk groups. The model, including the genes, coefficients, and threshold, is shown in the Data Supplement. In the training cohort, the high-risk group, comprising 29% of the cohort, had a worse OS than the low-risk group (5-year OS: 75% *v* 94%, respectively; Fig 2A). Although this model was trained on OS, the high-risk group also had a worse FFS that the low-risk group (5-year FFS: 51% *v* 77%, respectively; Data Supplement).

Validation Cohort

Gene expression data that met quality criteria were produced in 78 (95%) of the 82 patients selected as the independent validation cohort of patients with advanced-stage cHL uniformly treated with ABVD. Of the 52 genes shown to be associated with OS by univariate analysis in the training cohort, 19 were also significantly associated with OS in the validation cohort at a 10% false discovery rate (Fig 1B).

The predictive model, illustrated in Figure 3, including the selected genes, coefficients, and threshold value established in the train-

ing cohort, was tested in the independent validation cohort. The global performance of the model in the validation cohort was similar to that produced in the training cohort, with a C-statistic of 0.70. The area under the time-dependent ROC curve was stable from 3 to 7 years at a level between 0.74 and 0.77 (Data Supplement).

To estimate the predictor's performance in the population-based registry from British Columbia, a weighted analysis was performed by correcting the proportion of treatment failure in the validation cohort back to that observed in the registry. This analysis was made possible by the observation that the patients in the validation cohort were representative of the patients in the registry cohort when stratified by the presence or absence of treatment failure (Table 2). In the weighted analysis, the high-risk group, identified by the predictor model, had a significantly worse OS compared with the low-risk group (5-year OS: 63% *v* 92%, respectively; weighted log-rank *P* < .001; Fig 2B). The hazard ratio for death imparted by a high-risk score was 6.7 (95% CI, 2.6 to 17.4). The high-risk group also had a significantly worse FFS than the low-risk group (5-year FFS: 59% *v* 82%, respectively; weighted log-rank *P* = .04; Data Supplement).

Table 2. Demographics and Clinical Characteristics of the Validation Cohort and the British Columbia Population-Based Cohort According to Treatment Failure Status

Demographic or Clinical Characteristic	Patients Who Experienced Treatment Failure			Patients Who Did Not Experience Treatment Failure		
	Validation Cohort (n = 39)	British Columbia Cohort (n = 140)	<i>P</i> ^a	Validation Cohort (n = 39)	British Columbia Cohort (n = 399)	<i>P</i> ^a
Age, years			NS			NS
Median	36	36.5		30	32	
Range	19-82	16-85		16-74	16-85	
Male, %	51	56	NS	51	55	NS
WBC > 15 × 10 ⁹ /L, % ^b	24	17	NS	16	15	NS
Lymphocytes < 0.6 × 10 ⁹ /L, % ^c	11	13	NS	5	9	NS
Hemoglobin < 105 g/L, % ^d	28	28	NS	11	18	NS
Albumin < 40 g/L, % ^e	74	78	NS	64	61	NS
Stage IV, %	31	29	NS	23	21	NS
IPS ≥ 3 (high risk), % ^f	56	51	NS	28	35	NS
Nodular sclerosis subtype, % ^g	80	87	NS	92	89	NS
Mass ≥ 10 cm, % ^h	74	62	NS	51	42	NS
Systemic symptoms, % ⁱ	59	67	NS	41	34	NS
Follow-up, years			NS			NS
Median	7.4	7.4		5.1	6.2	
Range	5.5-16.5	2.3-16.5		1.5-12.6	0.1-17.2	
Overall survival, KM estimates, %			NS			
3 years	69	70		NA	NA	NA
5 years	54	59		NA	NA	NA
8 years	45	45		NA	NA	NA

Abbreviations: IPS, International Prognostic Score; KM, Kaplan-Meier; NA, not applicable; NS, not significant.

^aThe *P* values are for comparisons between characteristics of patients in the selected cohorts and the remainder of the patients in the total cohort.

^bData were unavailable for one and 15 patients from the treatment failure validation and population-based registry cohorts, respectively, and one and 16 patients from the no treatment failure validation and population-based registry cohorts, respectively.

^cData were unavailable for four and 22 patients from the treatment failure validation and population-based registry cohorts, respectively, and one and 30 patients from the no treatment failure validation and population-based registry cohorts, respectively.

^dData were unavailable for 14 patients from the treatment failure population-based registry cohort and one and 17 patients from the no treatment failure validation and population-based registry cohorts, respectively.

^eData were unavailable for 12 and 51 patients from the treatment failure validation and population-based registry cohorts, respectively, and six and 90 patients from the no treatment failure validation and population-based registry cohorts, respectively.

^fData were unavailable for seven and 32 patients from the treatment failure validation and population-based registry cohorts, respectively, and three and 36 patients from the no treatment failure validation and population-based registry cohorts, respectively.

^gThe biopsies for four and 27 patients from the treatment failure validation and population-based registry cohorts, respectively, and three and 56 patients from the no treatment failure validation and population-based registry cohorts, respectively, were classified as Hodgkin lymphoma not otherwise specified and were excluded from the comparison.

^hData were unavailable for seven and three patients from the treatment failure and no treatment failure population-based registry cohorts, respectively.

ⁱData were unavailable for one patient from the treatment failure population-based registry cohort.

Table 3. Weighted Pairwise Multivariate Analyses in the Independent Validation Cohort

Variable	Patients		Weighted Pairwise Multivariate Analysis*		
	No.	%	HR	95% CI	P
Pairwise analysis 1					
Predictor score high	17	21.8	8.4	3.3 to 21.6	< .001
IPS ≥ 3 (high risk)†	28	41.2	3.3	1.1 to 9.4	.03
Pairwise analysis 2					
Predictor score high	17	21.8	6.3	2.0 to 20.1	.002
$\geq 25\%$ CD68 ⁺ cells on IHC‡	30	40.0	1.0	0.3 to 3.2	NS
Pairwise analysis 3					
Predictor score high	17	21.8	5.1	2.1 to 12.6	< .001
Age ≥ 45 years	23	29.5	3.1	1.3 to 7.5	.01

Abbreviations: HR, hazard ratio; IHC, immunohistochemistry; IPS, International Prognostic Score; NS, not significant.

*Weighted pairwise multivariate analyses were performed with Cox proportional hazards regression models using only patients where both variables were available. The weighted analyses, performed according to Gray,¹³ provide unbiased estimates in the population registry-based cohort of 539 patient. *P* values are for the correlation between each factor and overall survival.

†There were insufficient data to assign an IPS category to 10 patients.

‡Data are from Steidl et al.⁸

Comparison between the characteristics of the patients in the low- and high-risk groups in both the training and validation cohorts shows that patients in the high-risk group are older and more likely to have high-risk IPS, have EBV-positive HRS cells, and have histologic subtypes other than nodular sclerosis (Fig 3; Data Supplement). EBV status (weighted log-rank $P = .17$) and non-nodular sclerosis histology (weighted log-rank $P = .38$) were not significantly associated with OS in the weighted validation cohort. In contrast, IPS ≥ 3 (weighted log-rank $P = .01$), age ≥ 45 years (weighted log-rank $P = .001$), and CD68 staining $\geq 25\%$ of the cells in the tumor (weighted log-rank $P = .05$) were variables significantly associated with OS in this cohort. To determine whether the predictor's prognostic ability was independent of each of these previously reported and validated prognostic factors at diagnosis, pairwise weighted multivariate analyses were performed (Table 3). The predictor was associated with OS independent of these variables, with IPS ($P = .03$) and age ≥ 45 years ($P = .01$) also reaching statistical significance.

DISCUSSION

We have developed a gene expression-based predictor of OS in advanced-stage cHL applicable to RNA from FFPET that is routinely obtained for diagnosis. It identifies a sizeable proportion of patients at diagnosis with an increased risk of death when treated up-front with ABVD or Stanford V with planned intensified treatment with high-dose chemotherapy and autologous hematopoietic stem-cell transplantation (auto-SCT) for relapsed or refractory disease in younger patients (age < 65 years). Application of the model in a cohort treated similarly with ABVD and planned auto-SCT for younger patients but enriched for primary treatment failure validated this biomarker's ability to identify a population at higher risk of death. The use of a validation cohort enriched for events of interest provides a study design that reduces the use of precious tissue resources. Additionally, in this instance, where the cohort had been assembled for a previous study, it reduces the impact of the potential biases introduced by selecting additional biopsies aimed at reversing this enrichment. The weighted analysis used to estimate the performance of the biomarker in the population-based registry cohort demonstrated that patients in

the high-risk group were at a statistically significant and clinically relevant greater risk of death.

The predictor was developed on, and for, the recently described NanoString platform.⁹ This technology has the potential to penetrate into clinical laboratory diagnostic practice because it has proven to be robust and reliable for quantification of RNA species extracted from FFPET²⁰ and, therefore, might be a suitable platform for a gene expression-based clinical test. Despite the FFPET blocks used in this study being more than 5 years old, sufficient quality of gene expression was obtained in 95% of samples. Used in a prospective manner, where the tissue has been recently fixed, we anticipate that a predictor score would be able to be determined for all patients. Furthermore, the 36-hour turnaround time achieved during this study would make the information produced available to inform decisions regarding up-front treatment.

The predictive model shows that features present in the diagnostic biopsy can portend failure of the treatment package and expands on our previous demonstration that increased numbers of macrophages in the diagnostic biopsy, now validated in numerous studies,²¹ are associated with inferior outcomes, with overexpression of *CD68*, *IL15RA*, *LYZ*, and *STAT1* in patients who die as a result of cHL. The gene signature is consistent with a Th₁ response with relative overexpression of *INFG* and genes regulated by this cytokine, namely *CXCL11*, *IRF1*, *STAT1*, *TNFSF10*, and the genes of major histocompatibility complex class I. Genes associated with cytotoxic T cells/natural killer cells are also overexpressed in patients who die. Demonstration of which cells within the tumor express the genes of the signature and a better understanding of how this relates to mechanisms by which front-line and salvage regimens fail to cure the patient are areas of ongoing research.

Increased numbers of CD68⁺ macrophages²² and a gene expression signature suggestive of a Th₁ immune response¹⁹ have been previously reported in biopsies from patients with EBV-positive cHL. Patients with EBV-positive cHL are over-represented in the high-risk group identified by the predictor, but this signature is also seen in patients who are EBV negative. Thus, latent EBV infection of HRS cells is not the only potential mechanism by which these responses are elicited in cHL. EBV positivity has been associated with reduced

OS,^{23,24} a relationship that seems to be confined to patients older than 45 years of age.^{24,25}

The genes examined in developing this predictor were drawn from a rich literature describing not only individual genes associated with outcome but also representative genes from components of the microenvironment that have been identified by immunohistochemistry and gene expression profiling.⁷ In this way, the predictor harnesses and integrates the prognostic ability of the multitude of previously described biomarkers,^{7,19,26-28} as is illustrated by the loss of significance of the CD68 immunohistochemistry in the pairwise multivariate analysis. Furthermore, the predictor demonstrates prognostic significance independent of the IPS and age ≥ 45 years, which remains one of the few independent factors encompassed by that tool.^{5,29} It is likely that the predictor encompasses multiple aspects of tumor biology and the interaction between the tumor and host immune system.

The two competing approaches to the treatment of advanced-stage CHL that are currently being examined are age specific; for the majority of patients whose age is less than 65 years, the approach is ABVD followed by planned auto-SCT for relapsed or refractory disease or dose-intense up-front treatments such as escalated bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, and prednisone; for patients older than 65 years of age, the approach is ABVD alone. The lack of prognostic biomarkers reliably detectable at diagnosis translates into an inability to safely discriminate between patients for whom the age-appropriate overall treatment ensures a high likelihood of long-term survival and patients in whom it will often fail.³⁰ This information would inform an educated selection of up-front treatment, balancing the risk of treatment failure with that of treatment adverse effects for the individual patient. The predictor model performs this task by identifying a group of patients who have excellent OS with standard treatment, where ABVD could be administered with confidence, and a group in whom this treatment fails in a significant proportion.

The current study establishes the model as a prognostic biomarker in patients treated with ABVD. Further studies are required to determine whether the model is a predictive biomarker, whereby the high risk of death identified by the model can be overcome by dose-intense regimens or novel agents. Once this model has been externally validated and the platform technology shown to be portable, the path

forward for finally introducing a robust biologic outcome predictor into routine clinical practice will be realized, paving the way to truly personalized therapy in Hodgkin lymphoma.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Although all authors completed the disclosure declaration, the following author(s) and/or an author's immediate family member(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: Randy D. Gascoyne, Celgene (C), Roche Canada (C) **Stock**

Ownership: None **Honoraria:** Randy D. Gascoyne, Roche Canada

Research Funding: Randy D. Gascoyne, Roche Canada, Seattle Genetics

Expert Testimony: None **Other Remuneration:** None

AUTHOR CONTRIBUTIONS

Conception and design: David W. Scott, Christian Steidl, Randy D. Gascoyne

Financial support: Randy D. Gascoyne

Administrative support: Randy D. Gascoyne

Provision of study materials or patients: Richard I. Fisher, Lisa M. Rimsza, Nancy L. Bartlett, Bruce D. Cheson, Lois E. Shepherd, Ranjana H. Advani, Joseph M. Connors, Brad S. Kahl, Leo I. Gordon, Sandra J. Horning, Randy D. Gascoyne

Collection and assembly of data: David W. Scott, King L. Tan, Barbara Meissner, Susana Ben-Neriah, Merrill Boyle, Robert Kridel, Adele Telenius, Bruce W. Woolcock, Pedro Farinha, Richard I. Fisher, Lisa M. Rimsza, Nancy L. Bartlett, Bruce D. Cheson, Lois E. Shepherd, Ranjana H. Advani, Joseph M. Connors, Brad S. Kahl, Leo I. Gordon, Sandra J. Horning

Data analysis and interpretation: David W. Scott, Fong Chun Chan, Fangxin Hong, Sanja Rogic, Joseph M. Connors, Christian Steidl, Randy D. Gascoyne

Manuscript writing: All authors

Final approval of manuscript: All authors

REFERENCES

- Borchmann P, Engert A: The past: What have we learned in the past decade. *Hematology Am Soc Hematol Educ Program* 2010:100-107, 2010
- Diehl V, Franklin J, Pfreundschuh M, et al: Standard and increased-dose BEACOPP chemotherapy compared with COPP-ABVD for advanced Hodgkin's disease. *N Engl J Med* 348:2386-2395, 2003
- Engert A, Diehl V, Franklin J, et al: Escalated-dose BEACOPP in the treatment of patients with advanced-stage Hodgkin's lymphoma: 10 years of follow-up of the GHSG HD9 study. *J Clin Oncol* 27:4548-4554, 2009
- Viviani S, Zinzani PL, Rambaldi A, et al: ABVD versus BEACOPP for Hodgkin's lymphoma when high-dose salvage is planned. *N Engl J Med* 365:203-212, 2011
- Moccia AA, Donaldson J, Chhanabhai M, et al: International Prognostic Score in advanced-stage Hodgkin's lymphoma: Altered utility in the modern era. *J Clin Oncol* 30:3383-3388, 2012
- Hasenclever D, Diehl V: A prognostic score for advanced Hodgkin's disease: International Prognostic Factors Project on Advanced Hodgkin's Disease. *N Engl J Med* 339:1506-1514, 1998
- Steidl C, Connors JM, Gascoyne RD: Molecular pathogenesis of Hodgkin's lymphoma: Increasing evidence of the importance of the microenvironment. *J Clin Oncol* 29:1812-1826, 2011
- Steidl C, Lee T, Shah SP, et al: Tumor-associated macrophages and survival in classic Hodgkin's lymphoma. *N Engl J Med* 362:875-885, 2010
- Geiss GK, Bumgarner RE, Birditt B, et al: Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nat Biotechnol* 26:317-325, 2008
- Gordon LI, Hong F, Fisher RI, et al: A randomized phase III trial of ABVD vs. Stanford V +/- radiation therapy in locally extensive and advanced stage Hodgkin's lymphoma: An intergroup study coordinated by the Eastern Cooperative Oncology Group (E2496). *Blood* 116:415, 2010 (abstr)
- Swerdlow SH, Campo E, Harris NL, et al: World Health Organization Classification of Tumours of Haematopoietic and Lymphoid Tissues (ed 4). Lyon, France, IARC Press, 2008
- Uno H, Cai T, Pencina MJ, et al: On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. *Stat Med* 30:1105-1117, 2011
- Heagerty PJ, Lumley T, Pepe MS: Time-dependent ROC curves for censored survival data and a diagnostic marker. *Biometrics* 56:337-344, 2000
- Benjamini Y, Hochberg Y: Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc B* 57:289-300, 1995
- Gray RJ: Weighted analyses for cohort sampling designs. *Lifetime Data Anal* 15:24-40, 2009

16. Devillard E, Bertucci F, Trempat P, et al: Gene expression profiling defines molecular subtypes of classical Hodgkin's disease. *Oncogene* 21:3095-3102, 2002
17. Sánchez-Aguilera A, Montalbán C, de la Cueva P, et al: Tumor microenvironment and mitotic checkpoint are key factors in the outcome of classic Hodgkin lymphoma. *Blood* 108:662-668, 2006
18. Sánchez-Espiridión B, Sánchez-Aguilera A, Montalbán C, et al: A TaqMan low-density array to predict outcome in advanced Hodgkin's lymphoma using paraffin-embedded samples. *Clin Cancer Res* 15:1367-1375, 2009
19. Chetaille B, Bertucci F, Finetti P, et al: Molecular profiling of classical Hodgkin lymphoma tissues uncovers variations in the tumor microenvironment and correlations with EBV infection and outcome. *Blood* 113:2765-2775, 2009
20. Reis PP, Waldron L, Goswami RS, et al: mRNA transcript quantification in archival samples using multiplexed, color-coded probes. *BMC Biotechnol* 11:46, 2011
21. Steidl C, Farinha P, Gascoyne RD: Macrophages predict treatment outcome in Hodgkin's lymphoma. *Haematologica* 96:186-189, 2011
22. Kamper P, Bendix K, Hamilton-Dutoit S, et al: Tumor-infiltrating macrophages correlate with adverse prognosis and Epstein-Barr virus status in classical Hodgkin's lymphoma. *Haematologica* 96:269-276, 2011
23. Diepstra A, van Imhoff GW, Schaapveld M, et al: Latent Epstein-Barr virus infection of tumor cells in classical Hodgkin's lymphoma predicts adverse outcomes in older adult patients. *J Clin Oncol* 27:3815-3821, 2009
24. Jarrett RF, Stark GL, White J, et al: Impact of tumor Epstein-Barr virus status on presenting features and outcome in age-defined subgroups of patients with classic Hodgkin lymphoma: A population-based study. *Blood* 106:2444-2451, 2005
25. Keegan TH, Glaser SL, Clarke CA, et al: Epstein-Barr virus as a marker of survival after Hodgkin's lymphoma: A population-based study. *J Clin Oncol* 23:7604-7613, 2005
26. Alvaro T, Lejeune M, Salvadó MT, et al: Outcome in Hodgkin's lymphoma can be predicted from the presence of accompanying cytotoxic and regulatory T cells. *Clin Cancer Res* 11:1467-1473, 2005
27. Sánchez-Espiridión B, Montalbán C, López A, et al: A molecular risk score based on 4 functional pathways for advanced classical Hodgkin lymphoma. *Blood* 116:e12-e17, 2010
28. Oudejans JJ, Jiwa NM, Kummer JA, et al: Activated cytotoxic T cells as prognostic marker in Hodgkin's disease. *Blood* 89:1376-1382, 1997
29. Guisado-Vasco P, Arranz-Saez R, Canales M, et al: Stage IV and age over 45 years are the only prognostic factors of the International Prognostic Score for the outcome of advanced Hodgkin lymphoma in the Spanish Hodgkin Lymphoma Study Group series. *Leuk Lymphoma* 53:812-819, 2012
30. Borchmann P, Engert A, Diehl V: Chemotherapy: Hodgkin lymphoma—Absence of evidence not evidence of absence! *Nat Rev Clin Oncol* 8:636-637, 2011

Affiliations

David W. Scott, Fong Chun Chan, King L. Tan, Barbara Meissner, Susana Ben-Neriah, Merrill Boyle, Robert Kridel, Adele Telenius, Bruce W. Woolcock, Joseph M. Connors, Christian Steidl, and Randy D. Gascoyne, Centre for Lymphoid Cancer, BC Cancer Agency; Sanja Rogic, Centre for High-Throughput Biology, University of British Columbia, Vancouver, British Columbia; Lois E. Shepherd, National Cancer Institute of Canada Clinical Trials Group, Queen's University, Kingston, Ontario, Canada; Fangxin Hong, Dana-Farber Cancer Institute, Harvard School of Public Health, Boston, MA; Richard I. Fisher, James P. Wilmot Cancer Centre, University of Rochester, Rochester, NY; Lisa M. Rimsza, University of Arizona, Tucson, AZ; Nancy L. Bartlett, Washington University School of Medicine, St Louis, MO; Bruce D. Cheson, Georgetown University Hospital, Washington, DC; Ranjana H. Advani, Stanford University, Stanford; Sandra J. Horning, Genentech, South San Francisco, CA; Brad S. Kahl, University of Wisconsin, Madison, WI; Leo I. Gordon, Northwestern University Feinberg School of Medicine, Chicago, IL; and Pedro Farinha, Centro Hospitalar Lisboa Central, Lisbon, Portugal.