

A DNA Repair Pathway–Focused Score for Prediction of Outcomes in Ovarian Cancer Treated With Platinum-Based Chemotherapy

Josephine Kang, Alan D. D'Andrea, David Kozono

Manuscript received September 19, 2011; revised February 17, 2012; accepted February 28, 2012.

Correspondence to: David Kozono, MD, PhD, Department of Radiation Oncology, Dana-Farber Cancer Institute, 450 Brookline Ave, Boston, MA 02215 (e-mail: dkozono@lroc.harvard.edu).

Background New tools are needed to predict outcomes of ovarian cancer patients treated with platinum-based chemotherapy. We hypothesized that a molecular score based on expression of genes that are involved in platinum-induced DNA damage repair could provide such prognostic information.

Methods Gene expression data was extracted from The Cancer Genome Atlas (TCGA) database for 151 DNA repair genes from tumors of serous ovarian cystadenocarcinoma patients ($n = 511$). A molecular score was generated based on the expression of 23 genes involved in platinum-induced DNA damage repair pathways. Patients were divided into low (scores 0–10) and high (scores 11–20) score groups, and overall survival (OS) was analyzed by Kaplan–Meier method. Results were validated in two gene expression microarray datasets. Association of the score with OS was compared with known clinical factors (age, stage, grade, and extent of surgical debulking) using univariate and multivariable Cox proportional hazards models. Score performance was evaluated by receiver operating characteristic (ROC) curve analysis. Correlations between the score and likelihood of complete response, recurrence-free survival, and progression-free survival were assessed. Statistical tests were two-sided.

Results Improved survival was associated with being in the high-scoring group (high vs low scores: 5-year OS, 40% vs 17%, $P < .001$), and results were reproduced in the validation datasets ($P < .05$). The score was the only pretreatment factor that showed a statistically significant association with OS (high vs low scores, hazard ratio of death = 0.40, 95% confidence interval = 0.32 to 0.66, $P < .001$). ROC curves indicated that the score outperformed the known clinical factors (score in a validation dataset vs clinical factors, area under the curve = 0.65 vs 0.52). The score positively correlated with complete response rate, recurrence-free survival, and progression-free survival (Pearson correlation coefficient [r^2] = 0.60, 0.84, and 0.80, respectively; $P < .001$ for all).

Conclusion The DNA repair pathway–focused score can be used to predict outcomes and response to platinum therapy in ovarian cancer patients.

J Natl Cancer Inst 2012;104:670–681

Epithelial ovarian cancer is the leading cause of gynecologic cancer death and the fifth most common cause of cancer mortality among women in the United States (1). The majority of patients are diagnosed with advanced (ie, stages III and IV) disease, for which the standard treatment is aggressive surgical debulking followed by six to eight cycles of platinum-based chemotherapy, which is typically delivered concurrently with a taxane (2). Because of high toxicity, up to 42% of patients are unable to complete therapy as initially prescribed (3). To date, no good clinical measures for prediction of response exist; as a result, approximately 30% of patients undergo multiple cycles of therapy, with little or no benefit, before they are identified as being chemoresistant. The remaining 70% of patients initially achieve a complete response (CR), but more than 75% relapse within a few years (3). Therefore, it is important to develop prognostic tools to identify patients with worse predicted outcomes

and redirect them to alternate therapies that may be potentially more efficacious, such as radiation (4) or alternate chemotherapeutic agents (eg, topotecan) (5).

Several groups have used microarray-based gene expression profiling to generate prognostic and/or molecular subtype signatures (6–14). For example, Berchuck et al. (9) analyzed 65 patients with serous ovarian cancers, including 54 stage III and stage IV cancers, and built classifiers incorporating up to 26 genes that distinguished short- and long-term survivors (10). These genes represented diverse cellular functions, such as cleavage stimulation factor subunit 3 (CSTF3), which is involved in mRNA processing (15), and ATP-binding cassette, subfamily D, member 3 (ABCD3), which is involved in peroxisome assembly (16). Tothill et al. (11) analyzed 285 serous and endometrioid tumors of the ovary, peritoneum, and fallopian tube, and performed k -means clustering, a

bioinformatics tool for unsupervised class discovery of molecular subtypes (17). They identified six molecular subtypes that represented varying tumor serous low malignant potential tumors and showed relative overexpression of mitogen-activated protein kinase pathway genes (dual specificity phosphatase 4 [DUSP4], dual specificity phosphatase 6 [DUSP6], serine [or cysteine] peptidase inhibitor 5A [SERPINA5], mitogen-activated protein kinase 5 [MAP3K5], and sprout homolog 2 [SPRY2]), and molecular subtype C5 represented a high-grade (ie, histological grade 3) serous ovarian tumor subtype defined by genes expressed in mesenchymal cell development (homeobox A7 [HOXA7], homeobox A9 [HOXA9], homeobox A10 [HOXA10], homeobox D10 [HOXD10], and sex-determining region Y-box 11 [SOX11]). These gene signatures, however, have not yet achieved widespread use. This may be partly because of concerns of reproducibility, which may stem from incorporation of sets of unrelated genes, and computer-based algorithms that do not incorporate biologic rationale during gene selection. In these published studies, the biological importance of the selected genes is discussed retrospectively rather than examined prospectively.

In this study, we hoped to show the strength of using a hypothesis-driven approach to improve reproducibility of gene expression-based scores. We hypothesized that ovarian cancer patients with poor vs favorable outcomes, following platinum-based chemotherapy, have tumors that show differential expression of genes involved in repair of platinum-induced DNA damage. The ataxia-telangiectasia-mutated (ATM) pathway is known to be activated in response to platinum-induced DNA damage (18), resulting in phosphorylation and stabilization of tumor protein p53 (TP53) (19), and may function in concert with Fanconi anemia (FA) genes to maintain genomic integrity (20). It is well known that FA/homologous recombination (FA/HR) pathway genes are critical in repair of complex double-stranded lesions induced by platinum agents (21,22). Nucleotide excision repair (NER) is the principal pathway through which intra- and interstrand crosslinks are repaired (23), along with translesion synthesis (TLS), which allows DNA replication to continue despite blockage from DNA damage (24). Our goal, therefore, was to develop a molecular score based on the expression levels of genes in these pathways that could be predictive for survival and other measures of clinical outcome for ovarian cancer patients treated initially with a standard platinum-based chemotherapy regimen.

Methods

Identification of DNA Repair Pathway Genes

Based on a literature review and our knowledge of DNA repair pathways, we devised a set of 151 DNA repair genes with documented roles in the following DNA repair pathways: ATM (18), base excision repair (BER) (25–28), FA/HR (21,22), mismatch repair (MMR) (29–31), non-homologous end joining (NHEJ) (32–34), NER (23), TLS (24), cross-link repair (XLR) (35–37), recQ helicase pathway (RECQ) (38,39), and other (40–43) (Supplementary Table 1, available online). Based on published data, genes in the ATM (18), FA/HR (21,22), NER (23), and TLS (24) pathways were selected from the set of 151 genes for the development of a score predictive of outcomes following platinum-based therapy.

CONTEXT AND CAVEATS

Prior knowledge

At present, there are no effective prognostic tools for prediction of response in ovarian cancer patients, a majority of whom are diagnosed with an advanced stage (stages III and IV) and undergo surgical debulking followed by and platinum-based chemotherapy.

Study design

Gene expression data was extracted from The Cancer Genome Atlas (TCGA) database for patients with advanced ovarian cancer, and a molecular score was developed by focusing exclusively on the genes involved in platinum-induced DNA damage repair pathways. Patients were divided into low (0–10) and high (11–20) scores, and the prognostic value of the score for overall survival, recurrence-free survival, and progression-free survival was assessed. Data were validated in two independent datasets.

Contribution

Patients with high scores showed statistically significant associations with improved overall survival compared with patients with low scores. The score was predictive of overall survival, recurrence-free survival, and progression-free survival in ovarian cancer patients who received first-line platinum-based chemotherapy.

Implication

This score has the potential to become an important prognostic tool to determine whether advanced-stage ovarian cancer patients will benefit from first-line platinum-based chemotherapy.

Limitation

The score has not been tested prospectively in a clinical trial.

From the Editors

Patient Samples

We extracted clinical data for 511 patients with serous ovarian cystadenocarcinoma from The Cancer Genome Atlas (TCGA) database (44) website (<http://tcga-data.nci.nih.gov>) on February 17, 2011, representing the largest available dataset of epithelial ovarian cancer gene expression profiles (see Supplementary Table 2, available online, for further details on which ovarian cancer samples were included in this study). These were all the patients for whom full sets of tumor gene expression data were available for download. Ovarian cancer samples were categorized by the TCGA into four molecular subtypes on the basis of gene cluster content: 1) immunoreactive, characterized by increased representation of T-cell chemokine ligand genes; 2) differentiated, characterized by genes suggestive of more mature development; 3) proliferative, characterized by high expression of proliferation markers and transcription factors; and 4) mesenchymal, characterized by stromal component genes (44). Clinicopathologic characteristics, and data on molecular subtype, are shown in Table 1. Median year of pathological diagnosis was 2004 (range = 1992–2009), and median age at diagnosis was 59 years (range = 30–89 years). Because our aim was to develop a score predictive of outcomes following standard platinum-based chemotherapy, we focused on the 304 patients with advanced (stages III and IV) ovarian cancer who received a platinum and taxane regimen as first-line treatment. The control group consisted of advanced-stage (stages III and IV) patients (n = 161) who did not receive platinum and taxane therapy.

Table 1. Clinicopathologic characteristics of ovarian cancer patients in The Cancer Genome Atlas (TCGA) dataset*

Characteristic	All patients (n = 511)		Advanced (stage III–IV) patients (n = 464)	
	No. of patients	Median OS (95% CI), y	No. of patients	Median OS (95% CI), y
Age (median 59, range 30–89), y				
≤59	262	4.1 (3.7 to 4.5)	242	4.1 (3.6 to 4.4)
≥60	242	3.2 (2.8 to 3.5)	222	3.0 (2.8 to 3.4)
Stage†				
I	15	6.7 (6.7 to ∞)	—	—
II	23	5.9 (3.7 to ∞)	—	—
III	383	3.7 (3.3 to 4.0)	383	3.7 (3.3 to 4.0)
IV	81	2.7 (2.2 to 4.2)	81	2.7 (2.2 to 4.2)
Grade‡				
1	5	5.4 (4.5 to 6.2)	5	5.4 (4.5 to 6.2)
2	61	4.6 (3.5 to 6.8)	49	4.0 (3.2 to 5.1)
3	428	3.5 (3.2 to 4.0)	398	3.4 (3.0 to 3.8)
Initial chemotherapy				
Platinum and taxane	331	4.0 (3.6 to 4.5)	304	3.8 (3.4 to 4.2)
Cisplatin	43	3.8 (3.0 to 4.7)	42	3.8 (3.0 to 4.7)
Carboplatin	286	4.0 (3.5 to 4.4)	260	3.7 (3.2 to 4.1)
Oxaliplatin	2	—	2	—
Non-platinum and taxane	126	3.7 (3.1 to 4.1)	121	3.5 (3.0 to 4.0)
Response to primary therapy				
Platinum and taxane	289	4.2 (3.7 to 4.6)	264	4.0 (1.8 to 3.0)
CR	208	4.8 (4.1 to 5.9)	188	4.8 (4.1 to 5.7)
Non-CR	81	2.8 (1.8 to 3.2)	78	2.7 (1.8 to 3.0)
Non-platinum and taxane	121	4.0 (3.3 to 4.5)	101	3.6 (3.0 to 4.3)
CR	80	4.7 (4.0 to 5.9)	67	4.4 (3.7 to 4.9)
Non-CR	41	2.5 (1.9 to 3.2)	34	2.5 (1.7 to 3.1)
Surgical debulking				
0–10 mm residual tumor	329	3.7 (3.4 to 4.0)	300	3.5 (3.2 to 3.8)
≥11 mm residual tumor	128	3.0 (2.6 to 3.5)	116	3.1 (2.6 to 3.6)
TCGA tumor/gene subtype				
Differentiated	134	3.7 (3.2 to 4.0)	83	3.5 (2.9 to 4.0)
Immunoreactive	104	4.3 (3.3 to 5.9)	68	4.2 (3.0 to 5.9)
Mesenchymal	107	3.4 (2.6 to 4.1)	67	3.8 (2.4 to 5.2)
Proliferative	135	3.5 (3.0 to 4.0)	79	4.1 (3.2 to 5.0)
N/A	24	6.2 (2.8 to ∞)	7	3.6 (0.8 to 6.2)

* We extracted clinical data for 511 patients with serous ovarian cystadenocarcinoma from TCGA (44) website (<http://tcga-data.nci.nih.gov>) on February 17, 2011. Ovarian cancer samples were categorized by the TCGA into four molecular subtypes on the basis of gene cluster content: immunoreactive, characterized by increased representation of T-cell chemokine ligand genes; differentiated, characterized by genes suggestive of more mature development; proliferative, characterized by high expression of proliferation markers and transcription factors; and mesenchymal, characterized by stromal component genes. Clinicopathologic characteristics are shown, and Kaplan–Meier survival analyses were performed to estimate OS. Because our aim was to develop a score predictive of outcomes in advanced-stage (stages III and IV) ovarian cancer patients, we also present data that is specific for this group. CR = complete response; CI = confidence interval; OS = overall survival; — = not applicable.

† Stage based on International Federation of Gynecology and Obstetrics (FIGO) (46).

‡ Grade based on histological features (47).

Gene Expression Data

Gene expression profiles are routinely acquired by the TCGA for tumors included in their database using three widely used microarray platforms: Agilent Custom 244K (Agilent Technologies, Santa Clara, CA), Affymetrix Human Exon 1.0, and Affymetrix HT_HG-U133A (Affymetrix, Santa Clara, CA). We downloaded from the TCGA website all available data as of February 17, 2011. For our analyses, we used the level 3 data, which provides, for each tumor, expression levels for each gene profiled by the three microarray platforms. The expression values for each gene were normalized by subtracting from it the mean value of the 511 tumor specimens. The three microarray platforms assessed many, but not all, of the same genes. For our analyses, the median normalized values among the three platforms were used.

Construction of a DNA Repair Pathway–Focused Score

Kaplan–Meier log–rank *P* values were calculated to identify a subset of genes whose expression values showed a trend associated with overall survival (OS) (*P* < .15) (Supplementary Figure 1, available online). This cut point, which included the top quartile of genes by *P* value, was selected because representation of each of the DNA repair pathways was lost with more stringent cut points. Also, other cut points were not examined to minimize retrospective optimization. For each patient’s tumor, a point was given for each gene for which higher than median expression was associated with longer survival, and vice versa. The sum of these points constituted our score. The DNA repair pathway–focused score (referred to as “the score”) included only genes in the ATM, FA/HR, NER, and TLS pathways (Table 2). The scores were categorized

Table 2. Genes in platinum-specific DNA repair pathways that were used to construct the score*

Gene	Pathway	Survival†	P‡
<i>ATM</i>	ATM	high	.12
<i>H2AFX</i>	ATM	high	.026
<i>MDC1</i>	ATM	high	.10
<i>RNF8</i>	ATM	high	.020
<i>TOP2A</i>	ATM	high	.11
<i>BRCA2</i>	FA/HR	low	.069
<i>C17orf70</i>	FA/HR	high	.059
<i>FANCB</i>	FA/HR	high	.11
<i>FANCE</i>	FA/HR	high	.055
<i>FANCF</i>	FA/HR	high	.006
<i>FANCG</i>	FA/HR	high	.047
<i>FANCI</i>	FA/HR	high	.13
<i>PALB2</i>	FA/HR	high	.034
<i>MUS81</i>	FA/HR	high	.11
<i>NBN</i>	FA/HR	low	.083
<i>SHFM1</i>	FA/HR	high	.12
<i>DDB1</i>	NER	high	.045
<i>ERCC8</i>	NER	low	.11
<i>RAD23A</i>	NER	high	.034
<i>XPA</i>	NER	low	.14
<i>MAD2L2</i>	TLS	high	.11
<i>POLH</i>	TLS	high	.15
<i>UBE2I</i>	TLS	high	.049

* The 23-gene set was constructed with genes in the ATM (18), Fanconi Anemia (FA)/homologous recombination (HR) (21,22), nucleotide excision repair (NER) (23), and translesion synthesis (TLS) (24) pathways. *ATM* = ataxia telangiectasia mutated; *BRCA2* = breast cancer 2, early onset; *C17orf70* = chromosome 17 open reading frame 70 (FAAP100); *DDB1* = damage-specific DNA binding protein 1; *ERCC8* = excision repair cross-complementing rodent repair deficiency, complementation group 8; *FANCB* = Fanconi anemia, complementation group B; *FANCE* = Fanconi anemia, complementation group E; *FANCF* = Fanconi anemia, complementation group F; *FANCG* = Fanconi anemia, complementation group G; *FANCI* = Fanconi anemia, complementation group I; *H2AFX* = histone 2A family, member X; *MDC1* = mediator of DNA damage checkpoint 1; *RNF8* = ring finger protein 8; *MAD2L2* = mitotic arrest deficient-like 2; *MUS81* = MUS81 endonuclease homolog (*S. cerevisiae*); *NBN* = nibrin; *PALB2* = partner and localizer of BRCA 2 (FANCN); *POLH* = polymerase (DNA directed), eta; *RAD23A* = RAD23 homolog A (*S. cerevisiae*); *SHFM1* = split hand/foot malformation (ectrodactyly) type 1; *TOP2A* = topoisomerase (DNA II alpha); *UBE2I* = ubiquitin-conjugating enzyme E2I; *XPA* = xeroderma pigmentosum, complementation group A.

† For each gene, “high” means higher than median gene expression was associated with improved overall survival in The Cancer Genome Atlas dataset, and “low” means higher than median gene expression was associated with worse overall survival (*P* values are presented). Kaplan–Meier analyses were performed to estimate survival.

‡ *P* values were calculated using a two-sided log-rank test.

as “low” (ie, scores 1–10) or “high” (ie, scores 11–20), based on the range of scores observed in the study, which divided the patients into two categories with equal ranges; other thresholds were not examined except, as noted, to minimize retrospective optimization.

Some pathways (BER, NHEJ, MMR, RECQ, XLR, and “other”) were excluded because of less-established or no known role in platinum-induced DNA damage repair. To further explore the use of biologic rationale in molecular score construction, we used the 17 genes from these pathways to form an alternate 17-gene score.

Validation Datasets

Two previously published ovarian cancer datasets with openly accessible microarray and clinical data [Berchuck et al. (9) and Tothill et al. (11); referred as “Berchuck dataset,” and “Tothill

dataset,” respectively] were used as validation datasets. After completion of all analyses using the TCGA dataset, the analyses were replicated in these two datasets for validation of the results, to the extent feasible based on available data.

Statistical Analysis

Median and 5-year OS for each scoring group (high vs low) were calculated using the Kaplan–Meier method, and statistical significance was assessed using the log-rank test. Hazard ratios (HRs) for all-cause mortality and 95% confidence intervals (CIs) were calculated using the Cox proportional hazards analysis. Assumptions of proportionality were confirmed for the Cox proportional hazards analyses by generating Kaplan–Meier survival estimate curves (eg, for high vs low scoring groups), and observing that the curves did not intersect with each other. These analyses were performed with the TCGA dataset, four TCGA molecular subtype subsets (immunoreactive, differentiated, proliferative, and mesenchymal), and the two validation datasets (Berchuck and Tothill). Distribution of scores among the four molecular subtypes was assessed as a continuous and as a categorical variable using analysis of variance (ANOVA). CR, partial response (PR), stable disease (SD), and progressive disease (PD) were assessed per Response Evaluation Criteria In Solid Tumors (RECIST) criteria (45). Univariate and multivariable analyses were performed using Cox proportional hazards models incorporating the score (low vs high) and known prognostic clinical factors, including response to primary therapy (CR vs non-CR), age at diagnosis (≤ 59 vs ≥ 60 years), International Federation of Gynecology and Obstetrics (FIGO) stage (46) (III vs IV), histological grade (47) (1–2 vs 3), and extent of surgical debulking (0–10 vs ≥ 11 mm residual tumor), as categorical variables.

To assess the potential clinical impact of the score, multivariable analyses using the Cox proportional hazards models were performed with the score as a continuous variable, and/or the following set of clinical variables: age as a continuous variable, grade 1–4 or unknown as a categorical variable, and stages IIIA–C or stage IV as a categorical variable (Table 3). None of the clinical variables showed *P* values less than .05, unlike the score, but all were included so that receiver operating characteristic (ROC) curves could be generated to compare the models with and without inclusion of the score. To generate the ROC curves, patients were classified as surviving either longer or shorter than the median OS, excluding patients who were alive for durations less than the median OS at last follow-up. Logistic fit of low vs high survival category by cumulative hazard (the product of the hazard ratios of each incorporated variable) was performed. Area under the curve (AUC) values were calculated from the ROC curves. For validation, the HRs calculated above were applied to the Berchuck dataset, which was the one that included data on the same clinical variables.

The Pearson correlation coefficient was calculated to assess the relationship between the score and each of the following: median OS, recurrence-free survival (RFS), progression-free survival (PFS), and CR, excluding the single patient in the TCGA dataset with a score of 3, which appeared to be an outlier. The Kaplan–Meier method and the log-rank test were used to assess RFS after achieving a CR, as well as PFS, in patients with low vs high scores. The proportion of patients in each treatment response group (CR, PR, SD, or PD) was compared between low- and high-scoring

Table 3. Cox proportional hazards model using relevant pretreatment predictors*

Variables	HR (95% CI)	P†
Age, grade, and stage		
Age, y‡	1.0051 (0.99 to 1.021)	.53
Grade		
1	1.00 (referent)	
2	1.36 (0.38 to 8.67)	.67
3	1.45 (0.44 to 8.93)	.59
4	2.47 (0.11 to 26.24)	.49
Unknown	1.54 (0.29 to 11.34)	.62
Stage		
IIIA	1.00 (referent)	
IIIB	0.93 (0.22 to 6.27)	.92
IIIC	0.87 (0.27 to 5.35)	.85
IV	0.93 (0.27 to 5.88)	.93
Age, grade, stage, and score		
Age, y‡	1.0039 (0.99 to 1.021)	.65
Grade		
1	1.00 (referent)	
2	1.41 (0.39 to 9.0063)	.64
3	1.64 (0.50 to 10.087)	.47
4	1.34 (0.061 to 14.40)	.82
Unknown	1.72 (0.33 to 12.72)	.53
Stage		
IIIA	1.00 (referent)	
IIIB	0.68 (0.16 to 4.66)	.65
IIIC	0.53 (0.16 to 3.26)	.42
IV	0.64 (0.18 to 4.049)	.58
Score§	0.86 (0.81 to 0.91)	<.001
Score alone		
Score§	0.86 (0.82 to 0.91)	<.001

* Multivariable analyses using Cox proportional hazards models were performed in advanced-stage ovarian cancer patients who received platinum and taxane therapy as first-line treatment, to compare models for overall survival with and without inclusion of the score, and score alone. Score was assessed as a continuous variable, age as a continuous variable, grade 1–4 or unknown as a categorical variable, and stage IIIA–C or stage IV as a categorical variable. CI = confidence interval; HR = hazard ratio.

† Two-sided *P* values were calculated using the Cox proportional hazards models.

‡ HR of death for each year of age at pathological diagnosis greater than 34.

§ HR of death for each unit score greater than 2.

patients. Observational data suggests that ovarian cancer patients with *BRCA1* and/or *BRCA2* mutations may have better overall outcome (48); thus, the association between likelihood of CR based on score and *BRCA1* and/or *BRCA2* mutation status was assessed using the likelihood ratio test. Here, the score was divided into three categories (≤ 7 , 8–13, and ≥ 14) to better distinguish patients with differing CR rates. All statistical analyses were performed using SAS JMP 9.2 (SAS Institute Inc, Cary, NC). All statistical tests were two-sided, and all *P* values less than .05 were considered statistically significant.

Results

Prognostic Value of the Score in Survival of Ovarian Carcinoma Patients Treated with Platinum-Based Chemotherapy

From the initial set of 151 DNA repair genes, we selected 23 genes from four pathways associated with platinum-induced DNA damage (ATM, FA/HR, NER, and TLS pathways), whose expression levels showed an association ($P < .15$; this cut point was chosen

because representation of each DNA repair pathway was lost with more stringent cut points) with OS in serous ovarian cystadenocarcinoma patients ($n = 511$) in the TCGA dataset (Table 2 and Supplementary Figure 1, available online). Median OS for the entire dataset was 3.7 years (range = 8 days to not reached). A molecular score was devised for each ovarian cancer patient by assigning a point for each gene for which higher than median expression was associated with longer survival, and vice versa, and obtaining the sum of these points. We divided patients into two groups by the median score—patients with low scores (scores 1–10) and patients with high scores (scores 11–20). This cut point seemed to be the simplest to apply clinically and because it divided a similar number of patients into each group. The association between OS (median OS and 5-year OS) and low vs high score was analyzed. Low scores were associated with worse OS (low vs high score, median OS = 3.0 years [95% CI = 2.8 to 3.3 years] vs 4.5 years [95% CI = 4.0 to 4.9 years]; 5-year OS = 17% [95% CI = 11% to 24%] vs 40% [95% CI = 33 to 49%]; log-rank $P < .001$). The score demonstrated greater ability to distinguish worse vs improved survival outcomes in patients who received a platinum and taxane regimen as first-line chemotherapy (low vs high score, median OS = 3.2 years [95% CI = 2.7 to 3.5 years] vs 4.7 years [95% CI = 4.2 to 5.7 years]; 5-year OS = 17% [95% CI = 10% to 28%] vs 46% [95% CI = 36% to 57%]; log-rank $P < .001$) (Figure 1, A) compared with those who did not receive this treatment (low vs high score, median OS = 3.3 years [95% CI = 2.8 to 4.0 years] vs 4.0 years [95% CI = 3.0 to 4.8 years]; 5-year OS = 19% [95% CI = 10% to 34%] vs 31% [95% CI = 29% to 46%]; log-rank $P = .017$) (data not shown in the figure).

Next, we performed a univariate analysis to assess whether the score was associated with OS for patients treated with a platinum and taxane regimen in the TCGA dataset from which it was derived. Low scores were associated with worse OS and high scores with improved OS (high vs low scores, HR of death = 0.44, 95% CI = 0.33 to 0.61, $P < .001$). We then assessed whether the score was associated with OS in two validation sets and found a similar statistically significant association in the Berchuck and Tothill datasets (high vs low scores, Berchuck dataset, HR of death = 0.33, 95% CI = 0.13 to 0.86, $P = .013$; Tothill dataset, HR of death = 0.61, 95% CI = 0.36 to 0.99, $P = .044$) (Figure 1, B). All patients in the Berchuck dataset received platinum-based therapy, and the score identified low vs high survival subtypes (low vs high score, median OS, 1.8 years [95% CI = 0.8 to 2.8 years] vs 2.9 years [95% CI = 2.3 to ∞]; 5-year OS, 0% vs 35% [95% CI = 15% to 63%], $P = .013$). The Tothill dataset had a mixed population containing patients treated with platinum and non-platinum-based chemotherapy, and as a whole did not show a statistically significant association between the score and OS (low vs high score, log-rank $P = .88$). However, by narrowing the analysis to patients who received a platinum and taxane regimen, we found that low score was statistically significantly associated with poor prognosis and high score with better prognosis (low vs high score, median OS = 3.6 years [95% CI = 2.8 to 4.3 years] vs 4.1 years [95% CI = 3.1 to 5.7 years]; 5-year OS, 18% [95% CI = 7% to 42%] vs 47% [95% CI = 32% to 63%]; log-rank $P = .044$).

We hypothesized that there would be decreased reproducibility of a score with incorporation of genes without biologic rationale.

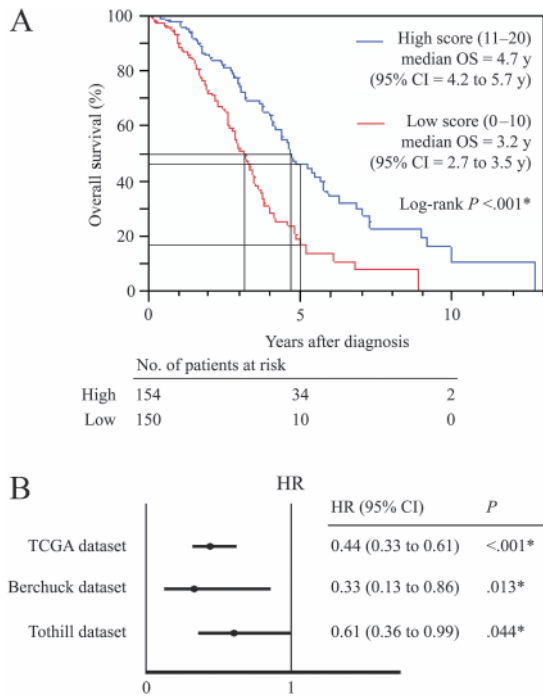


Figure 1. DNA repair pathway–focused score in prognosis of overall survival. For each patient’s tumor, a point was given for each DNA repair gene for which higher than median expression was associated with longer survival, and vice versa. The sum of these points constituted the score. Only genes in pathways related to platinum-induced damage repair were included. Advanced-stage (stages III and IV) ovarian cancer patients from The Cancer Genome Atlas (TCGA) dataset who received a platinum and taxane regimen as first-line chemotherapy ($n = 304$) were arbitrarily divided into low (scores 1–10) vs high (scores 11–20) scores. **A**) Kaplan–Meier analysis was used to assess median overall survival (OS) and 5-year OS (indicated by black lines) from time of pathological diagnosis in low- and high-scoring subgroups. $P < .001$, calculated using a two-sided log-rank test. **B**) Univariate analysis was performed using the Cox proportional hazards regression analyses to assess whether the score was prognostic for OS in the TCGA dataset. **Solid circles** represent hazard ratio (HR) of death and **open-ended horizontal lines** represent the 95% confidence intervals (CIs). This was validated in the two published datasets by Berchuck et al. (9) and Tothill et al. (11). * $P < .05$; all P values were calculated using Cox proportional hazards analysis.

Therefore, we repeated our analyses using an alternate 17-gene score based on DNA repair pathways unrelated to repair of platinum DNA damage (Supplementary Table 3, available online). When this alternate score was applied to the two validation sets, we found no statistically significant associations between low vs high scores and OS (log-rank $P = .29$ and $.48$, in the Berchuck and Tothill datasets, respectively) (Supplementary Figure 2, available online).

Survival analyses were repeated in the four ovarian cancer molecular subtypes (immunoreactive, differentiated, proliferative, and mesenchymal) identified by the TCGA based on cluster gene content (44), to ascertain whether there was an association between subtype and survival outcomes. We found no statistically significant difference in survival between the four subtypes in our subset of advanced-stage platinum- and taxane-treated patients (log-rank $P = .33$), or in the entire TCGA dataset (log-rank $P = .084$), for which subtype information was provided. We then applied the score to each of the individual subtypes (Supplementary Figure 3,

available online). We observed the greatest statistical significance when the score was applied to the mesenchymal subtype (low vs high score, median OS, 2.6 years [95% CI = 1.7 to 4.0 years] vs 7.3 years [95% CI = 3.7 to 9.9 years]; 5-year OS, 18% [95% CI = 8% to 40%] vs 79% [95% CI = 50% to 93%]; log-rank $P = .0016$). Distribution of score was also assessed as a continuous and categorical variable; statistically significant differences in score between the subtypes were observed by ANOVA ($P < .001$ for both).

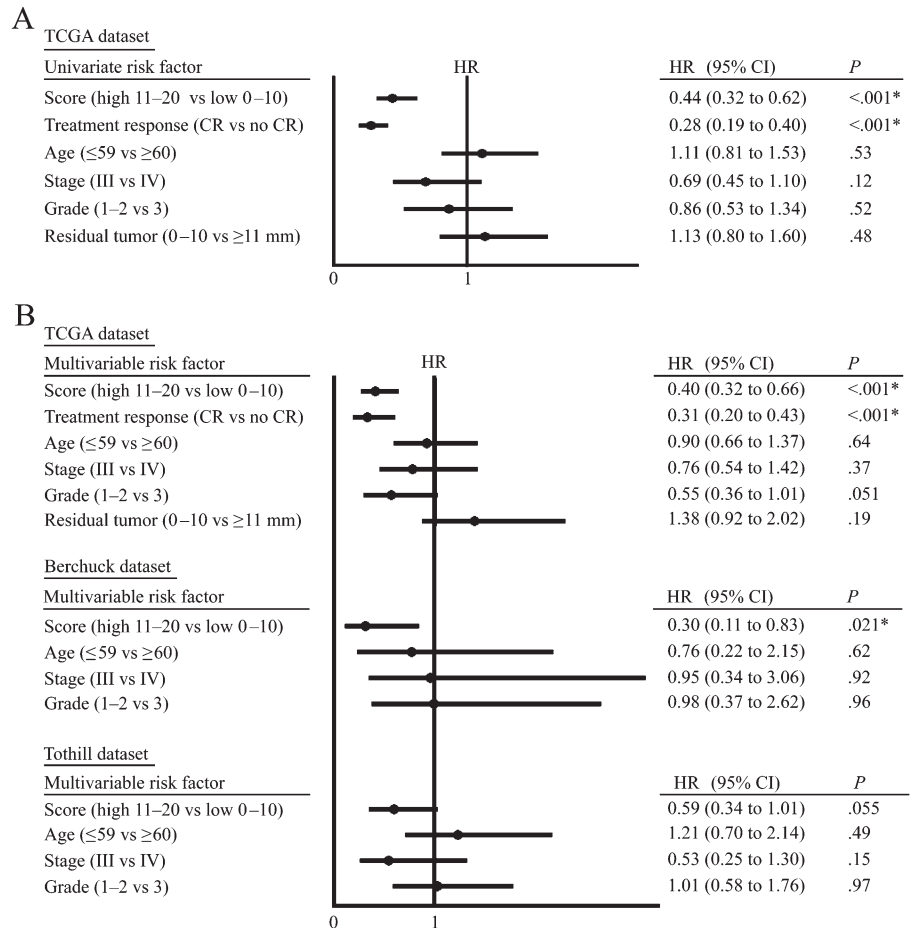
Predictive Accuracy of the Score Compared With Prognostic Clinical Factors

We examined the predictive accuracy of the score compared with other clinical factors by performing the Cox proportional hazards univariate (Figure 2, A) and multivariable (Figure 2, B) analyses using the following factors as categorical variables: score (high vs low), treatment response (CR vs none), age (≤ 59 vs ≥ 60 years), FIGO stage (III vs IV), histological grade (1–2 vs 3), and extent of surgical debulking (0–10 vs ≥ 11 mm residual tumor). In the TCGA dataset, only the score and treatment response were statistically significantly associated with OS (high vs low scores, HR of death = 0.40, 95% CI = 0.32 to 0.66, $P < .001$; CR vs no CR to treatment, HR of death = 0.31, 95% CI = 0.20 to 0.43, $P < .001$). Multivariable analysis was repeated in the two validation sets. The score outperformed other pretreatment clinical covariates; low score was consistently associated with poor prognosis and high score with better prognosis in both validation datasets (Berchuck dataset, HR of death = 0.30, 95% CI = 0.11 to 0.83, $P = .021$; and Tothill dataset, HR of death = 0.59, 95% CI = 0.34 to 1.01, $P = .055$) (Figure 2, B), whereas all other clinical covariates failed to show a consistent association with prognosis. Although we used high vs low score as a categorical variable in this analysis, the score also correlated with OS when used as a continuous variable (Pearson correlation coefficient [r^2] = 0.47) (Supplementary Figure 4, available online).

To assess the contribution of the score toward prediction of response to first-line platinum and taxane therapy, the Cox proportional hazards statistical models containing relevant pretreatment predictors (age, grade, stage, and score) were constructed (Table 3), and ROC analyses were performed using the following: age, grade, and stage (AGS); age, grade, stage, and score (AGS + score); and score alone. For these analyses, survival was classified as either higher than or lower than the median OS of 3.7 years, excluding patients who were alive for durations less than the median OS at last follow-up. This demonstrated AUC values of 0.60, 0.71, and 0.70, for AGS, AGS + score, and score alone, respectively (Figure 3, A). The Berchuck dataset, with median OS of 2.4 years, was the only dataset with comparable details available for age, histological grade (1–4), and FIGO stage (IIIA, IIIB, IIIC, and IV), and was used for validation; the AUC values were 0.52, 0.65, and 0.65 for AGS, AGS and score, and score alone, respectively (data not shown in the figure).

Next, we analyzed the performance of the score in the four ovarian cancer subtypes identified in the TCGA dataset, to determine whether the score had particularly high predictive accuracy in certain subtypes. Thus, ROC analyses were repeated in the four ovarian cancer subtypes and demonstrated particularly high predictive accuracy of the score in the mesenchymal subtype, with AUC values of 0.60, 0.86, and 0.87 for AGS, AGS and score, and

Figure 2. Comparison of the score with prognostic clinical covariates. Univariate and multivariable Cox proportional hazards regression analyses incorporating the score (high [scores 11–20] vs low [scores 0–10]) and known prognostic clinical factors, including response to primary therapy (complete response [CR] vs non-CR) by Response Evaluation Criteria In Solid Tumors (RECIST) criteria (45), age at diagnosis (≤ 59 vs ≥ 60 years), International Federation of Gynecology and Obstetrics (FIGO) (46) stage (III vs IV), grade (1–2 vs 3), and extent of surgical debulking (0–10 vs ≥ 11 mm residual tumor); each as categorical variables. **Solid circles** represent the hazard ratio (HR) of death and **open-ended horizontal lines** represent the 95% confidence intervals (CIs). * $P < .05$; all P values were calculated using Cox proportional hazards analysis. **A)** Univariate analysis was performed using Cox proportional hazards regression analyses in The Cancer Genome Atlas (TCGA) dataset of patients with advanced-stage (stages III and IV) ovarian cancer treated with platinum and taxane chemotherapy. **B)** Multivariable analysis was performed in the TCGA and two validation datasets by Berchuck et al. (9) and Tothill et al. (11), adjusting for the same categorical variables.



score alone, respectively (Figure 3, B). The remaining subtypes had AUC ranges of 0.48–0.69, 0.60–0.69 and 0.58–0.69 for AGS, AGS and score, and score alone, respectively (data not shown in the figure).

Probability of Achieving CR Based on the Score

Overall, 188 (71%) of 264 of patients with advanced epithelial ovarian cancer achieved a CR to a platinum and taxane regimen (Table 1). To determine whether the score correlated with likelihood of CR, we plotted the percentage of patients achieving CR against the score (Pearson correlation coefficient [r^2] = 0.60, $P < .001$) (Figure 4). Patients were classified into lowest, middle, and highest tertiles of scores and showed low likelihood (score ≤ 7 ; 44% patients), intermediate likelihood (score 8–13; 73% patients), and high likelihood (score ≥ 14 ; 80%) of CR. Improved survival was noted with increasing likelihood of CR (low likelihood, median OS = 2.1 years, 95% CI = 1.6 to 2.6 years; intermediate likelihood, median OS = 3.8 years, 95% CI = 3.4 to 4.5 years; and high likelihood, median OS = 4.6 years, 95% CI = 4.1 to 5.9 years). Interestingly, for a score of 14 or higher, the likelihood of CR appeared to plateau at approximately 80%, though median OS continued to increase with higher score. For example, the subset of patients ($n = 43$) with score 14–15 and the subset of patients ($n = 32$) with score 16 or higher had similar likelihoods of CR, but a markedly improved median OS was observed in the higher-scoring group (scores 14–15 vs ≥ 16 , 79% vs 81% likelihood of CR, median OS, 4.6 years [95% CI = 2.9 to 5.9 years] vs 7.2 years [95% CI = 2.7 to ∞ years]).

Probability of CR and BRCA Mutation Status

In our dataset of advanced platinum- and taxane-treated patients ($n = 304$), 215 had DNA sequencing data available, with 34 (16%) germline and seven (3%) somatic *BRCA1* and/or *BRCA2* mutations. *BRCA* mutant tumors have defects in homologous recombination (49). Patients with germline *BRCA* mutations have tumors with increased sensitivity to platinum therapy and improved survival (50), likely secondary to compromised repair of DNA damage. Tumors with somatic *BRCA* mutations appear less frequently (51) but are hypothesized to exhibit a “BRCAness” phenotype with defective homologous recombination and similarly improved outcomes. We therefore examined the breakdown of *BRCA1* and *BRCA2* germline and somatic mutations by scoring category (≤ 7 , 8–13, and ≥ 14), which corresponded to low, intermediate, and high likelihood of CR (44%, 73%, and 80% likelihood of CR), and found that the percentage of *BRCA* germline and somatic mutations increased correspondingly (13% germline and 0% somatic mutations for scoring category ≤ 7 , 15% germline and 1% somatic mutations for scoring category 8–13, and 19% germline and 5% somatic mutations for scoring category ≥ 14), although this did not reach statistical significance (likelihood ratio test $P = .18$) (Supplementary Figure 5, available online).

Prediction of Recurrence-Free Survival and Progression-Free Survival

Despite achieving a CR, the majority of patients with advanced epithelial ovarian cancer eventually relapse and die of disease.

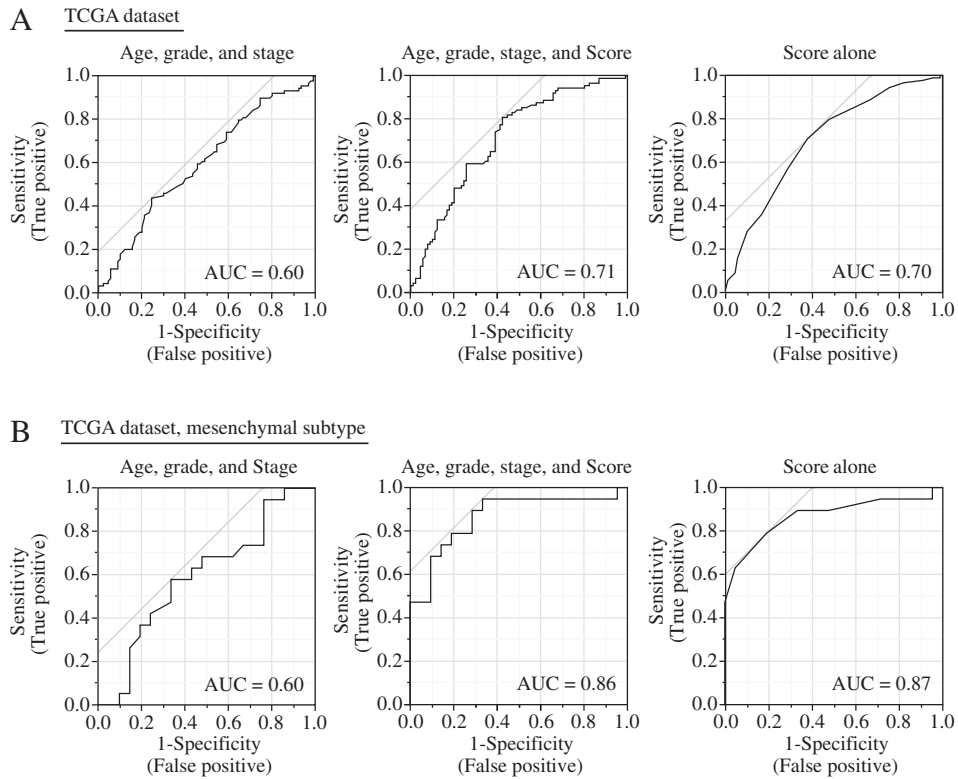


Figure 3. Receiver operating characteristic (ROC) analysis of the score and clinical covariates in predicting overall survival. The area under the curve (AUC) was calculated for ROC curves, and sensitivity and specificity was calculated to assess score performance. **A)** Using statistical models constructed based on multivariable Cox proportional hazards, ROC curves were calculated incorporating clinical variables of age, grade, and stage (left); age, grade, stage, and score (middle); and score alone (right). **B)** ROC

curves, including only patients with tumors of mesenchymal TCGA subtype, were also calculated incorporating clinical variables of age, grade, and stage (left); age, grade, stage, and score (middle); and score alone (right). Grey lines indicate the 45° angle tangent line marked at a point that provides best discrimination between true positives and false positives, assuming that false positives and false negatives have similar costs.

Although, historically 70% of patients achieve a CR (4), this is not the best indicator of overall outcome. This is reflected in our dataset; patient subsets with scores of 14–15 vs 16 both had relatively high likelihood of CR (Figure 4), but disparate median OS, as stated earlier. To explore whether the score could predict additional

measures of outcome, we examined RFS in patients who achieved a CR to platinum and taxane treatment (Figure 5, A and B). The score was statistically significantly associated with RFS in this cohort (low [0–10] vs high [11–20] scores, median RFS, 1.3 years [95% CI = 1.1 to 1.6 years] vs 1.8 years [95% CI = 1.6 to 2.2 years]; 5-year RFS, 11% [95% CI = 4% to 24%] vs 22% [95% CI = 14% to 34%]; log-rank $P = .021$); there was a statistically significant positive correlation between score and duration of RFS ($r^2 = 0.84$, $P < .001$) (Figure 5, B). We analyzed patients with scores of 14–15 vs 16 or higher and found a trend toward differences in RFS (score 14–15 vs score 16 or higher, median RFS, 1.8 years [95% CI = 1.3 to 2.2 years] vs 3.8 years [95% CI = 0.8 to ∞]; 5-year RFS, 13% [95% CI = 4 to 32%] vs 43% [95% CI = 20 to 67%]; log-rank $P = .076$). On multivariable analysis (Figure 5, C), only the score was statistically significantly associated with RFS (HR of recurrence = 0.66, 95% CI = 0.45 to 0.98, $P = .039$) after adjusting for age, stage, grade, and extent of surgical debulking. We did not have an appropriate dataset in which to validate these findings.

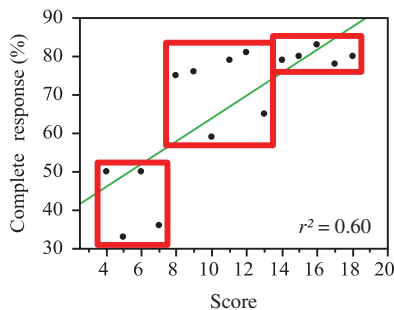


Figure 4. Correlation of score with complete response (CR). Advanced-stage (stages III and IV) ovarian cancer patients from The Cancer Genome Atlas (TCGA) dataset who received platinum and taxane as first-line chemotherapy ($n = 304$ patients) were analyzed based on their individual scores. For each patient's tumor, a point was given for each DNA repair gene for which higher than median expression was associated with longer survival, and vice versa. The sum of these points constituted our score. The percentage of patients achieving CR based on the Response Evaluation Criteria In Solid Tumors (RECIST) criteria was calculated for each score value and is represented by the **black solid circles**. The Pearson correlation coefficient (r^2) was used to assess the relationship between the score and likelihood of CR. Patients were classified into lowest (score ≤ 7), middle (score 8–13), and highest (score ≥ 14) tertiles (shown in **boxes**). The straight line depicts the least squares linear regression line through the data points.

We also examined the ability of the score to predict PFS and found a statistically significant difference between scoring groups (low [scores 0–10] vs high [scores 11–20] scoring groups, median PFS, 1.1 years [95% CI = 0.9 to 1.2 years] vs 1.6 years [95% CI = 1.3 to 1.8 years]; 5-year PFS, 2.3% [95% CI = 0.4 to 13%] vs 19% [95% CI = 13 to 28%]; log-rank $P < .001$) (Figure 5, D). There was a positive correlation between score and duration of PFS (Pearson correlation coefficient [r^2] = 0.80, $P < .001$) (Figure 5, E).

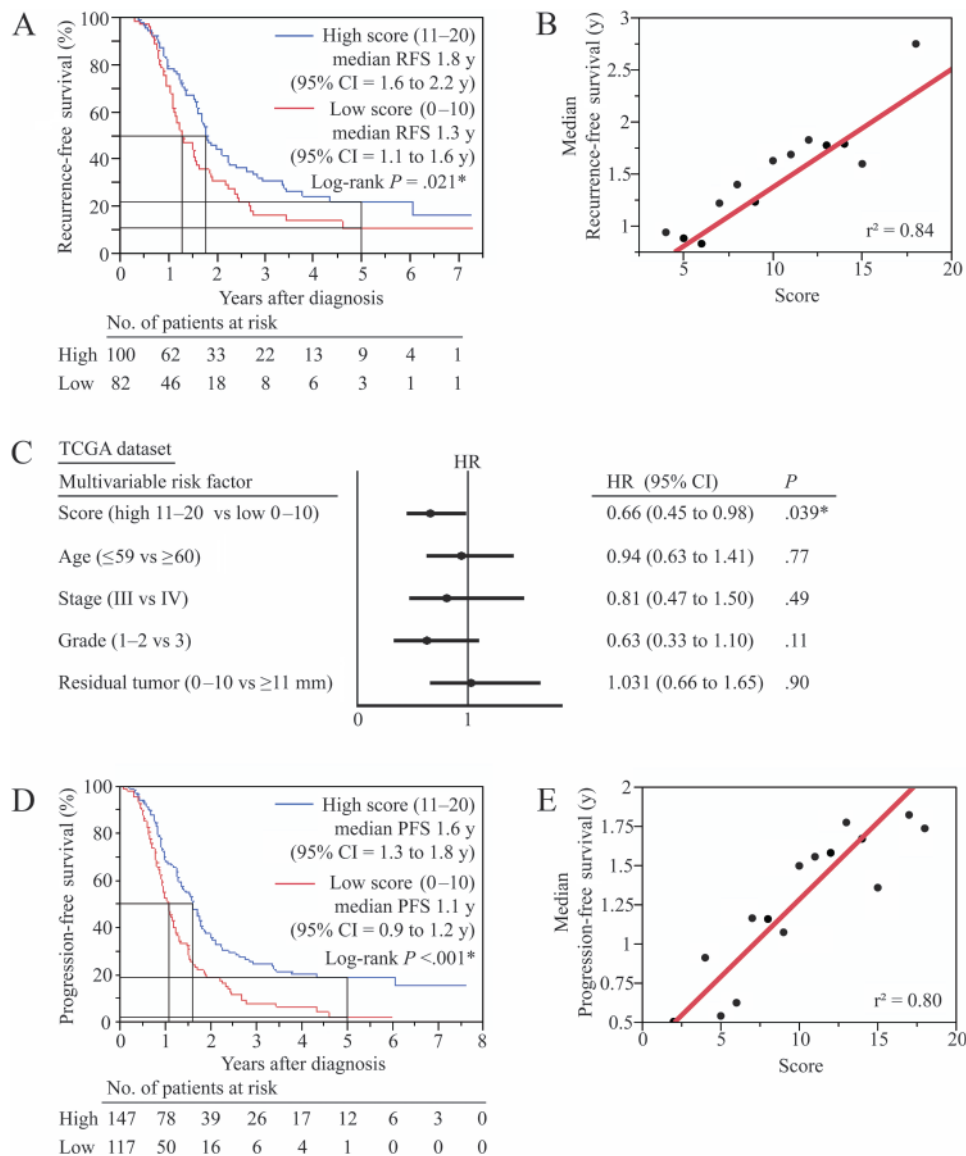


Figure 5. Ability of the score to predict recurrence-free survival (RFS) and progression-free survival (PFS). For each patient's tumor, a point was given for each DNA repair gene for which higher than median expression was associated with longer survival, and vice versa. The sum of these points constituted our score. Only genes in pathways related to platinum-induced damage repair were included. Advanced-stage (stages III and IV) ovarian cancer patients from The Cancer Genome Atlas (TCGA) dataset who received a platinum and taxane regimen as first-line chemotherapy ($n = 304$ patients) were analyzed to assess the relationship between score and RFS and PFS. **A**) The association of score and RFS was assessed in the TCGA ovarian cancer patients who achieved a complete response (CR) after receiving the first-line platinum and taxane therapy. The Kaplan-Meier method was used to compare RFS in patients with low (scores 1–10) vs high (scores 11–20) scores. * P value was calculated using a two-sided log-rank test. **B**) TCGA ovarian cancer patients who achieved a CR after the first-line platinum and taxane therapy were analyzed, by calculating median RFS for each score subgroup, represented by the **black solid circles**. The Pearson correlation coefficient (r^2) was calculated to assess the relationship between score and RFS. The straight line depicts the least

squares linear regression line through the data points. **C**) Multivariable analysis of factors that impact RFS was performed in TCGA ovarian cancer patients who achieved a CR following platinum and taxane chemotherapy. Cox proportional hazards regression was performed for score (high, 11–20, vs low, 1–10), treatment response (CR vs no CR), age (≤ 59 vs ≥ 60 years), International Federation of Gynecology and Obstetrics (FIGO) stage (III vs IV), grade (1–2 vs 3), and extent of surgical debulking (0–10 vs ≥ 11 mm). **Solid circles** represent the hazard ratio (HR) of death and **open-ended horizontal lines** represent the 95% confidence intervals (CIs). * P values were calculated using a two-sided log-rank test. **D**) The association of score and PFS was assessed in the TCGA ovarian cancer patients with available data. The Kaplan-Meier method was used to compare PFS in patients with low (scores 1–10) vs high (scores 11–20) scores. * P value was calculated using a two-sided log-rank test. **E**) TCGA ovarian cancer patients with PFS data were analyzed, by calculating median PFS for each score subgroup, represented by the **black solid circles**. The Pearson correlation coefficient (r^2) was calculated to assess the relationship between score and PFS. The straight line depicts the least squares linear regression line through the data points.

We then examined the distribution of responses in the low (score 1–10) vs high (score 11–20) scoring group (Supplementary Figure 6, available online), to evaluate the association between the lower score and likelihood of poor response to treatment as measured by RECIST criteria. Patients with non-CR to a platinum

and taxane regimen were categorized as having a PR, SD, or PD by RECIST criteria. Out of 265 patients with known response data, 21 (8%) patients had PD, with 13 patients (11%) in the low-scoring group and eight patients (5%) in the high-scoring group, suggesting enrichment of patients with PD in the low-scoring

group. This is consistent with the poor prognosis of patients with progressive or platinum-refractory disease.

Discussion

In this study, we demonstrate that a hypothesis-driven approach can be used to produce a reproducible gene expression–based score. By selecting genes in pathways known to be involved in repair of platinum-induced DNA damage, we produced a score that is predictive of OS, PFS, and RFS in ovarian cancer patients following platinum-based chemotherapy. Moreover, we were able to validate our results in two additional datasets of advanced-stage ovarian cancer patients treated with platinum-based chemotherapy and demonstrate that the score is prognostic for survival only when it is composed of genes from relevant DNA repair pathways, further strengthening the use of biologic rationale in molecular signature construction.

We were also able to demonstrate that the score outperforms other known clinical factors in predicting OS, not only in the TCGA dataset but also in two additional validation sets. Developing the ability to predict OS and outcomes to chemotherapy using prognostic markers such as the score is critical, particularly in ovarian cancer, because there are presently no other good clinical measures to predict response to standard platinum-based chemotherapy. As a result, although 30% of patients receive no benefit from standard platinum-based chemotherapy, many patients undergo multiple cycles of futile, potentially toxic treatment.

With further validation, the score can become an important tool that can be used in such advanced-stage ovarian cancer patients before initiation of first-line therapy to help direct them toward treatments with the greatest benefit/risk ratio. Patients in the lowest scoring category of 7 or less, with a median OS of 2.1 years, stand to benefit the most from alternate therapies, such as enrollment in phase I clinical trials or treatments that target alternate DNA repair pathways, such as radiation. A Gynecologic Oncology Group (GOG) phase I study, GOG-9915, was recently published using low-dose abdominal radiation along with docetaxel in recurrent epithelial ovarian cancer patients, with acceptable toxicity and promising results in a subset of patients; 30% of patients with measurable disease achieved greater than 6 months PFS (4). In addition, topotecan, a topoisomerase poison that has an alternate mechanism of action than platinum, has been tested in clinical trials with demonstrated efficacy in platinum-resistant cohorts (5).

Conversely, patients in the highest scoring category of 14 and above derive the greatest most durable benefit from platinum and taxane chemotherapy. The enrichment of *BRCA1/2* germline and somatic mutations in this group suggests this subset exhibits a BRCA-like phenotype (“BRCAness”). High CR rates to platinum, improved OS, and defective HR are characteristics of “BRCAness” that predict for excellent response to poly-ADP ribose polymerase (PARP) inhibition (51,52). This is consistent with recently published results showing even among *BRCA*-mutant epithelial ovarian cancer tumors, platinum-sensitive tumors exhibit the highest response rates (61%) compared with platinum-resistant and platinum-refractory *BRCA*-mutant tumors (41% and 15%, respectively) (53). Therefore, we hypothesize that patients in our highest scoring category are the subset of patients that would benefit most from

PARP inhibitor therapy, regardless of BRCA status. This would greatly expand the number of patients who stand to benefit from this promising treatment.

Gene expression profiling has immense potential to aid in predicting patient outcomes. The recent release of TCGA gene expression data in epithelial ovarian cancer is unprecedented in size and comprehensiveness (44). By mining this resource, and applying biologic rationale, we have created a durable score that provides predictive information regarding a tumor’s intrinsic sensitivity or resistance to first-line platinum and taxane chemotherapy. It outperforms traditional predictors of clinical outcome (age, grade, and stage) for both OS and RFS, with high predictive accuracy. A statistical model including age, grade, and stage was notably improved by addition of the score and performed trivially better than the score alone based on AUC results. The fact that the score performs particularly well in the mesenchymal tumor subtype suggests that this group is especially dependent on DNA repair capability and merits further investigation.

The study has a few limitations. Although we recapitulated our findings in two published datasets to the extent possible based on data availability, the score has not yet been tested prospectively in a clinical trial. We believe the score is ready for such testing, which must be performed before more widespread adoption in ovarian cancer patients as a prognostic tool. Also, gene expression profiling captures only a subset of cancer genetic and epigenetic changes. Many genes are not regulated at the transcriptional level, and thus their expression levels are noninformative. Other mechanisms of regulation include microRNAs (54), protein phosphorylation (55), and ubiquitination (56). To address this problem, we selected only those genes whose expression levels were associated with survival. High vs low levels of relative gene expression do not necessarily reflect function; for instance, a high level of expression may reflect an attempt to compensate for a defective pathway. More direct assays of DNA repair pathway activity may address this but would be more complex to implement in routine clinical practice.

In conclusion, by using a hypothesis-driven approach, we have generated a DNA repair pathway–focused score that could be validated in two additional datasets. With additional prospective validation in clinical trials, we hope that the score can become a powerful tool that is useful in stratifying advanced-stage ovarian cancer patients toward optimal treatments incorporating new treatment regimens vs current standard of care.

References

1. Jemal A, Siegel R, Xu J, et al. Cancer statistics, 2010. *CA Cancer J Clin*. 2010;60(5):277–300.
2. Pignata S, Cannella L, Leopardo D, et al. Chemotherapy in epithelial ovarian cancer. *Cancer Lett*. 2011;303(2):73–83.
3. Barakat RR, Markman M, Randall ME. *Principles and Practice of Gynecologic Oncology*. 5 ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2009.
4. Kunos CA, Sill MW, Buekers TE, et al. Low-dose abdominal radiation as a docetaxel chemosensitizer for recurrent epithelial ovarian cancer: a phase I study of the Gynecologic Oncology Group. *Gynecol Oncol*. 2011;120(2):224–228.
5. Herzog TJ. Clinical experience with topotecan in relapsed ovarian cancer. *Gynecol Oncol*. 2003;90(3, pt 2):S3–S7.
6. Schwartz DR, Kardia SL, Shedden KA, et al. Gene expression in ovarian cancer reflects both morphology and biological behavior, distinguishing

- clear cell from other poor-prognosis ovarian carcinomas. *Cancer Res.* 2002;62(16):4722–4729.
7. Zorn KK, Bonome T, Gangi L, et al. Gene expression profiles of serous, endometrioid, and clear cell subtypes of ovarian and endometrial cancer. *Clin Cancer Res.* 2005;11(18):6422–6430.
 8. Shridhar V, Lee J, Pandita A, et al. Genetic analysis of early- versus late-stage ovarian tumors. *Cancer Res.* 2001;61(15):5895–5904.
 9. Berchuck A, Iversen ES, Lancaster JM, et al. Patterns of gene expression that characterize long-term survival in advanced stage serous ovarian cancers. *Clin Cancer Res.* 2005;11(10):3686–3696.
 10. Spentzos D, Levine DA, Ramoni MF, et al. Gene expression signature with independent prognostic significance in epithelial ovarian cancer. *J Clin Oncol.* 2004;22(23):4700–4710.
 11. Tothill RW, Tinker AV, George J, et al. Novel molecular subtypes of serous and endometrioid ovarian cancer linked to clinical outcome. *Clin Cancer Res.* 2008;14(16):5198–5208.
 12. Jazaeri AA, Awtrey CS, Chandramouli GV, et al. Gene expression profiles associated with response to chemotherapy in epithelial ovarian cancers. *Clin Cancer Res.* 2005;11(17):6300–6310.
 13. Hartmann LC, Lu KH, Linette GP, et al. Gene expression profiles predict early relapse in ovarian cancer after platinum-paclitaxel chemotherapy. *Clin Cancer Res.* 2005;11(6):2149–2155.
 14. Konstantinopoulos PA, Spentzos D, Karlan BY, et al. Gene expression profile of BRCAness that correlates with responsiveness to chemotherapy and with outcome in patients with epithelial ovarian cancer. *J Clin Oncol.* 2010;28(22):3555–3561.
 15. Park JS, Young Yoon S, Kim JM, et al. Identification of novel genes associated with the response to 5-FU treatment in gastric cancer cell lines using a cDNA microarray. *Cancer Lett.* 2004;214(1):19–33.
 16. Tanaka AR, Tanabe K, Morita M, et al. ATP binding/hydrolysis by and phosphorylation of peroxisomal ATP-binding cassette proteins PMP70 (ABCD3) and adrenoleukodystrophy protein (ABCD1). *J Biol Chem.* 2002;277(42):40142–40147.
 17. Do JH, Choi DK. Clustering approaches to identifying gene expression patterns from DNA microarray data. *Mol Cells.* 2008;25(2):279–288.
 18. Yazlovitskaya EM, Persons DL. Inhibition of cisplatin-induced ATR activity and enhanced sensitivity to cisplatin. *Anticancer Res.* 2003;23(3B):2275–2279.
 19. Efeyan A, Serrano M. p53: guardian of the genome and policeman of the oncogenes. *Cell Cycle.* 2007;6(9):1006–1010.
 20. Kennedy RD, Chen CC, Stuckert P, et al. Fanconi anemia pathway-deficient tumor cells are hypersensitive to inhibition of ataxia telangiectasia mutated. *J Clin Invest.* 2007;117(5):1440–1449.
 21. D'Andrea AD. The Fanconi Anemia/BRCA signaling pathway: disruption in cisplatin-sensitive ovarian cancers. *Cell Cycle.* 2003;2(4):290–292.
 22. Powell SN, Kachnic LA. Therapeutic exploitation of tumor cell defects in homologous recombination. *Anticancer Agents Med Chem.* 2008;8(4):448–460.
 23. Basu A, Krishnamurthy S. Cellular responses to cisplatin-induced DNA damage. *J Nucleic Acids.* 2010;2010. pii: 201367.
 24. Hicks JK, Chute CL, Paulsen MT, et al. Differential roles for DNA polymerases eta, zeta, and REV1 in lesion bypass of intrastrand versus interstrand DNA cross-links. *Mol Cell Biol.* 2010;30(5):1217–1230.
 25. Kim YJ, Wilson DM III. Overview of base excision repair biochemistry [published online ahead of print]. *Curr Mol Pharmacol.* 2011.
 26. Zharkov DO. Base excision DNA repair. *Cell Mol Life Sci.* 2008;65(10):1544–1565.
 27. Duncan T, Trewick SC, Koivisto P, et al. Reversal of DNA alkylation damage by two human dioxygenases. *Proc Natl Acad Sci U S A.* 2002;99(26):16660–16665.
 28. Hazra TK, Izumi T, Kow YW, et al. The discovery of a new family of mammalian enzymes for repair of oxidatively damaged DNA, and its physiological implications. *Carcinogenesis.* 2003;24(2):155–157.
 29. Lin Y, Wilson JH. Diverse effects of individual mismatch repair components on transcription-induced CAG repeat instability in human cells. *DNA Repair (Amst).* 2009;8(8):878–885.
 30. Martinez SL, Kolodner RD. Functional analysis of human mismatch repair gene mutations identifies weak alleles and polymorphisms capable of polygenic interactions. *Proc Natl Acad Sci U S A.* 2010;107(11):5070–5075.
 31. Acharya S, Wilson T, Gradia S, et al. hMSH2 forms specific mismatch-binding complexes with hMSH3 and hMSH6. *Proc Natl Acad Sci U S A.* 1996;93(24):13629–13634.
 32. Barker CA, Powell SN. Enhancing radiotherapy through a greater understanding of homologous recombination. *Semin Radiat Oncol.* 2010;20(4):267–273.e3.
 33. Gomes BC, Silva SN, Azevedo AP, et al. The role of common variants of non-homologous end-joining repair genes XRCC4, LIG4 and Ku80 in thyroid cancer risk. *Oncol Rep.* 2010;24(4):1079–1085.
 34. Takedachi A, Saijo M, Tanaka K. DDB2 complex-mediated ubiquitylation around DNA damage is oppositely regulated by XPC and Ku and contributes to the recruitment of XPA. *Mol Cell Biol.* 2010;30(11):2708–2723.
 35. Hemphill AW, Bruun D, Thrun L, et al. Mammalian SNM1 is required for genome stability. *Mol Genet Metab.* 2008;94(1):38–45.
 36. Hlavín EM, Smeaton MB, Miller PS. Initiation of DNA interstrand cross-link repair in mammalian cells. *Environ Mol Mutagen.* 2010;51(6):604–624.
 37. Marini F, Kim N, Schuffert A, et al. POLN, a nuclear PolA family DNA polymerase homologous to the DNA cross-link sensitivity protein Mus308. *J Biol Chem.* 2003;278(34):32014–32019.
 38. Khakhar RR, Cobb JA, Bjergbaek L, et al. RecQ helicases: multiple roles in genome maintenance. *Trends Cell Biol.* 2003;13(9):493–501.
 39. Hickson ID. RecQ helicases: caretakers of the genome. *Nat Rev Cancer.* 2003;3(3):169–178.
 40. Bradsher J, Coin F, Egly JM. Distinct roles for the helicases of TFIIH in transcript initiation and promoter escape. *J Biol Chem.* 2000;275(4):2532–2538.
 41. Marini F, Wood RD. A human DNA helicase homologous to the DNA cross-link sensitivity protein Mus308. *J Biol Chem.* 2002;277(10):8716–8723.
 42. Broderick S, Rehmet K, Concannon C, et al. Eukaryotic single-stranded DNA binding proteins: central factors in genome stability. *Subcell Biochem.* 2010;50:143–163.
 43. Balajee AS, May A, Dianova I, et al. Efficient PCNA complex formation is dependent upon both transcription coupled repair and genome overall repair. *Mutat Res.* 1998;409(3):135–146.
 44. Cancer Genome Atlas Research N. Integrated genomic analyses of ovarian carcinoma. *Nature.* 2011;474(7353):609–615.
 45. Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer.* 2009;45(2):228–247.
 46. Benedet JL, Bender H, Jones H III, et al. FIGO staging classifications and clinical practice guidelines in the management of gynecologic cancers. FIGO Committee on Gynecologic Oncology. *Int J Gynaecol Obstet.* 2000;70(2):209–262.
 47. International Federation of Gynecology and Obstetrics. Classification and staging of malignant tumours in the female pelvis. *Acta Obstet Gynecol Scand.* 1971;50(1):1–7.
 48. Gallagher DJ, Konner JA, Bell-McGuinn KM, et al. Survival in epithelial ovarian cancer: a multivariate analysis incorporating BRCA mutation status and platinum sensitivity. *Ann Oncol.* 2011;22(5):1127–1132.
 49. Tutt AN, Lord CJ, McCabe N, et al. Exploiting the DNA repair defect in BRCA mutant cells in the design of new therapeutic strategies for cancer. *Cold Spring Harb Symp Quant Biol.* 2005;70:139–148.
 50. Cass I, Baldwin RL, Varkey T, et al. Improved survival in women with BRCA-associated ovarian carcinoma. *Cancer.* 2003;97(9):2187–2195.
 51. McCabe N, Turner NC, Lord CJ, et al. Deficiency in the repair of DNA damage by homologous recombination and sensitivity to poly(ADP-ribose) polymerase inhibition. *Cancer Res.* 2006;66(16):8109–8115.
 52. Tan DS, Rothermundt C, Thomas K, et al. “BRCAness” syndrome in ovarian cancer: a case-control study describing the clinical features and outcome of patients with epithelial ovarian cancer associated with BRCA1 and BRCA2 mutations. *J Clin Oncol.* 2008;26(34):5530–5536.
 53. Fong PC, Yap TA, Boss DS, et al. Poly(ADP)-ribose polymerase inhibition: frequent durable responses in BRCA carrier ovarian cancer correlating with platinum-free interval. *J Clin Oncol.* 2010;28(15):2512–2519.

54. Landau DA, Slack FJ. MicroRNAs in mutagenesis, genomic instability, and DNA repair. *Semin Oncol.* 2011;38(6):743–751.
55. Keith CT, Schreiber SL. PIK-related kinases: DNA repair, recombination, and cell cycle checkpoints. *Science.* 1995;270(5233):50–51.
56. Dianov GL, Meisenberg C, Parsons JL. Regulation of DNA repair by ubiquitylation. *Biochemistry (Mosc).* 2011;76(1):69–79.

Funding

This work was supported by the Department of Radiation Oncology, Dana-Farber Cancer Institute, Boston, MA (internal funds).

Notes

We thank Matthew Oesting for help with data processing and Dr Akila Viswanathan for helpful comments during article preparation. The authors are solely responsible for the study design, data collection, analysis and interpretation of the data, writing the article, and decision to submit the article for publication.

Affiliations of authors: Harvard Radiation Oncology Program, Boston, MA (JK); Department of Radiation Oncology, Dana-Farber Cancer Institute, Boston, MA (AD'A, DK).