DOI: 10.1093/jnci/djs203
Advance Access publication on April 13, 2012.

Molecular Scores to Predict Ovarian Cancer Outcomes: A Worthy Goal, but Not Ready for Prime Time

Elizabeth M. Swisher, Toshiyasu Taniguchi, Beth Y. Karlan

Correspondence to: Elizabeth M. Swisher, MD, Department of Obstetrics and Gynecology, University of Washington, 1959 NE Pacific St, Box 356460, Seattle, WA 98195-6460 (e-mail: swishere@u.washington.edu).

Ovarian carcinoma has the highest mortality of all gynecological cancers. The American Cancer Society estimates that in 2012, about 22 280 new cases of ovarian cancer will be diagnosed, and 15 500 women will die of ovarian cancer in the United States (1). Despite achieving high rates of remission following radical surgery and platinum-based chemotherapy, most women relapse

and ultimately die of chemoresistant disease. Advances in chemotherapy lengthen survival for women with advanced-stage (stages III and IV) disease but have not changed the likelihood of cure. Biomarkers, such as the rate of decline of serum cancer antigen 125 (CA-125, also known as mucin-16) level or the absolute CA-125 nadir, can be predictors of progression-free and overall survivals (2–4); however, when faced with a slowly declining level of CA-125 during primary treatment, the oncologist has few effective alternatives. With almost 80% of primary ovarian cancers initially responding to platinum-based therapy, a prospective biomarker would need to be highly predictive of treatment failure because alternative treatments would be purely experimental.

The advent of high-throughput molecular profiling technologies has led to numerous studies that defined ovarian carcinoma phenotypes by various profiles, including mRNA, protein, microRNA, DNA copy number, etc. In unsupervised analyses, the molecular profiles are used to sort the carcinomas into different categories by statistical algorithms without assumptions about the type or number categories present. In contrast, in supervised analyses, the investigator defines the characteristics of interest (such as response to chemotherapy) in a test set and determines which subset components of the molecular profile are differentially expressed, thereby defining a signature for the phenotype of interest. Because of the large number of comparisons made, statistically significant associations can occur by chance, and the "novel" signature will invariably perform well within the test set from which it was derived. Therefore, to assess reproducibility and clinical performance, the signature must be assessed in one or more "validation sets" of unrelated specimens.

The Cancer Genome Atlas (TCGA) network recently compiled detailed molecular datasets on more than 500 serous ovarian carcinomas and identified an expression profile signature that was associated with survival (5). This signature was then validated in several publicly available datasets. In this issue of the Journal, Kang et al. (6) have taken a different approach to the publicly available TCGA data. They applied a hypothesis-driven approach to test whether expression of specific DNA repair genes can predict clinical outcomes after platinum-based therapy. A score was developed based on the expression of 23 genes in DNA repair pathways that were thought to be responsible for repair of platinum-induced DNA damage and had shown a moderate association with survival ($P < .15$) in the TCGA set. They defined the score as a simple sum (up to 23) for each sample based on whether expression of each of the 23 genes deviated from the median expression in the direction favoring survival. The score was then categorized as "low" (ie, scores 1–10) or "high" (ie, scores 11–20), and clinical performance was assessed in two independent publicly available validation sets. The authors conclude that their DNA repair score can be used to predict outcomes and response to platinum therapy in ovarian cancer patients and is ready for application in a prospective clinical trial. Strengths of their analyses include the use of a large well-annotated set of serous ovarian carcinomas, use of comprehensive profiling with multiple platforms, and the development of a hypothesis-driven predictor.

The underlying hypothesis was that malignant cells deficient in DNA repair (such as homologous recombination and the Fanconi anemia pathway) will demonstrate a more durable response to platinum-based chemotherapy. This is a reasonable hypothesis, given that carcinomas from women with germline *BRCA1/2* mutations are deficient in homologous recombination and have better overall survival compared with women with sporadic ovarian carcinoma (7). This phenotype has been somewhat elusively referred to as "BRCAness" (8); however, no one knows how to best define such a homologous recombination–deficient or platinum-sensitive phenotype. Konstantinopoulos et al. (9) tried to define a *BRCA1/2*-deficient (BRCA-like) phenotype with expression profiling, but interestingly, none of the 60 genes defined in this signature overlap with any of the 23 genes used in the score developed by Kang et al. (6).

Contrary to the initial hypothesis of Kang et al. (6), high expression of the DNA repair genes, rather than low expression, was associated with improved overall survival [table 2 in (6)]. This finding is counterintuitive and without an obvious biological basis. The authors suggest that high expression of DNA repair genes may reflect an attempt to compensate for a defective pathway, a worthwhile and testable hypothesis. Alternatively, increased expression of many DNA repair genes may actually lead to reduced DNA repair efficiency by disrupting the tight regulation normally present (10). Finally, alterations in one or more transcription factors could affect expression of multiple genes in the score (11).

The score was intended to predict platinum response and thus help direct therapy. But does their score reliably predict outcome and should we test it in a clinical arena? We would expect the score to perform well in the "test set," that is, the TCGA samples from which it was derived. An alternative score derived from 17 DNA repair genes that were "not of interest" (which means the DNA repair genes that have not been implicated in cellular resistance to platinum) was also predictive in the TCGA samples [$P < .001$, supplementary Figure 2, A, available online, in (6)] demonstrating that evaluating the biomarker within the test set is not a valid assessment of performance. The authors' conclusion that the 23-gene score outperformed known prognostic factors in TCGA samples is not meaningful, given that none of the known prognostic factors were actually predictive in the TCGA dataset. Clinical performance can only be meaningfully assessed in independent sample sets (the validation sets). Two validation sets, Tothill dataset (12) and Berchuck dataset (13), were used to establish clinical validity, but both were small. The score was not statistically significantly associated with survival in the larger (ie, Tothill) dataset until the authors eliminated one-third of the cases treated with platinum but not taxane, an analysis counterintuitive to the underlying hypothesis. And even in that selected subset, the score was not statistically significantly associated with survival in multivariable analysis ($P = .055$). In the second (ie, Berchuck) dataset, the score was statistically significantly associated with survival in multivariable analysis ($P = .021$), but that dataset included only 54 advanced-stage cancers, too few to rely on exclusively for validation. The receiver operating characteristic curves could be determined only for the smaller validation set, generating a limited area under the curve of 0.65, not sufficiently predictive for clinical decision making. In fact, the DNA repair score (6), turned out to be a less powerful prognosticator than the predictive expression profile already reported by TCGA using an unsupervised discovery approach and using same validation sets (5).

So what lessons have we learned from this study? First, the vast amount of TCGA data is challenging to access and analyze. TCGA used multiple platforms including three for expression profiling, two for exome sequencing, and five for copy number analyses, making integration of data difficult. Kang et al. (6) do not provide a detailed explanation of how they integrated data from the three expression platforms, but their methods appear to differ substantially from those used in the published TCGA analyses (5). This could explain discrepancies in data from the two analyses. For example, in evaluating the association of *BRCA1* expression with survival (5), *BRCA1* expression met the TCGA cut point ($P < .01$) and was included in the TCGA prognostic predictor. However, in the analysis by Kang et al. (6), *BRCA1* expression was not associated with survival despite using a much lower cut point ($P < .15$) and despite overlap of the vast majority of samples between the two studies. We conclude that the application of different analytical methods to TCGA data will generate different findings. The lack of impact of known prognostic factors (stage, grade, age, and surgical cytoreduction) in TCGA is troublesome and will further complicate attempts to develop biomarkers from this dataset. TCGA samples were collected for molecular annotation, not chosen to be representative or to answer specific outcomes questions, likely explaining why the TCGA set behaves differently than most other collections. Therefore, the utility of TCGA in the identification of generalizable biomarkers remains uncertain.

Another important lesson is the need to better understand the molecular phenotype to successfully apply a hypothesis-driven approach. Our understanding of the regulatory mechanisms (somatic mutation, copy number alteration, changes in epigenetic or microRNA regulation, etc.) impacting the various DNA repair pathways and their interactions with each other remains incomplete. Ultimately, it may be more accurate to directly measure functional markers of DNA repair in neoplastic cells (such as DNA damage–induced formation of RAD51 nuclear foci), which might require fresh tissue samples and exposure to DNA-damaging agents (14,15). The molecular heterogeneity of high-grade serous ovarian carcinoma may result in one phenotype dominating in the early platinum responsive malignancy and another in the resistant recurrence, further complicating the picture. It may be important to define a homologous recombination-deficient phenotype in "real-time," perhaps, even during the course of treatment. To date, the only clinically validated biomarker of homologous recombination function in ovarian carcinomas is *BRCA1/2* germline mutation status, but even that can be circumvented in malignant cells by the development of secondary mutations that restore functional *BRCA1/2* in the tumor (16–21). Therefore, the plasticity of response to chemotherapy may be as important as the initial phenotype.

The analysis of the TCGA dataset by Kang et al. (6) was unable to prove their hypothesis or validate their counterintuitive DNA repair score in a large independent dataset. Although their score is not yet ready for clinical application, it has raised some interesting questions: 1) what is the best way to assess DNA repair function in solid malignancies, and 2) with such redundancy in the DNA repair pathways, how do we define the driver vs passenger events and classify the important vs unimportant pathways and genes? The development of predictive biomarkers is critical to the successful application of individualized therapy. The study by Kang et al. (6) is an important effort in that direction but demonstrates the challenges we face as we attempt to utilize large datasets to develop personalized genomic medicines. The premature application of inadequately validated biomarkers may adversely impact the successful implementation of individualized therapies.

References

1. American Cancer Society. http://www.cancer.org/Cancer/OvarianCancer/DetailedGuide/ovarian-cancer-key-statistics Accessed March 27, 2012.
2. van Altena AM, Kolwijck E, Spanjer MJ, et al. CA125 nadir concentration is an independent predictor of tumor recurrence in patients with ovarian cancer: a population-based study. *Gynecol Oncol.* 2010;119(2):265–269.
3. Juretzka MM, Barakat RR, Chi DS, et al. CA125 level as a predictor of progression-free survival and overall survival in ovarian cancer patients with surgically defined disease status prior to the initiation of intraperitoneal consolidation therapy. *Gynecol Oncol.* 2007;104(1):176–180.
4. Gadducci A, Zola P, Landoni F, et al. Serum half-life of CA 125 during early chemotherapy as an independent prognostic variable for patients with advanced epithelial ovarian cancer: results of a multicentric Italian study. *Gynecol Oncol.* 1995;58(1):42–47.
5. Network TCGA. Integrated genomic analyses of ovarian carcinoma. *Nature.* 2011;474(7353):609–615.
6. Kang J, D'Andrea A, Kozono D. A DNA repair pathway-focused score for prediction of outcomes in ovarian cancer treated with platinum-based chemotherapy. *J Natl Cancer Inst.* 2012;104(9):670–681.
7. Bolton KL, Chenevix-Trench G, Goh C, et al. Association between BRCA1 and BRCA2 mutations and survival in women with invasive epithelial ovarian cancer. *JAMA.* 2012;307(4):382–390.
8. Turner N, Tutt A, Ashworth A. Hallmarks of 'BRCAness' in sporadic cancers. *Nat Rev Cancer.* 2004;4(10):814–819.
9. 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.
10. Ciccia A, Elledge SJ. The DNA damage response: making it safe to play with knives. *Mol Cell.* 2010;40(2):179–204.
11. Meier D, Schindler D. Fanconi anemia core complex gene promoters harbor conserved transcription regulatory elements. *PLoS One.* 2011;6(8):e22911.
12. 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.
13. 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.
14. Willers H, Taghian AG, Luo CM, et al. Utility of DNA repair protein foci for the detection of putative BRCA1 pathway defects in breast cancer biopsies. *Mol Cancer Res.* 2009;7(8):1304–1309.
15. Mukhopadhyay A, Elattar A, Cerbinskaite A, et al. Development of a functional assay for homologous recombination status in primary cultures of epithelial ovarian tumor and correlation with sensitivity to poly(ADP-ribose) polymerase inhibitors. *Clin Cancer Res.* 2010;16(8):2344–2351.
16. Sakai W, Swisher EM, Karlan BY, et al. Secondary mutations as a mechanism of cisplatin resistance in BRCA2-mutated cancers. *Nature.* 2008;451(7182):1116–1120.
17. Edwards SL, Brough R, Lord CJ, et al. Resistance to therapy caused by intragenic deletion in BRCA2. *Nature.* 2008;451(7182):1111–1115.
18. Swisher EM, Sakai W, Karlan BY, et al. Secondary BRCA1 mutations in BRCA1-mutated ovarian carcinomas with platinum resistance. *Cancer Res.* 2008;68(8):2581–2586.
19. Sakai W, Swisher EM, Jacquemont C, et al. Functional restoration of BRCA2 protein by secondary BRCA2 mutations in BRCA2-mutated ovarian carcinoma. *Cancer Res.* 2009;69(16):6381–6386.

20. Norquist B, Wurz KA, Pennil CC, et al. Secondary somatic mutations restoring BRCA1/2 predict chemotherapy resistance in hereditary ovarian carcinomas. *J Clin Oncol.* 2011;29(22):3008–3015.
21. Dhillon KK, Swisher EM, Taniguchi T. Secondary mutations of BRCA1/2 and drug resistance. *Cancer Sci.* 2011;102(4):663–669.

Funding

This work is supported, in part, by the National Institutes of Health (R01CA125636 to EMS and TT; R01CA131965 to EMS; P50CA83636 to EMS, TT, and BYK) and Howard Hughes Medical Institute (to TT). American Cancer Society (SIOP-06-258-06-COUN to BYK).

Notes

The authors declare no conflict of interest.

Affiliations of authors: Department of Obstetrics and Gynecology, University of Washington, Seattle, WA (EMS); Howard Hughes Medical Institute, Fred Hutchinson Cancer Research Center, Seattle, WA (TT); Division of Human Biology and Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA (TT); Cedars-Sinai Medical Center, Women's Cancer Program at the Samuel Oschin Comprehensive Cancer Institute, Los Angeles, CA (BYK).