

NextGen

*Research advances
New technologies
Future practice*

Credit: Nathan L. Galli, University of Utah.



38–41

Big Data, Hidden Knowledge
Orly Alter explains how she applied a mathematical model to the problem of tumor assessment; resulting in a new tool to analyze ovarian cancer prognosis, and to predict how patients will respond to platinum therapy.

42–44

Color Power for Cancer Management
Bladder cancer is one of the most costly malignancies to diagnose and manage – could a new method of staining urothelial cells offer a much-needed alternative to cystoscopy?

Big Data, Hidden Knowledge

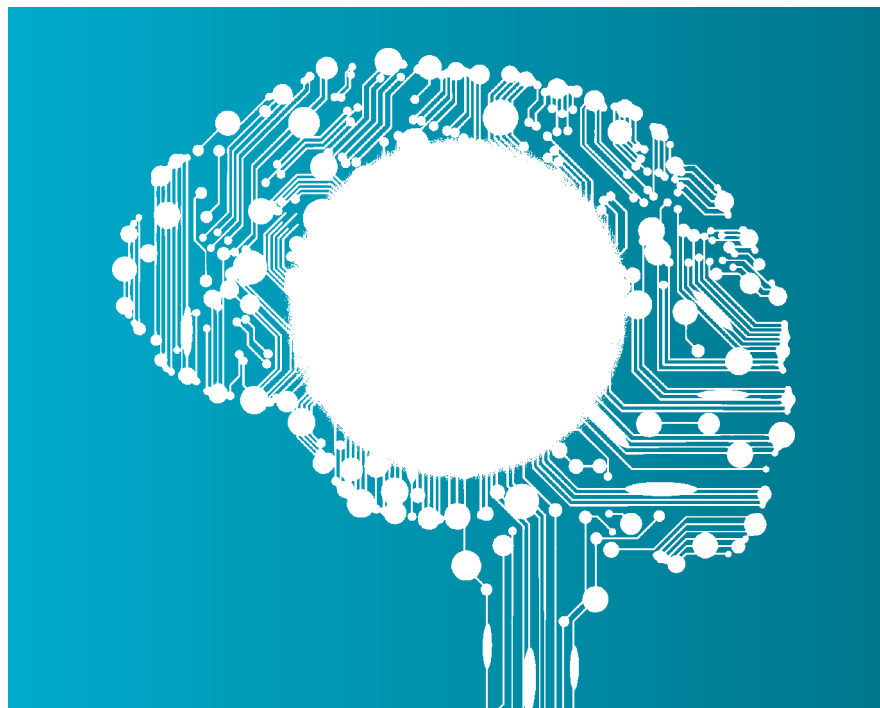
Novel mathematical tools show promise in predicting ovarian cancer patient survival and sensitivity to platinum-based chemotherapy.

By Fedra Pavlou

Ovarian cancer is an unusually harsh form of cancer. Almost 80 percent of the patients diagnosed – about 50,000 annually in the US and Europe alone – are in advanced tumor stages at diagnosis, and most are expected to die within five years. The statistics already look gloomy, and they are compounded

At a Glance

- *Most ovarian cancer cases are diagnosed in advanced stages, and no diagnostic exists that distinguishes tumors that are resistant to traditional chemotherapy from those that aren't*
- *A team based at the University of Utah have developed a mathematical tool that analyzes DNA profiles from the Cancer Genome Atlas to discover patterns of DNA anomalies*
- *Using this approach, they have been able to predict a woman's outcome significantly better than can be done with the tumor's stage; it is also the first known indicator of how well a woman will respond to platinum therapy*
- *These DNA patterns could be the basis of a personalized prognostic and diagnostic laboratory test. The researchers continue to assess the patterns in ovarian cancer and plan to expand their mathematical modeling to other tumor types*



by the frightening fact that about 25 percent of primary ovarian cancer tumors are resistant to platinum-based chemotherapy, the first-line treatment for over 30 years now. And no pathology laboratory diagnostic exists that distinguishes between resistant and sensitive tumors before treatment. Not only has treatment remained largely unchanged for three decades, but so too has the diagnosis and prognosis of ovarian cancer. Until now, the best indicator for how a woman will fare and how her cancer should be treated, has been the tumor's stage at diagnosis.

Orly Alter and her students at the University of Utah's Genomic Signal Processing Lab have been working on possible solutions to the problem of tumor assessment, by developing mathematical tools for interpreting interrelated datasets. They hope that their tools will improve understanding of cancer at the molecular level, and be the basis of personalized prognostic and

“Treatment remained largely unchanged for three decades, but so too has the diagnosis...”

diagnostic laboratory tests. One of their first disease targets: ovarian cancer.

“Our algorithms extend a mathematical technique called the singular value decomposition or SVD. The SVD helps us understand data arranged in two-dimensional tables, known as matrices, by breaking the data down into individual components. In physics, for example, the SVD describes the activity of a prism, which splits white light into its component colors,” explains Alter, who has a PhD in applied physics. “So it seems natural to

me that generalizations of the SVD can separate the multidimensional data that arise in medicine into mathematical patterns that have biological meaning.”

The team decided to try their most recent algorithm on ovarian cancer data, after they had successfully tested a previous algorithm on glioblastoma (GBM) data (1). Why ovarian cancer? “To be honest, simply because it was the next disease after GBM in the Cancer Genome Atlas [or TCGA, a US national database containing data from thousands of cancer patients]. It was only after we started our work in this area that we really appreciated why this type of ovarian cancer, ovarian serous cystadenocarcinoma, is one of the initial diseases to be studied by TCGA,” Alter says.

The value of patterns

Alter and her team develop algorithms to uncover patterns in datasets arranged in multidimensional tables, known as tensors (Figure 1). Rather than simplifying big data (a common approach), they have actually made use of the complex structure of the data to tease out the patterns within. So, for example, by modeling DNA profiles of tumor and normal cells from the same set of patients, they were able to separate the patterns of DNA anomalies – which occur only in tumor genomes – from those that occur in the genomes of normal cells in the body, and from variations caused by experimental inconsistencies.

According to Alter, their mathematical tools uncovered patterns of DNA anomalies that predict a woman’s outcome significantly better than tumor stage (Figure 2) (2). “These patterns are the first known indicator of how well a patient will respond to platinum therapy,” she says. “We found, for example, that among patients that were diagnosed at late stages, the DNA

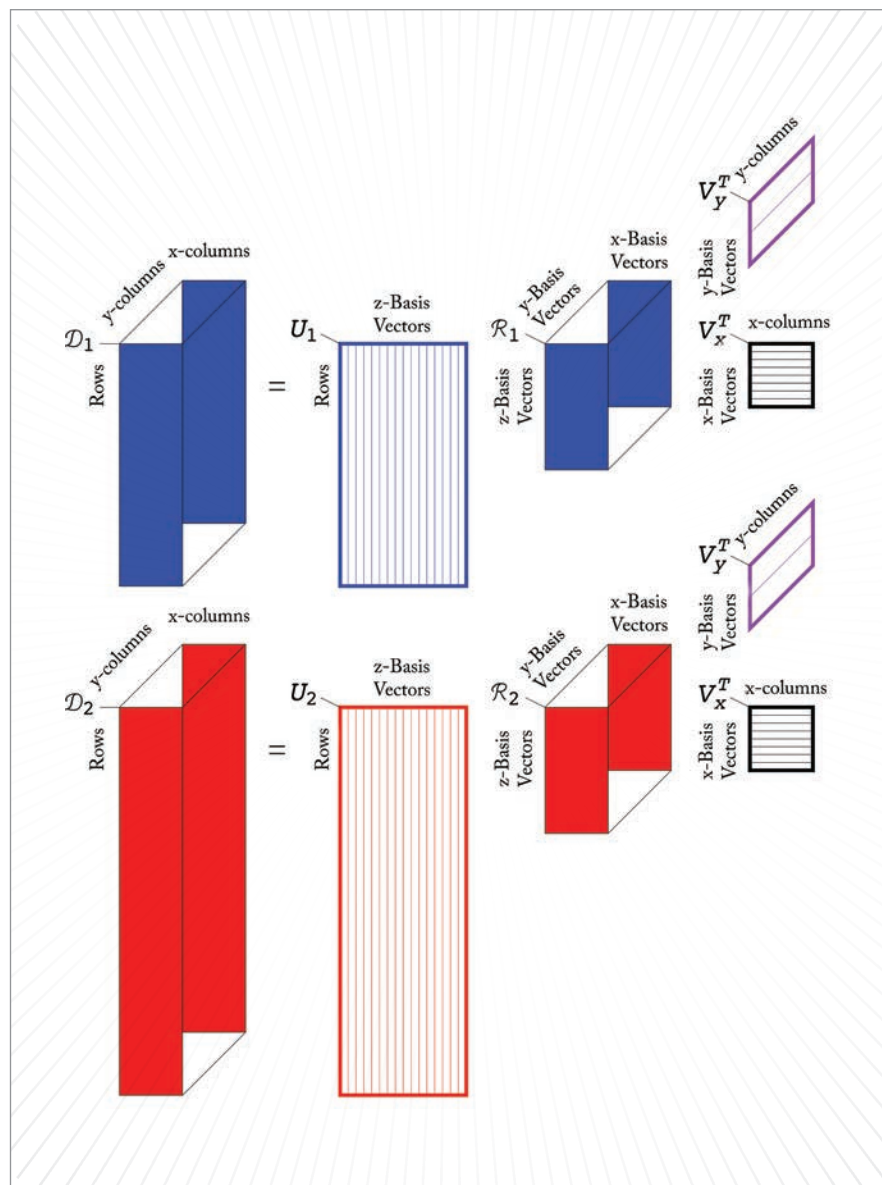


Figure 1. The tensor generalized SVD. Alter and her team develop algorithms to uncover patterns in datasets arranged in multidimensional tables, known as tensors. Rather than simplifying the big data, as is commonly done, the algorithms make use of the complexity of the data in order to tease out the patterns within them. Credit: Theodore E. Schomay and Orly Alter, University of Utah.

patterns distinguished about 60 percent short-term survivors, with a median survival time of three years, from about 10 percent long-term survivors, with a median survival time almost twice as long. Among patients treated with platinum-based chemotherapy drugs,

the DNA patterns distinguished those with platinum-resistant tumors (about 55 percent), with a median survival time of three years, from those with platinum-sensitive tumors (about 15 percent), with a median survival time of more than seven years. We then computationally

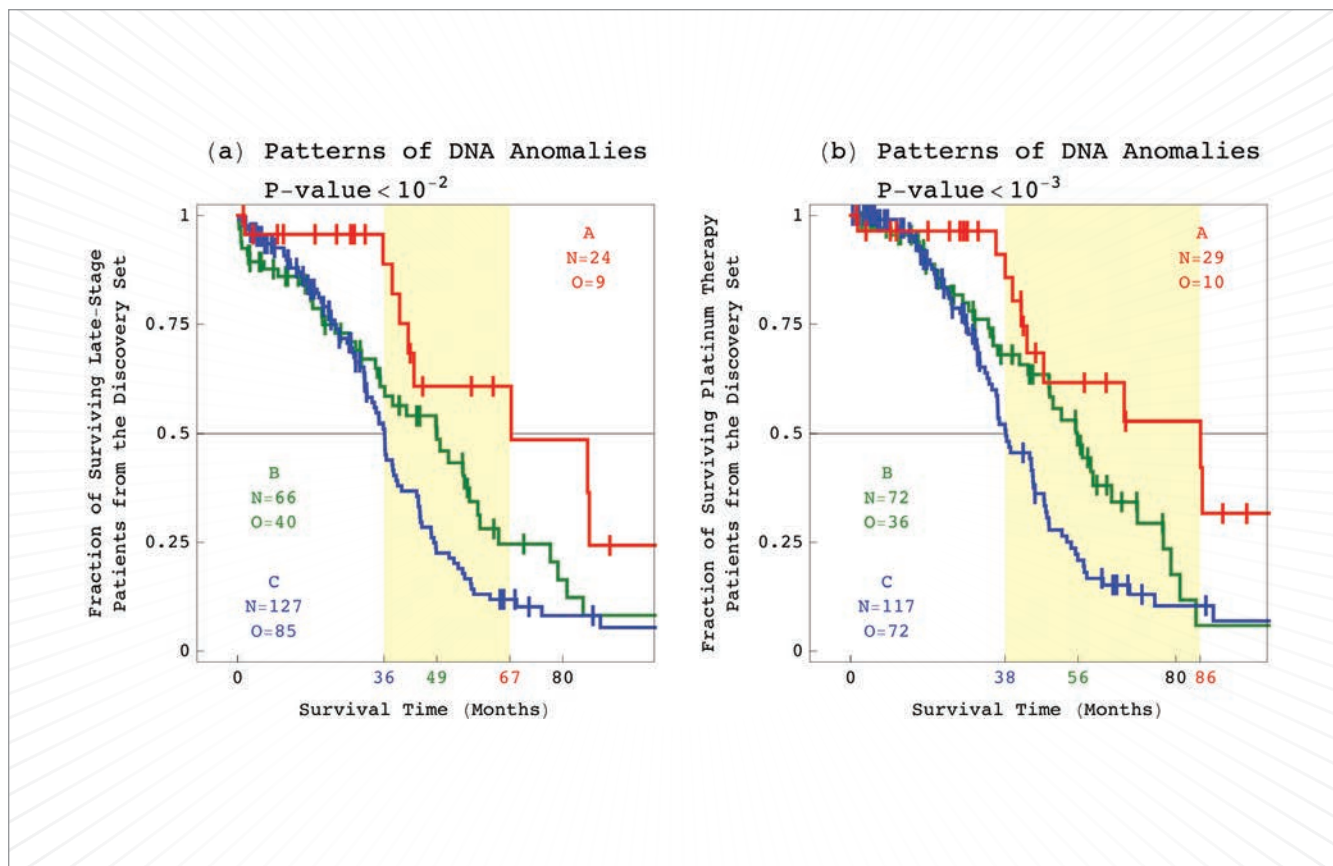


Figure 2. A personalized prognostic and diagnostic laboratory test for ovarian cancer would predict both the patient's survival and the tumor's sensitivity to platinum-based chemotherapy; doctors could tailor treatment accordingly. Credit: Katherine A. Aiello and Orly Alter, University of Utah.

validated the results by using data from independent sets of patients.

"Because these patterns link a tumor's genome with a patient's phenotype, they offer insights into the cancer's formation and growth. For example, one of the patterns points to a combination of genetic changes, the cellular function of which is similar to one that was shown by Robert Weinberg's lab at MIT to convert human normal cells to tumor cells" (3,4).

What this means for patients

How does this translate into benefit for the patient? "Based on our results so far, we believe that our DNA patterns could be the basis of a personalized prognostic

and diagnostic laboratory test, pending experimental revalidation in the clinic," Alter says. "This test would predict both the patient's survival and the tumor's sensitivity to platinum-based chemotherapy. Doctors could then tailor treatment accordingly."

The hope is that, for those with a poor prognosis, doctors can focus on taking measures to improve quality of life. For those with platinum-resistant tumors, doctors can suggest other appropriate, approved therapies. "Because no diagnostic currently exists that distinguishes between resistant and sensitive tumors before the treatment," says Alter, "these drugs can only be administered after the platinum-based

treatment fails. A pathology laboratory test based upon the DNA patterns we uncovered would, therefore, eliminate a lot of unnecessary suffering and expense."

How would this work in practice? Alter explains, "The test would analyze DNA, which can be robustly measured from formalin-fixed paraffin-embedded [or FFPE] samples. We believe this is more reliable than pathology laboratory tests that depend on measuring such easily degradable biomarkers as RNA. The test can also be used with all existing, off-the-shelf platforms for measuring DNA profiles, such as DNA microarrays and next-generation sequencing."

“This test would predict both the patient’s survival and the tumor’s sensitivity to platinum-based chemotherapy.”

The specific genes found to be perturbed could be the basis for drug therapies. Alter explains, “For example, some of the genes that we assessed during our research – the p21-encoding *CDKN1A* and the p38-encoding *MAPK14* on 6p, and *RAD51AP1* on 12p – are already known to interact with existing drugs, but were not recognized previously as targets for therapy in ovarian cancer. Pending clinical trials, these existing drugs may be found to benefit some of the patients.”

Next steps?

What direction will this research move in next? “First, we are working to translate our basic science to the clinic. To this end, we are setting up collaborations with medical doctors and pathologists to experimentally revalidate the patterns,” says Alter.

“Second, we are working to develop prognostic and diagnostic tests for other cancers. Ultimately we plan to cover most cancers studied by TCGA and similar consortia, such as the International Cancer Genome Consortium. There are currently data available for at least 14 additional cancers at TCGA. In developing these tests, we plan to make use of data from the X chromosome,” she adds.

The US National Human Genome

Research Institute noted that fewer than 1 percent of genomic associations with a disease map to the X chromosome (5). This is because the X chromosome is regularly excluded from most genomic data analyses, such as the 2011 TCGA report on ovarian cancer. The normal female genome includes two copies of the X chromosome, whereas the male genome includes just one. Excluding the X chromosome removes the normal variation between the female and male patients from the data, but may also remove variations that are due to the differences among the tumors, which may be linked to variations among the disease outcomes.

“Our algorithms not only find patterns of DNA variation, but also tell us which patterns are exclusive to the tumors, and which are common to the normal and tumor genomes,” explains Alter. “This means that we can separate the normal variation in the number of copies of the X chromosome, which is common to the normal and tumor genomes, from the tumor-exclusive patterns, which do not occur in the normal cells, while still including the X chromosome in the analyses. When the X chromosome is associated with a disease, our mathematical tools would be able to identify this association as a tumor-exclusive pattern that maps to the X chromosome. For example, analyzing the ovarian cancer data, one of the patterns of DNA anomalies we uncovered maps to the X chromosome, which is perhaps not surprising for this gynecological cancer, but would be missed by X chromosome-excluding analyses of the same data.

“Third, we are developing additional algorithms that extend the SVD to more than two datasets arranged in tensors. These algorithms will enable the comparison of more than just two cell types. For example, we could use these algorithms to model data from recurrent tumor cells together

with data from primary tumor and normal cells. These models will identify not just what patterns are exclusive to the tumor cells, but also what patterns are similar and dissimilar between the primary and recurrent tumor cells,” she says.

Alter has high hopes for these mathematical tools: “It may very well be that the data needed to better treat cancer are already published. The ovarian cancer data, for example, were published back in 2011. The bottleneck to discovery is in the analysis of the data, and we hope to have found a way to overcome the bottleneck and provide the means for improved prognostics, diagnostics and disease management.”

Orly Alter is associate professor of bioengineering, adjunct associate professor of human genetics, and faculty member of the Scientific Computing and Imaging Institute, University of Utah, US.

References

1. CH Lee, et al., “GSVD comparison of patient-matched normal and tumor aCGH profiles reveals global copy-number alterations predicting glioblastoma multiforme survival,” *PLOS One*, 7, e30098 (2012). PMID: 22291905.
2. P Sankaranarayanan, et al., “Tensor GSVD of patient- and platform-matched tumor and normal DNA copy-number profiles uncovers chromosome arm-wide patterns of tumor-exclusive platform-consistent alterations encoding for cell transformation and predicting ovarian cancer survival,” *PLOS One*, 10, e0121396 (2015). PMID: 25875127.
3. WC Hahn, et al., “Creation of human tumour cells with defined genetic elements,” *Nature*, 400, 464–468 (1999). PMID: 10440377.
4. AE Karnoub, Weinberg RA, “Ras oncogenes: split personalities,” *Nat Rev Mol Cell Biol*, 9, 517–531 (2008). PMID: 18568040.
5. Notice of the National Human Genome Research Institute’s Interest in Receiving Applications to Analyze and Develop Methods for X Chromosome Genome-wide Association (GWA) Data; <http://1.usa.gov/1AqZHE9>.