The computer will see you now

'Computational pathologist' diagnoses different grades of breast cancer.

Heidi Ledford

09 November 2011

A computer program has been trained to grade breast cancer, predicting which tumours are associated with worse outcomes and, therefore, deserve more aggressive treatment.

In doing so, the program aims to improve on a technique that has remained essentially unchanged for more than 80 years. The method for grading breast tumours was established in 1928, and is largely based on three criteria: how well the tumour cells resemble healthy tissue, how many cells have abnormal nuclei, and how many cells in the sample are dividing.

But after being fed a 'training set' of breast tumour images and survival data, the computer program — named C-Path, for computational pathologist — developed a new list of features that best predicted patient outcome. Instead of focusing on the tumour cells themselves, C-Path determined that the most predictive features were found in the cells surrounding the tumour, in a region called the stroma. The results were published today in *Science Translational Medicine*[1].

"It tells us that cancer really is an ecosystem," says Daphne Koller, a computer scientist at Stanford University in California, and lead author on the paper. "Yet currently, pathologists don't look at the stroma at all."

Science/AAAS

The program C-Path scans stained tumour images (bottom left and top right) and labels them (top left and bottom right) to predict which tumours need more aggressive treatment.

Quantity and quality

To walk into a pathology lab can feel like a step back in time, with pathologists hunched over microscopes analysing slides of tissue stained with dyes that have been used for decades. But the field is changing, says David Rimm, a pathologist at Yale University School of Medicine in New Haven, Connecticut. "It's not just about the classical stains anymore," he says. "It's about the molecular events."

That transition enticed Andrew Beck, a trained pathologist and an author on the C-Path paper, to seek out graduate training in bioinformatics. He knew that pathology labs process thousands of tissue slides every year — a "treasure trove of data", he says — but that new techniques were needed if those morphological findings were ever to be correlated with the emerging molecular data that increasingly influence treatment decisions.

C-Path is not the first attempt to quantify and automate the traditional art of pathology, says Beck, now at Harvard Medical School in Boston, Massachusetts. "But it's difficult to do," says Beck. "The images are very complex and heterogeneous — not an easy type of information to convert into quantitative data."

Beck and his colleagues decided to take a different approach. Rather than training a computer to recognize features that pathologists had already deemed important, they decided to allow the machine to pick them out. C-Path's training set consisted of images of tumour tissue and five-year survival data from 248 women with breast cancer. The program looked for 6,642 features in that set, chose a subset of features that reliably predicted survival, and then applied those to a 'validation' set of data from 328 women.

The computationally selected features were strongly associated with overall survival in both sets of data.

An objective approach

The technique is a significant step for a rapidly modernizing field, says Rimm. But it will need to be tested in larger studies before it can be deployed clinically.

Rimm also notes that, during C-Path's training, the team had to include 42 samples from the validation set to account for differences in the staining technique used in the two data sets. This variability is common, he says: "Some labs like a little more red, some like a little
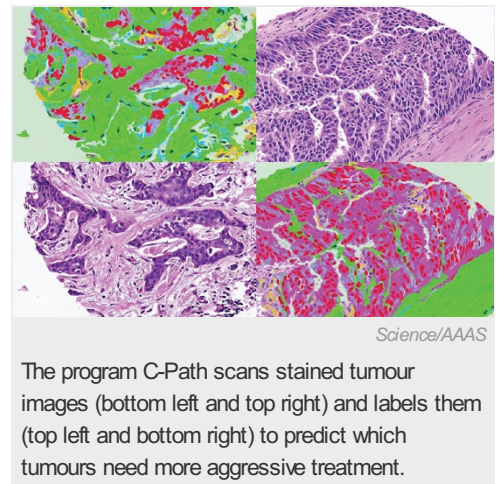
more blue."

That could complicate efforts to make a program that can be widely applied in clinical settings, acknowledges Beck. It is possible that labs would have to adhere to a specific staining protocol if C-Path is involved in data analysis.

Meanwhile, Rimm is hopeful that the C-Path method can be applied to other cancers, such as prostate and bladder cancer, for which grading tumours reliably has proved difficult for even well-trained human eyes. "In those cases, if you show the sample to a different pathologist, you often get a different grade," he says. "And if you're going to treat on the basis of grade, then subjectivity is a problem."

References

1. Beck, A. H. *et al. Science Trans. Med.* **3**, 108ra113 (2011).