YALE JOURNAL OF BIOLOGY AND MEDICINE 80 (2007), pp.183-185.
Copyright © 2007.

INTERVIEW

Dr. David Rimm

Feras Akbik

Yale University School of Medicine, New Haven, Connecticut

Dr. David Rimm, MD PhD, is a professor of Pathology at the Yale University School of Medicine specializing in developing quantitative, diagnostic techniques. His lab recently engineered a fluorescence-based algorithm, Automated Quantitative Analysis (AQUA), to analyze tissue microarrays in the hope of moving toward personalized medicine and diagnoses.

As a physician-scientist, you're a practicing pathologist and an investigator interested in developing diagnostic techniques. What is it about diagnostics, as opposed to pursuing a more therapeutic approach, that drew you to pathology and your research field?

Diagnostics is really where pathologists live. We realize the importance of proper diagnosis and the concept that even though two people have the same disease, it may manifest itself very differently in each individual. Our goal as pathologists is to look at the tissue and more carefully subclassify the diseases. When I first encountered pathology in medical school, I was amazed at how subjective this process was. Pathology seemed more of an art form than an actual scientific discipline. That was when I became interested in the use of molecular tools to improve the diagnosis of disease. I wanted to develop a more objective way to treat patients appropriately as a function of their disease sub-class.

In what ways did your experiences as a pathologist help you think of developing something like AQUA?

As a pathologist, I've seen an urgent need to move beyond subjective methods of diagnosis and treatment. The best example is Herceptin, a drug immensely valuable for the treatment of breast cancer. Fifteen percent of breast cancers are highly responsive to Herceptin, but the drug costs $100,000 and somewhere between 5 and 10 percent of the patients will have significant cardiac side effects. So with an outcome that important, don't you think we ought to have a really good test that we don't guess at? You need to have a quantitative test that's reproducible and get a number and make a medical decision based on a reproducible number.

Back in the late 1990s, molecular diagnostics was just coming of age; there were only a handful of tests you could do. We relied heavily on immunohistochemistry, but because it was so fraught with variables and poorly controlled, it was not respected as a molecular test. "Molecular" had come to mean a DNA-based test, and we were making great progress in deciphering the genetic markers for various subclasses of disease. The use of fluorescence instead of antibody stains was the first step in quantifying these markers.

The true turning point was the invention of the tissue microarray in 1998. This technique was revolutionary because it reduced a specimen to a 0.6mm diameter piece of tissue on a slide. We regarded that one slide as representative of the whole tumor, even though it comprised 0.03 percent of the whole tumor. With such a small sample, we thought it possible to be rigorously quantitative and count every single molecule in that little spot. The goal was to do away with the traditional, arbitrary scoring system of +2, +3, etc. and calculate a specific concentration, such as 2 µg, for a given molecular marker. That was the vision, but we could not find a piece of software to do it. That was when Dr. Richard Camp, a senior associate research scientist, wrote the first version of AQUA in visual basic.

What makes AQUA so different from the software that came before it?

Today, most software utilizes what is known as feature extraction. That's when you have a contrast generating edge, and the best example of this is the nucleus of a cell. You look through the cell and anytime you find something round, it's a nucleus. AQUA, on the other hand, makes use of fluorescence to define a location within the cell. We want everything to be defined by a molecular definition rather than by contrasting shapes. In the case of the nucleus, we take advantage of the fluorescent dye, DAPI, which binds to DNA. By quantifying DAPI-positive pixels, AQUA gives the nucleus a molecular definition without using feature extraction.

So how do we arrive at a concentration? Well, a measurement is simply a numerator over a denominator. We have our denominator, the compartment, and now we need our numerator, which is just the sum of all the intensities of our target molecule. For example, we can take the sum of all the intensities of estrogen receptors signals and divide that by the sum of the areas, and we have a virtual concentration.

In what settings have you been able to test AQUA against pathologists or have validated the system in a clinical model?

Since the first publication in 2002, AQUA has been used in many different settings. Many times, it allows us to see things that the pathologist can't see. One reason is that pathologists, and human beings in general, are able to tell dark from light, but they are generally bad at reproducibly judging shades of gray. AQUA tends to do better than the pathologist at measuring subtle differences. One of the first studies looked at Her-2. A pathologist can see the high level peak, probably only missing 10 to 15 percent of those. But what happens with low-level expression? AQUA is especially useful when you have to tell the difference between low and very low or high and very high. Another advantage of AQUA stems from the fact that human beings can't calculate ratios in their heads. AQUA can measure the signals from the nucleus and the cytoplasm and then ratio the two. By internally normalizing the scores, we can get rid of individual variation in protein expression levels. Finally, the most important feature of AQUA is that we can multiplex. It's clear that if you want to classify and do diagnostic things, one protein is not going to be enough. For example, we now have an algorithm of six different proteins to predict a patient's response to Herceptin.

So what you're saying is that using AQUA in a multiplex setting is going to lead to a more specific or personalized diagnosis?

Absolutely. I think AQUA will lead us to a more personalized diagnosis of tumor subclass. For example, we can use the expression profile of multiple markers to identify those patients with the type of tumor that is more likely to recur. We can then give those patients extra treatment after their resection as a protective measure. The first DNA-based test, Oncotype DX, is being used now. It analyzes the sequence of 21 genes to fractionate patients with the same disease into sub-classes. We can tell which patients need more therapy. AQUA will be different in that it analyzes protein expression rather than a DNA or RNA sequence. Instead of looking at 21 genes like

Oncotype DX, we now can analyze the expression of only five proteins. Looking at the protein level is more accurate approach.

So why do you use a proteomic approach instead of a genomic one?

Well, for two reasons. Genomics is the grammar, but proteins are the literature. We need all the letters to make the words, but the functional things are the responsibilities of the proteins. Even if you had RNA expression, it doesn't mean there's protein expression, or, if it was, it might not still be around while the RNA is still there. Also, assessing the activity of a protein by looking at post-translational modifications is a lot more information-rich than simply knowing if the protein is there or not. Using AQUA, we can measure phosphorylation states and see it change as we give patients a particular drug. We certainly can't do that by eye. We really need to have a number and see if that number goes down.

What do you think it's going to take to make AQUA, and technology like it, a more standard fixture in clinical practice and research? How do you think it fits into the future of medicine?

That's the real challenge, because if you can do something in your lab that's really cool but nobody else in the world is doing it, it's not that valuable. The next step is for AQUA to get out there and be employed in clinical trials. We're involved now in a trial aimed at predicting the response to Herceptin. If we can demonstrate that using AQUA is better at predicting a response to Herceptin than by looking at Her-2 expression alone, that might make other people want to try it or buy it. Oncologists may want to do the AQUA test instead of the traditional immunohistochemistry test to decide whether to prescribe Herceptin.

What's the limitation for this technology, or what kinds of things could AQUA not do?

AQUA requires a tissue sample, which means performing a biopsy. This is a bit more invasive than imaging-based diagnoses, like MRI, PET, etc. Another disadvantage is that you can't just look under the microscope and get the answer right away. AQUA takes time and standardization. It may be more expensive, but it costs nothing compared to the drug at $100,000 per year.

Can you tell us where AQUA goes from here?

In 2004, HistoRx was founded, and they are the exclusive licensee of the software. Sometime later this year, they will release the first test so pathologists can use it in the clinical setting to predict responses to therapies in more subtle ways than we can currently do by eye. Not too long after that, there may be tests that can predict whether patients do or don't need chemotherapy or if a tumor will recur or not. The commercialization of AQUA will allow it to reach a greater number of patients.

How long will it take to get this technology to everyday use? What needs to happen?

I initially guessed it was going to be a few years, but it's clearly going to take a lot longer than that. It turns out that a diagnostic test has to go through the same trials as drugs have to go through to be accepted. A drug can take 10 to 12 years, and we might see the same thing with diagnostics. We're pretty far along the pathway, so we might see the first inkling of it this year, but my hope is we'll see AQUA as a prominent fixture in pathology in three to five more years.