Debate Rages Over Proteomic Patterns

One controversy aired in numerous sessions and hallways at this year’s American Association for Cancer Research annual meeting was the future of proteomic patterns in cancer diagnosis. Can protein “signatures,” generated by mass spectrometers, accurately and reliably distinguish cancer from noncancer?

This idea has generated a tremendous amount of excitement among researchers—as well as wide national publicity—because it could, in theory, lead to a rapid, universal blood test for cancer. But critics have recently begun to raise their voices against the pattern concept. “The validity of the method is questionable at the moment,” said Eleftherios Diamandis, M.D., Ph.D., a cancer biomarkers expert at the University of Toronto.

Diamandis, who has penned several critiques of the protein pattern approach (for example, see JNCI, Vol. 96, No. 5, p. 353), was speaking at a packed AACR forum. Emanuel “Chip” Petricoin, Ph.D., the Food and Drug Administration scientist who helped pioneer the profiling method, also spoke. Petricoin co-authored the landmark paper that appeared in the British medical journal The Lancet in February 2002, which showed that mass spectrometry, coupled with an artificial intelligence algorithm, could distinguish ovarian cancer from normal controls with 100% sensitivity (in other words, no false negatives) and 95% specificity (5% false positives).

The paper created a sensation, even leading to a Congressional resolution supporting the method, but Petricoin is quick to acknowledge its limitations. “I view this as kind of like, ‘Can we do heavier than air flight?’” he said. “Technology marches on. And what we are doing now is not taking the Wright Flyer across the Atlantic. A one-minute flight was enough to say, ‘This is a possibility, this could work.’ Now we need to do the hard work of making sure this can happen over and over again, and [start] building trans-Atlantic aircraft.”

But Diamandis and Edward Patz, M.D., a radiologist and lung cancer specialist at Duke University Medical Center, argued that the usefulness of mass spectrometry patterns for diagnosis is still completely unproven. Diamandis contended that mass spectrometry alone is incapable of distinguishing cancer in blood from noncancer. “This hypothesis is actually wrong,” he said, “for the simple reason that mass spectrometry does not have the adequate sensitivity to pick up the changes.”

All agree that mass spectrometry is a powerful tool for discovering novel biomarkers (see sidebar, p. 817). The argument is over mass spectrometry as a clinical tool. Can the peaks themselves, without knowing what proteins they represent, be diagnostic? “We want to leave that possibility open,” said O. John Semmes, Ph.D., director of the Center for Biomedical Proteomics at Eastern Virginia Medical School.

Petricoin agrees. “The pattern itself can be used as a diagnostic independent of the [protein] identities,” he contended in an interview. Mass spectrometry, he said, is “better and faster, with higher accuracy, than any other test.” But Patz is skeptical. “Without these proteins, with simply using mass spec patterns, [we] reside in the black box.”

The Hamburger Effect

One objection to patterns is that they are hard to reproduce. “There are issues of reproducibility, how you handle specimens, how you put them down on the chips,” said Patz. “There are serious concerns.” Because most mass spectrometers were not designed as clinical tools, it is hard to generate the same results from machine to machine or even from operator to operator. Petricoin admitted that the machines used in the study published in The Lancet did not measure up, but he expects his new machines to perform better. And he contends that rigorously enforced sample handling procedures should eliminate those reproducibility problems.

Reproducibility of patterns is another issue. Diamandis, for example, pointed out that patterns published in five different prostate cancer studies were completely different. “No group has ever discovered the same discriminatory peak in any of the five studies,” he reported. “This is unbelievable. The same occurs for ovarian cancer as well.”

That hardly makes the concept invalid, said Petricoin. “Everyone’s using a different methodology,” he pointed out. “Thus it’s self-evident to us why you would get entirely different features.” As for why the same group cannot come up with the same pattern every time, Petricoin said that there are many patterns from the same sample that will identify the cancer. “There’s a fish tank with a million fish, and you’re dipping your net in different regions of the fish tank,” he said. “And you’re asking, ‘how come you don’t catch the same fish?’”

But even if the diagnostic pattern is perfect, could fluctuations in the patient’s protein pattern upset the diagnosis? “Somebody eats a hamburger, a Big Mac, it may change your profile dramatically, because if you add one protein to a complex mixture, you may suppress other proteins,” said Patz.

Diamandis agreed. “If I eat a hamburger, will my proteomic pattern change? Nobody knows,” he said. “If I take a sample from a fridge that has been freeze-thawed many times, will that change the proteomic pattern?"
Nobody knows.” The validation trials should put these concerns to rest, countered Semmes, because the final pattern algorithm will be common to patients whether or not they have eaten hamburgers (or anything else) and will be tested rigorously. “Every appropriately designed validation test is certainly going to take simplistic factors like that into consideration,” he said. “We’ll make sure folks are coming in that have eaten whatever they want.”

**Finding Needles in Haystacks**

Reproducibility, in theory, can be achieved with the right procedures and the right equipment. A more fundamental objection to patterns is the possibility that mass spectrometry cannot detect the necessary discriminatory proteins to begin with. That is because cancer markers may only be present in the blood in very low amounts, which mass spectrometry has trouble picking up. “This technology … detects high-abundance molecules and is missing the low abundance molecules, and—in my opinion—most of the information,” said Diamandis. He added that none of the studies to date have detected known biomarkers, like prostate-specific antigen (PSA) and CA-125, in their patterns. “They discovered other things, but not the qualified cancer biomarkers, and I think this is a major limitation.”

The failure to find known biomarkers is a legitimate concern, said Semmes. “Are we going to be able to look deep enough to find the right things?” he asked. “The only way to find out is, if our validation fails, then we know we didn’t find them. You can’t guess ahead of time.”

Semes’ group is validating a prostate cancer algorithm at several sites around the country. Results for this trial should be available within a year, he said, and if the algorithm works then a definitive validation would be launched. “Are we going to find biomarkers that are going to be diagnostic?” he asked. “There’s no reason to think that we won’t.”

As for why he is not finding PSA, “there’s no easy answer,” said Semmes. He conceded that Diamandis may be right: Mass spectrometers have trouble picking up low abundance proteins such as PSA. But that does not mean that they are not going to find other proteins that, together, can reliably diagnose early cancer. The ongoing clinical trials should settle the question. “The only way to do it,” Semmes said, “is to do it.”

**Noise Pollution**

The third objection to patterns is that they are mostly “noise,” not discriminating biological information. That is because of the technology. In mass spectrometry, the serum sample is spotted on a plate or chip and fixed to a matrix. Then the proteins are ionized with a laser, and the ions fly through the

---

**Proteomics: From Signature to Protein Identification**

A proteomics technique for diagnosing early-stage ovarian cancer developed 2 years ago by researchers at the U.S. Food and Drug Administration and the National Cancer Institute relied solely on the pattern generated by mass spectrometry analysis to distinguish between serum samples from healthy women and those with ovarian cancer.

In early studies, the technique enabled researchers to diagnose patients without having to know what proteins differ between the sample types. Now, however, the team reported in March at the annual meeting of the American Association for Cancer Research in Orlando that they have modified their approach and have started identifying the proteins that underlie the mass spectrometry pattern.

As blood flows through capillaries, tissues, and vessels, it picks up small proteins and protein fragments. Most of these nonblood peptides should be cleared by the kidney within a few minutes. However, based on mathematical modeling, Arpita Mehta, a student at Tufts University School of Medicine and a Howard Hughes Medical Institute research fellow in Boston, deduced that most of the peptides stick to large carrier proteins, like albumin, rather than floating free in the serum.

Therefore, to learn what proteins account for the peaks in the mass spectrometry analysis, Mehta isolated the albumin serum fraction from 127 unaffected women and 115 women with ovarian cancer, including 31 with stage I disease. She used these samples to identify a diagnostic proteomic pattern, much like the one NCI’s Lance Liotta, M.D., Ph.D., and FDA’s Emanuel Petricoin, Ph.D., had previously identified in whole serum. By focusing on the albumin carrier and its binders, the scientists improved their ability to detect stage I ovarian cancer. But rather than stopping with that diagnostic pattern, as the group has previously done, Mehta sequenced candidate biomarkers from discriminating ion peaks in the mass spectrometry sample. Each peak consists of numerous proteins, providing her with a substantial amount of material.

With the new method, Mehta, who did this research while working with the NCI–FDA Clinical Proteomics Program, and colleagues have identified more than 800 proteins that are bound to the albumin carrier protein in the blood. Nearly three-quarters of these have not been found in blood before. They fall into numerous functional categories, including transcription factors, motor proteins, cell adhesion molecules, and proteins involved in apoptosis. Half of the proteins are of unknown function. The scientists have thus far sequenced 223 of the proteins that were differentially expressed in patients with stage I ovarian cancer relative to both stage III disease and high-risk controls.

When asked why the team has moved from pattern-based diagnostics to identifying proteins, Liotta said that it was simply the natural progression of the project. “This is like CA125; it was just a band on a gel” when people started using it as a marker for ovarian cancer. “In a year or so, we will be looking at labeled, known peaks in the pattern rather than just unknown ion peaks,” he said.

—Rabiya S. Tuma
mass spectrometer, which records their time of flight. The resulting pattern of peaks reflects their relative masses and charges. But the results are influenced by “ion suppression,” the tendency of some ions in the protein-rich gas hovering above the plate to keep other proteins on the plate from being ionized. “When you put thousands of proteins down there, on any one [readout] you’re probably seeing only 100 to 150 proteins,” said Patz. “Nobody really understands why.”

In Semmes’ view, this objection is irrelevant. Although a lot of information is lost off the chip, he said, the ion suppression is consistent, not sporadic. “Yes, it’s not accurately representing what’s in the serum, but it’s reproducibly representing what’s in serum,” he said. “If you have an ion that is suppressed and won’t come off, it’s not going to come off anytime.” So as long as the machines are properly calibrated and the procedures standardized, he added, the same peaks should be generated each time. When Patz implied that the process is undependable, Semmes said, “he’s absolutely wrong.”

Patz also pointed out that the mass spectrometer generates many peaks that do not represent proteins at all but are artifacts or “noise.” For example, events like sample contamination or even the humming of the detector plate could generate peaks. “Trying to decide what is actually a [protein] peak and what’s noise can be very difficult,” said Patz. True, said Petricoin, but the answer is to test the pattern in a blinded fashion. If the pattern distinguishes cancer from noncancer, it cannot be relying on noise. “Noise, by definition, can’t discriminate, because of its unpredictable and random nature,” said Petricoin. A systematic bias, he said, could be responsible, but bias can be eliminated by standardizing procedures and then validating. “The only way to completely eliminate it is just larger, expanded, blinded validation sets, and then the pretenders from the contenders are separated, period,” said Petricoin.

Even after the noise is weeded out, the pattern may still contain biological information irrelevant to the tumor, Diamandis pointed out. Enzymes in the blood, for example, may chop up proteins and generate irrelevant peaks. While this is true, said Semmes, properly designed “training” trials will remove those peaks and generate a biologically relevant and cancer-specific diagnostic pattern. It is important to test cancer against benign and inflammatory conditions, not just normal controls, to make sure that the final pattern reflects only cancer. Also crucial is to test the pattern with samples taken well before cancer diagnosis, if early detection is the goal.

The patterns have yet to prove themselves in blinded, multicenter clinical trials. Answers should come fairly soon, though. The ovarian cancer pattern is now being compared against CA-125 in a prospective clinical trial for the ability to detect recurrence in women who have undergone surgery and are in remission. Results could be available within 2 years. Meanwhile, Semmes’ prostate cancer pattern is in the process of being optimized and could enter a final validation, using blinded serum samples, sometime next year, with the trial lasting perhaps another year.

Will patterns prove able to discriminate between cancer and noncancer? Or should researchers instead put their energy into identifying the peaks and generating marker panels of known proteins and then testing those? The clinical trials, assuming they pass scientific scrutiny, should settle the debate. Abandoning patterns now, said Petricoin, would be irresponsible. “It’s … wrong to basically say we shouldn’t investigate anything that seems to be able to discriminate disease from nondisease,” he said. “We have an obligation to at least track that down.”

—Ken Garber