THE EVERYTHING TEST

↳ MOST DISEASES GIVE OFF A MOLECULAR SIGNAL BEFORE THEY SHOW SYMPTOMS. CAN THE SCIENCE OF PROTEOMICS LEARN TO SPOT THEM?

↳ BY MICHAEL BEHAR

↳ PG.00026

THE SERGEANT WITH THE MOUNT CRESTED BUTTE POLICE DEPARTMENT IN COLORADO APPEARED SHORTLY AFTER 9 P.M. IT WAS AUG. 1, 2017,

and I was with my wife and our two young children, ages 2 and 7, at Lake Irwin, a remote campsite at 10,200 feet in the Rocky Mountains. When the officer stepped out of his S.U.V. cruiser, its blue and red emergency strobes piercing the darkness, I thought that perhaps a neighboring camper had summoned him to silence my dissonant guitar strumming beside the campfire.

"I'm looking for Mr. Behar," the sergeant announced. My cousin, who knew our whereabouts, had called the county sheriff, who dispatched the sergeant. His name was Brad Phelps, and he had navigated a dirt road at night through rugged alpine terrain to our location, because there was no cell reception where we were. After I identified myself, Phelps read from a palm-size paper notepad: "I'm sorry to have to tell you that your mother has passed away."

"Impossible," I remember thinking. My mother, at 73, was a competitive tennis player who often went up against (and defeated) women years younger. She was exceptionally fit, never smoked and ate like an ascetic thanks to her father, who ran a produce business and raised his daughter on fruits and vegetables. "She was the picture of health," I would hear from friends at the funeral.

"I was absolutely floored," her physician, Milah Frownfelter, told me recently. "She had this tremendous vigor that was completely divorced from the acute process she suffered." In 2014, my mother had a coronary calcium scan, which uses X-rays to look for plaque in arteries. The results were plugged into a risk-crunching algorithm that compares a patient's individual health data (blood pressure, cholesterol, age and so on) with that of people who are similar, demographically and pathologically. While my mother had some minor plaque, the algorithm calculated her odds of having an "event" — a heart attack or stroke — at 7 percent over 10 years, hardly alarming for her age. She didn't even meet the threshold at which most doctors would prescribe cholesterol-lowering statin drugs, Frownfelter said.

On the day she died, she had only mild nausea and clamminess. "I think I'm going to throw up," she told my father before excusing herself to the bathroom. When he checked on her five minutes later, she was in the fetal position beside the toilet, dead from cardiac arrest. Cardiologists I consulted afterward surmised that she suffered a so-called widow-maker, a plaque blockage in the left anterior descending artery. It's an invariably fatal heart attack, one that disproportionately affects men and can kill an otherwise healthy person almost instantly and with no previous symptoms.

As an older father with young kids, I'm vigilant about my health, which led me to get a coronary scan in 2016, the same procedure my mother underwent. I was 47 and exercised fanatically. My physician thought I was crazy and refused to approve the procedure for insurance coverage. So I paid $270 to a medical-imaging outfit in Boulder, where I live, to get it done. I had smidgens of plaque in two arteries, including the left anterior descending. My doctor was surprised, but my risk was considered negligible — a 3.5 percent chance of an event within five years, based on a similar algorithm, developed by the American College of Cardiology, to the one that had established my mother's odds. "You are already doing everything right," he assured me. And yet my mom had been, too.

Six weeks after my mother's death, I visited Nelson Trujillo, a prominent cardiologist in Boulder. Coronary scans show only hard plaque, he explained; they don't reveal whether it's softening and ready to rupture, "like a big pimple," in his words. If it pops, its fragments could clog an artery, which would be dangerous, if not lethal. A coronary angiogram, done with a CT scanner, can detect soft plaque. But because a CT scan is expensive and exposes the patient to radiation, it's rarely performed on someone whose probability of having a heart attack is statistically low.

Trujillo had a recommendation. A privately held biotech company in town called SomaLogic was conducting a trial for a new blood-screening process to gauge cardiovascular health, and Trujillo happened to be leading the study. "You are the perfect candidate!" he declared with kid-like enthusiasm. Cardiologists make clinical decisions based on statistical factors gleaned from broad population studies. "We're always comparing you to big groups of people," Trujillo told me. "The problem is, we know that sometimes we're wrong. We have a miss rate — those six guys out of a hundred who we say are O.K. but who are not. I wanted to know who those six people were."

The test promised to search my blood for nine proteins associated with cardiovascular health. There are 20,000 or so known proteins in the human proteome, as the collective sum of proteins in any organism is called. Because it can signal when something is amiss inside a body, the proteome has the potential to serve as a diagnostic system — sort of like the ones in modern cars that alert mechanics when a fuel injector is plugged or a timing belt needs replacing. The SomaLogic screen wouldn't merely compare my health stats with those of other men like me. It would take a snapshot of what was happening inside my body at that moment. "It's not odds-based on people who look like you," Trujillo told me. "It's odds-based on you specifically."

Proteomics, or the study of proteins, has long offered the ability to identify many biological processes. But until recently, the sheer number of proteins and the complexity of their interactions made screenings impractical, if not impossible. Now, with the advent of more powerful computers and a form of artificial intelligence called machine learning, medical experts are imagining a future where proteomics will realize its power to tell us, to an incredible degree, what's transpiring inside our bodies. As Omid Farokhzad, a professor and physician recently at Harvard Medical School, puts it: "We'll be able to diagnose diseases such as cancers and Alzheimer's years before symptoms."

For a physician like Trujillo, who is still working with SomaLogic to give the protein test to patients, the promise of proteomics is already here. "I used to say to somebody like your mom, 'You don't have much heart disease; you don't need anything.' But that was B.S. I couldn't reassure her one way or the other," he said. "This is where proteomics comes in — and where it's fundamentally different than anything else we have." Proteomics might have saved her life — and it may yet save mine.

GENES TEND TO get more attention, but proteins might really deserve the limelight. As the workhorses in the human body, proteins play a role in nearly all of its biological processes. They make antibodies to battle infections. They grow bone and muscle and convert what you eat into nutrients. Proteins are constructed inside cells, from building blocks called amino acids. Humans have 20 commonly occurring amino acids: We produce 11 naturally and obtain nine more from what we eat (which is why high-protein foods are essential to our diet). The blueprints for proteins are stored inside genes, which contain DNA and RNA. RNA is like a courier, delivering the instructions from genes to cells for making proteins from specified assortments of amino acids.

Our proteins are in constant flux. They can appear and disappear, or shift in concentration, in response to our external environment or internal physiology. Our proteome reacts to what we eat, when we sleep, how we exercise, the smoke we inhale, the alcohol we drink. My proteins look different after I run a 10K than they do before the race; they change when I get the flu; they are altered by stress and emotions.

Researchers are now learning that diseases have their own unique proteomic patterns. An ailment like colon cancer might involve interactions with hundreds of proteins.

OPENING PAGES: CONCEPT BY DELCAN & COMPANY. PHOTO ILLUSTRATION BY JAMIE CHUNG. PROP STYLING BY PINK SPARROW.

Stephen Williams, the chief medical officer of SomaLogic, told me about a recent experiment that turned up at least 1,000 proteins associated with diabetes.

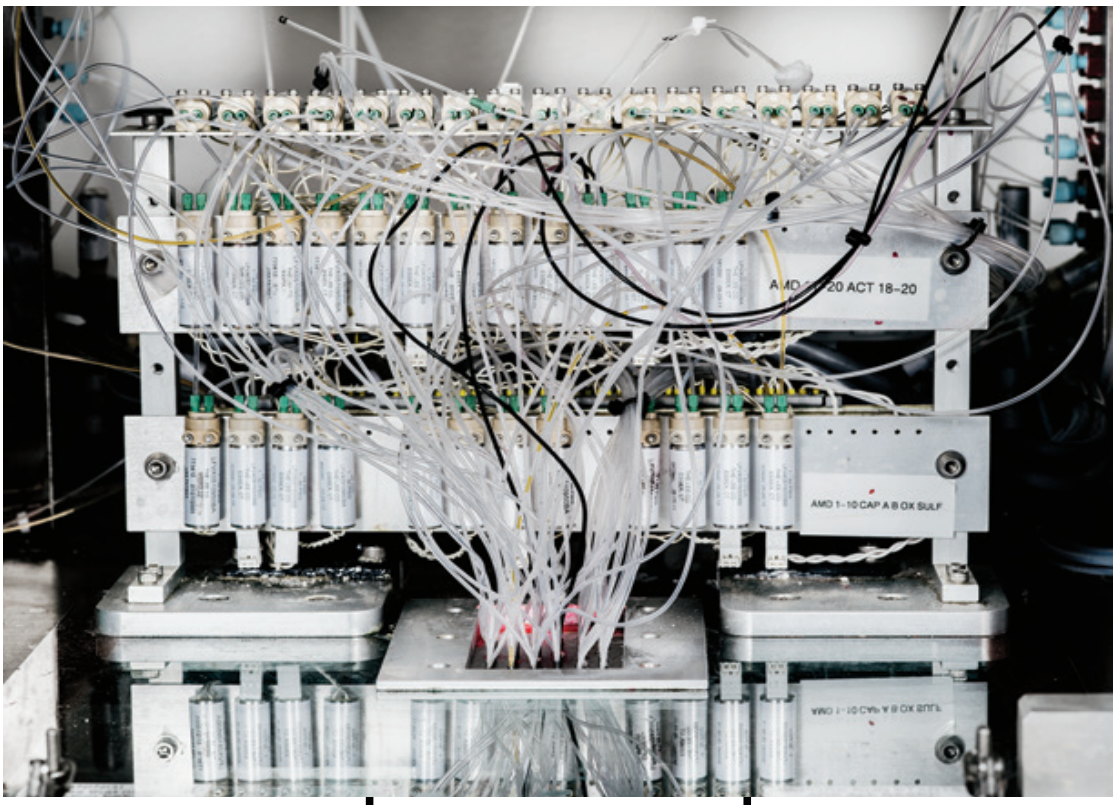
Proteins and genes have very different things to say about our health. Genes are static. We live and die with the same set we're born with: about 20,000 in the human genome that make proteins. Certain diseases are linked directly to genes, like those for breast cancer and Alzheimer's. But genes can inform us only of our odds. If my DNA resembles that of other men who have prostate cancer, that means I have a greater risk of suffering a similar fate. My genes can't tell me that I actually have cancer, or that I will, however, only that I have a predisposition for it.

Proteins can confirm an illness is underway, and they often appear in our blood long before we feel sick — months or years before symptoms, when many diseases are still curable. "In the vast majority of cases, it's the proteins we can measure before anything else," says Joshua LaBaer, who founded the Harvard Institute of Proteomics and now directs the Biodesign Institute at Arizona State University, where he is a professor.

This kind of information is invaluable to doctors, who want to know exactly what is going wrong in real time, not what might happen in the future — and it's also why they haven't learned as much from genomics as they had originally hoped. "We've looked a lot at genes in the past 15 years," says Jon Heimer, the chief executive of Olink Proteomics, a biotech firm based in Uppsala, Sweden, that sells protein panels to scan for ailments ranging from organ damage to inflammation. "But it makes more sense to look at proteins, because they are the biological machinery of human beings."

Today, physicians rely largely on pattern recognition to make clinical diagnoses: They match observed symptoms with associated ailments. But there's a pitfall to this approach — symptoms are sometimes absent. Our bodies can function well when they're ill, even gravely so. It's an evolutionary survival mechanism that prevented my mother's heart from divulging any clues to its infirmity until it failed.

In general, we don't go to the doctor until we're feeling lousy — a persistent cough, say, or a throbbing headache. But there are deadly conditions that don't always exhibit conspicuous signs. Cardiovascular disease, stroke, cancer, Alzheimer's, diabetes and kidney dysfunction — six of the top 10 things that kill people in the United States, accounting for 1.6 million deaths annually, according to the Centers for Disease Control and Prevention — can all manifest without symptoms. But proteins often show up first.



↳PART OF THE EQUIPMENT SOMALOGIC USES TO CREATE APTAMERS.

Hundreds of ailments, like strep throat, flu and H.I.V., among others, are already diagnosed through single-protein tests. The prostate-specific antigen protein, or P.S.A., can indicate if a man has prostate cancer. The blood drawn at your annual physical is analyzed for a variety of its proteins, like hemoglobin and lipoproteins, which flag cholesterol levels. Perhaps the most common (and reliable) test is the pregnancy pee stick, which measures human chorionic gonadotropin, or hCG, a hormonal protein produced by the placenta.

But complex disease fingerprints with swarms of proteins are exceptionally difficult (and time-consuming) to spot using conventional "wet lab" methods. "To your eyes and my eyes, we won't catch them," says Farokhzad, who helped start Seer, a biotech firm, in 2017 to spin off his academic research into marketable protein tests. "But to sophisticated machine-learning algorithms, these things pop out like daylight." Once algorithms have identified these fingerprints, researchers can use them to develop tests like the one I took for my heart.

"Diagnostic medicine has always been about proteins," says Philip Ma, Seer's president and chief business officer. "All proteomics is allowing you to do is to look at them in bunches instead of one at a time."

LARRY GOLD, a founder of SomaLogic, has wispy white hair and a bushy mustache, making him look a bit like Mark Twain. "We got our unicorn status," he said, when we met at his office in June, modeling a knit cap with a rainbow-banded horn protruding from its crown. "I guess that's kind of a big deal." The designation is given to privately held companies valued at more than $1 billion. He was wearing bluejeans and a long-sleeve black T-shirt, and he has 30 more identical ones that he keeps stacked in a drawer and in regular rotation. "I don't care what I look like," Gold said. "But I have a dress-up outfit somewhere in my closet for raising money."

Gold, who is 77, moved to Boulder after securing a professorship at the University of Colorado, a few years after earning a Ph.D. in biochemistry from the University of Connecticut. He ran his own lab — now adjacent to the 111,905-square-foot Gold Biosciences Building — and still teaches there occasionally. During the 1980s, one of his graduate students was a scientist named Craig Tuerk, currently a biochemistry professor at Morehead State University in Kentucky. At the time, there were several available methods to determine the presence of a single type of protein, among them the use of antibodies, which Gold and Tuerk were familiar with. A protein of interest can be injected into an animal, like a mouse, whereupon its immune system will make antibodies to fight the invading molecule. These "binding antibodies" attach themselves like Lego bricks to the protein, which becomes easier to spot as a result. The mouse antibodies are subsequently harvested for research and can be used to develop diagnostic screenings.

In 1989, four years after joining Gold's lab, Tuerk was conducting research for his Ph.D. thesis on viral

PHOTOGRAPH BY BENJAMIN RASMUSSEN FOR THE NEW YORK TIMES

DNA. While experimenting with a type of virus that infects bacteria, Tuerk noticed that strands from the virus's RNA had somehow bound themselves to specific proteins. DNA and RNA are nucleic acids. What Tuerk had discovered accidentally was a way to identify immense numbers of proteins at once with nucleic acids instead of antibodies, using a molecule found on both DNA and RNA called an aptamer. It was a momentous breakthrough. "We didn't start dancing because we were being science-y types," Gold said. They did, however, immediately begin drafting a patent, which Gold stayed up all night to write. The next evening, a biotech entrepreneur named David Brunel invited Gold over for Thanksgiving dinner. "I fell asleep on the floor of David's house," Gold said. "People were walking over me."

Brunel later invested in SomaLogic, and Gold has put at least $20 million of his own money into the firm, made from the sale of two companies he founded previously. At his venture before SomaLogic, during the early 1990s, Gold helped develop the first F.D.A.-approved aptamer-derived medication, for macular degeneration, which had roughly $200 million in sales in its first year on the market. But even then, Gold had bigger plans for aptamers — to use them to analyze proteins on an unprecedented scale. He would glean patterns from the data that could diagnose diseases. It was possible to do the same thing with antibodies, but that would require luck and patience, like trying to catalog every fish in the ocean with a net that captured only a single species at a time. And even then, in those years — right about the time when Marc Wilkins, a graduate student at Macquarie University in Australia, coined the word "proteomics" — it might have taken months for researchers to make sense of the voluminous data. Proteins are folded into intricate three-dimensional structures. Mapping one was challenging enough; mapping thousands was all but unmanageable.

"Antibodies are used mostly to measure one protein," Gold said. "They would never scale to what I thought proteomics would have to do to make it significant in health care" — that is, generate information quickly and cheaply enough to be practical for doctors to use in daily clinical practice. "I had this idea that if you measured enough proteins, you'd be able to get insights into human biology that were hard to get any other way."

SOMAscan is the manifestation of his vision. It's a twofold technology platform, combining machine learning with a chemical process to isolate 5,000 proteins from a single drop of blood. Williams says this scan has found fingerprint-like patterns for more than 50 diseases, including lung and pancreatic cancer, both notorious for their dismal survival rates.

What these fingerprints convey can be grouped into three categories: probability (your odds of getting sick), current state (you're already sick but don't know it) and trajectory (how soon you'll get worse). "There are patterns that do each of these jobs," Williams told me, adding that his ultimate goal is to find all these patterns, for any condition, in a single scan, and also measure whether they change in response to lifestyle improvements or medication.

PROTEOMICS HAS FACED something of a chicken-or-egg dilemma. Doctors won't embrace the technology until they are sure that protein screenings provide reliable results, but improving reliability is largely contingent on widespread adoption. Put another way, the greater the number of patients who are tested, the more accurate the fingerprints become.

Before it had patients whose proteins it could analyze, SomaLogic began building its disease database with biobanks, which store frozen blood and tissue for research. Using specimens from these repositories — the ones at the National Institutes of Health and many research hospitals are made accessible to scientists — the company began to hone its machine-learning algorithm, training it to search for key protein configurations. Donors to biobanks are anonymous, but their health data is not. By cross-indexing newly identified protein patterns with medical histories, researchers might find a new disease fingerprint in a subset of donors who had liver cancer, for example. SomaLogic can duplicate or confirm these findings by scanning different samples and using different biobanks, which the company is doing now.

Ideally, SomaLogic would run blinded trials, the gold standard for validating new medical drugs and diagnostics. But there is an ethical obstacle to doing so. SomaLogic typically focuses on a single disease when it's evaluating a disease fingerprint in a real-world setting, which is what Gold enlisted Trujillo to do with the heart test back in 2016. For a blinded trial, Trujillo would have had to randomly select some of his patients with a high-risk fingerprint and then — without informing them of the looming threat — wait to see if they had a heart attack.

Trujillo's work is limited to his patient pool, about 180 people so far. But SomaLogic's continuing studies typically entail thousands of patients, whose proteomes are sampled and tracked longitudinally. In lieu of a trial, SomaLogic is able to confirm data retrospectively. Over the span of, say, a lung-cancer study, some people in the cohort will get other ailments. If one of those subjects in the lung cancer study develops diabetes, for instance, SomaLogic can check whether she carried the fingerprint for that disease when her blood was drawn years earlier.

Once a fingerprint is found, it can still need refining. The cardiovascular test from SomaLogic intrigued Robert Gerszten, the chief of cardiovascular medicine at Boston's Beth Israel Deaconess Medical Center and also a Harvard Medical School professor. Gerszten had patients with a condition called hypertrophic cardiomyopathy, which causes the heart muscle to thicken abnormally. It can be treated by medically inducing a heart attack, which thins the affected tissue. "It's one of the few examples where you *know* the person is going to have a big heart attack," Gerszten says. He sampled blood from his patients with hypertrophic cardiomyopathy both before and after the procedure, as well as from people who'd had ordinary heart attacks.

Not only did Gerszten find proteins that matched with those that SomaLogic had identified, but he also came across ones never previously tied to cardiovascular health. "We found dozens and dozens of new proteins that no one had discovered," he says. While Gerszten had helped validate that protein-screening panels could presage heart attacks, he also illustrated that the underlying biology was remarkably more complicated.

Other factors also hinder proteomic investigation. One is a statistical anomaly known as "overfitting," which happens when trying to match a disease that involves scads of proteins with too few patients. As LaBaer puts it, "There is a chance you're going to find a set of markers that look real but are not." Another is the tendency of proteins to interact with other molecules and change after they're formed, a process known as post-translational

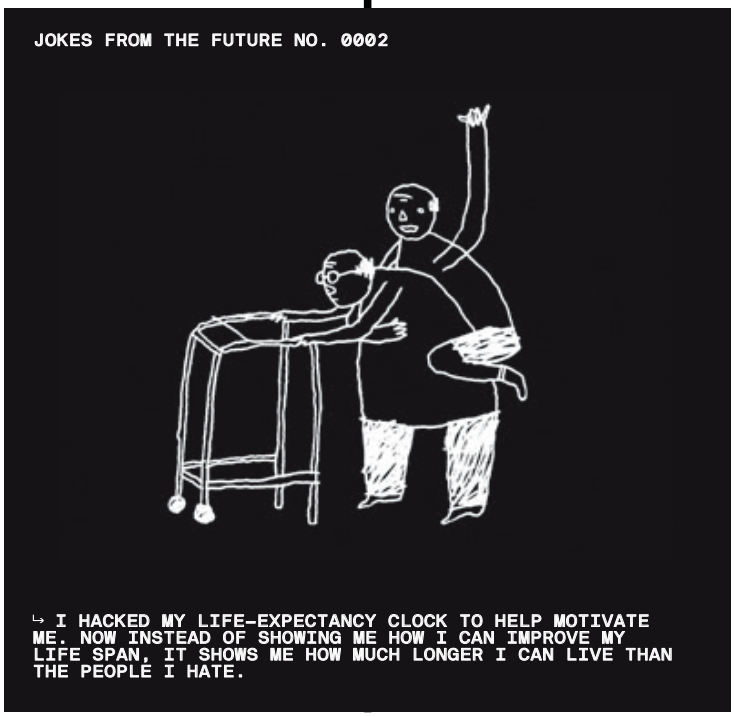


ILLUSTRATION BY BRIAN REA

modification. “There is a landscape of these modifications — a real zoo of molecules — most of which we don’t understand what their effect is,” says Steven Carr, senior director of proteomics at the Broad Institute of M.I.T. and Harvard. Scientists don’t always know whether it’s the actual proteins or the modified ones that are associated with trouble. Nor can they be sure if a blood test is necessarily distinguishing between these two protein structures — one of which may be malicious and the other benign — or even capable of detecting certain modified proteins at all. “Not everyone who has these proteins might get that disease,” Carr says. “And some who have the disease might not have that particular form of proteins.”

SomaLogic’s hunt for proteins starts in a sparse room at its headquarters, in eight-foot-tall upright freezers set to minus 80 degrees Celsius. They are filled with small trays holding dozens of inch-long plastic tubes, each containing a drop of serum (the clear liquid remaining after clotting compounds are removed from blood). At any given time, there are approximately 300,000 samples on site, stored in 13 freezers. Technicians oversee robotic arms programmed to add aptamers to the samples, with a fluorescent light-refracting tag. (SomaLogic has engineered 5,000 aptamers and rechristened them SOMAmers.) After more robotic juggling, the serum solution from the tubes is placed onto glass slides, and an imaging device measures the light intensity passing through the fluorescent aptamer to assess which proteins are present and in what concentrations. It’s here that the biological information is converted into digital data.

It takes 30 hours to acquire protein data from one sample, and the lab processes about 680 a day. The raw data — totaling some four million protein measurements every 24 hours — is fed into machine-learning algorithms, which are revised constantly based on the various disease patterns they’re interested in investigating.

While allocating most resources to the big three — cancer, heart disease and diabetes — SomaLogic also is delving into realms that traditionally haven’t been studied with proteomics, such as smoking, social deprivation, excessive alcohol consumption and fitness. There are hundreds of

↳TUBES HOLDING INDIVIDUAL SAMPLES AT SOMALOGIC.

proteins common among tobacco smokers, which Williams speculates could be used to expose those suspected of smoking but who deny it. “It’s the worst lifestyle health risk you can get,” he says. “But people lie about it.”

SomaLogic will commence its first large-scale beta test this year, collaborating with the Leeds Center for Personalized Medicine and Health in England. Williams told me that the project is starting with a diabetes trial, with other diseases to follow. More than one-third of British residents are prediabetic — meaning that they are not exhibiting symptoms but are at risk of developing the disease. “We don’t want those people to become diabetic,” Williams says. “But we don’t really know who is susceptible.”

SomaLogic will collect blood samples from patients in Leeds and then zero in on those whose protein fingerprints suggest that diabetes is imminent (the company recently set up a 3,000-square-foot lab in Oxford to process samples). Physicians, in turn, will instruct patients in strategies that have proved to pre-empt the need for diabetic medication, like exercise, weight loss and nutrition counseling. “The point is that this information may motivate people straightaway,” Williams says.

Gold wants SOMAscan to eventually evolve into something he calls the “wellness chip,” a do-it-all protein screening to replace annual physicals, which several studies have suggested rarely benefit health. In May, at an annual health care and science symposium that Gold hosts at the University of Colorado, SomaLogic’s chief corporate strategist, Mark Messenbaugh, displayed a slide during his lecture that showed a toilet equipped with a SOMAscan chip and captioned: “Our Ultimate Goal — SOMAwhiz.” He wasn’t joking. Williams confirmed later that the company is adapting its technology for “a noninvasive home-collected urine test” to look for many of the same disease fingerprints found by SOMAscan.

Eventually, SomaLogic hopes to sell its blood and urine scans directly to consumers, for as little as $100 per test, Gold says. But there is a fear percolating among some scientists and health care providers about what patients will do with the data. Will they demand treatment from their doctors based solely on a proteomic scan? More important, will their doctors comply?

“There are consequences to these measurements,” Carr says. “There are interventions taken, drugs people are put on, additional testing. That costs money, and it raises the anxiety level of those being tested. So you damn well better be sure you’re measuring what you say you’re measuring — and know that it matters — before you employ it in a clinical setting.”

WHEN MY RESULTS came back from SomaLogic, a month after my blood was drawn, they told me that I had an 11 percent chance of experiencing a cardiac event within five years — more than three times as high as what the coronary calcium scan and associated algorithm had forecast. Trujillo insisted that I immediately start taking a statin drug to halt further plaque buildup.

Because proteins react to external inputs, Trujillo asked everyone who got the heart screening to redo it one year later. From what we know about the proteome, any actions taken to prevent a heart attack — dietary improvements, exercise, medications — should nudge the odds lower. For now, it’s too early to draw conclusions from Trujillo’s follow-up data. But, he says, “the results have really allowed me to personalize care. They have motivated people to change their lifestyle or take medication.”

I told Gold that my anniversary date for the heart test was approaching and that I would soon be providing a second blood sample to Trujillo. He warned me that my risk might not fall precipitously, as evolving proteomic research is showing that cardiovascular disease, like so many other ailments, is a more intricate biological puzzle than once thought. The results from my second test confirmed as much: I did not drop a single percentage point. I had spent a year on statins, mostly eliminated red meat from my diet and added high-intensity interval training to my workouts. I also did a five-day fast that Trujillo recommended for reducing inflammation — a well-known contributor to heart disease — and in the course of it lost eight pounds. None of it appeared to help.

Trujillo was untroubled. “My work has led to a page of questions as it relates to how to use this test going forward,” he said. “Is it God’s word? Not even close. It’s just part of the armamentarium.” Trujillo reckoned that fathoming the proteome might spawn new mysteries while solving others.

Gold echoed a similar sentiment when we talked. “We know more about the proteome here than anyone on earth, and we think it’s a treasure trove of understanding human biology,” he said. “But I won’t lie about it. The science is hard — harder than I thought.” ◆

PHOTOGRAPH BY BENJAMIN RASMUSSEN FOR THE NEW YORK TIMES