

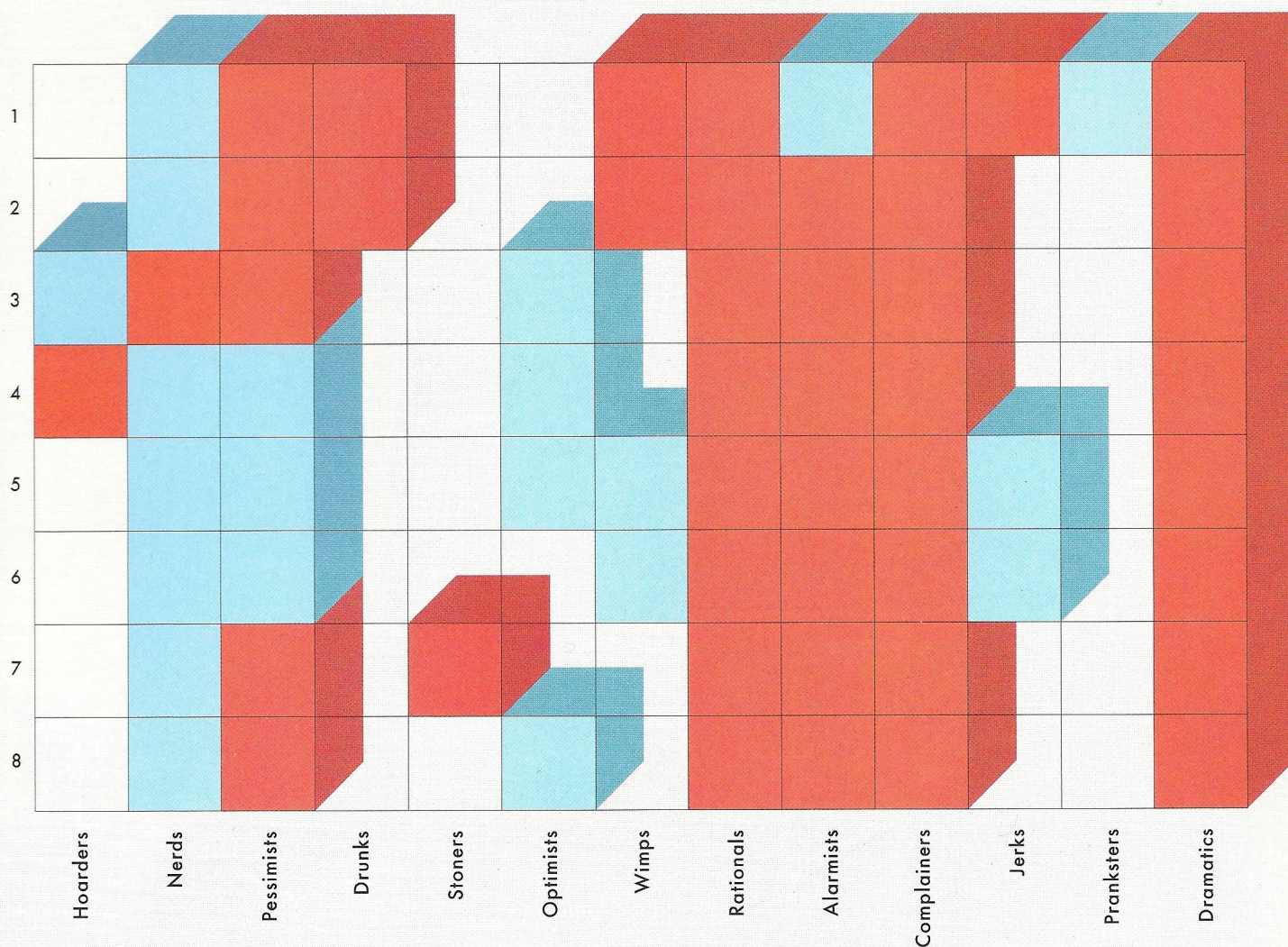
# GOOD

## The Data Issue

Qualifying our quantifying in the age of information overload

ISSUE 024 / FALL 2011

WHICH KINDS OF PEOPLE LIKE WHICH CHARTS?



### CHARTS ABOUT

1. Divorce
2. Health Care

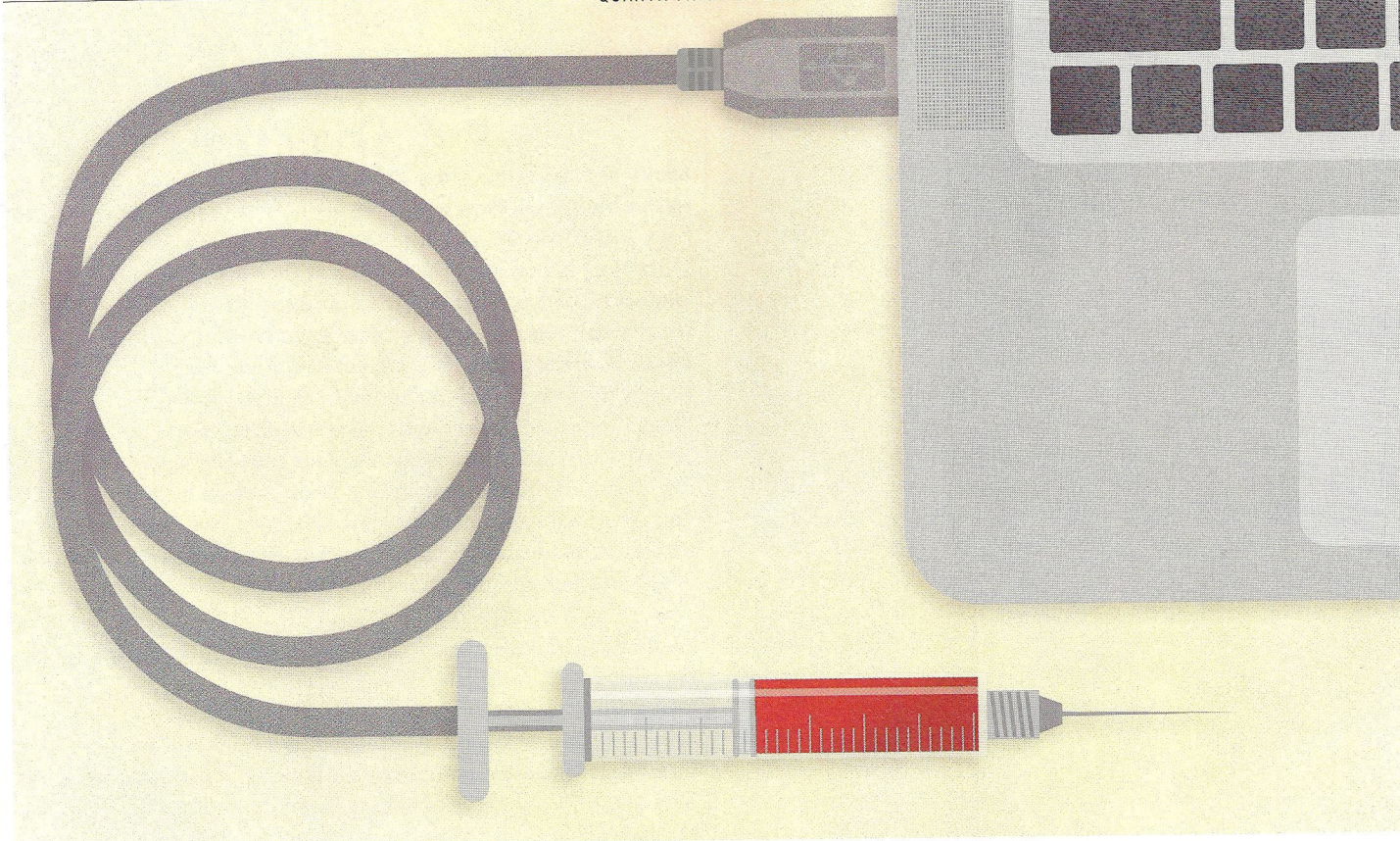
3. The U.S. Dollar
4. Real Estate

5. Foreign Policy
6. Gas Prices

7. The Ozone Layer
8. Natural Disasters

- Probably  
Definitely





## SELF-STORAGE

Data sharing is driving medical care, and we should all opt in.

by Steven Leckart  
illustration Trevor Burks

**One of the greatest undertakings** in modern medicine didn't begin in a lab or a doctor's office, but on the streets of a small town 20 miles southwest of Boston. In 1948, the U.S. Public Health Service chose Framingham, Massachusetts, as ground zero for a long-term study on heart disease. In their book, *A Change of Heart*, Daniel Levy and Susan Brink describe how a team of local volunteers went door to door, spoke at PTA meetings, and made countless phone calls to sign up their neighbors.

The Framingham Heart Study was unprecedented in scale and size. Researchers examined 5,209 adults, measuring each by 80 variables including cholesterol, weight, blood pressure, and lung capacity. The plan: Re-examine those same patients every two years,

write down everything on paper, and file it away. The idea: Look for patterns in the data over time. As Levy and Brink put it, "Science is a collaboration, not an isolated effort. It is like a relay race toward answers."

Ultimately, our understanding of the cardiovascular risks associated with smoking, poor diet, and lack of exercise all stem from the Framingham data, which have since been digitized. The numbers have been the foundation for more than 1,200 published studies. Currently, geneticists are evaluating and expanding on the data to see how heart disease might be wired into our DNA. In the next year, the National Heart, Lung and Blood Institute expects its Framingham-specific database to balloon from 1.3 terabytes to 100.



And back in Framingham, the granddaughters and grandsons of the original volunteers are getting pricked, poked, prodded, and X-rayed every two years in what is now called the Offspring Study.

We're on the cusp of what might eventually blossom into a dozen Framinghams, provided we as patients let ourselves be counted. There's been much hand-wringing about the need for more doctors to adopt electronic medical records, and a major benefit of going digital is that it can bolster research. One of the first HMOs to adopt e-records, Kaiser Permanente is quickly evolving from health care provider to health insight provider. In 2009, the National Institutes of Health awarded Kaiser 22 grants totaling \$54 million for research using its database of 9 million patients. Today, Kaiser's Division of Research oversees more than 250 ongoing studies on everything from diabetes and autism to epilepsy and back pain. Simply by seeing a doctor in the network, each of Kaiser's 9 million patients could be contributing.

Kaiser isn't alone. An even bigger data set is managed by Practice Fusion, a company that in 2007 began providing doctors with electronic-record and practice-management software free of charge. Four years later, the company's cloud-based database consists of 12 million patients being treated by more than 75,000 doctors. The data are backed up with three biometric security checks; the software is as secure as networks maintained by credit card companies.

But how can the San Francisco-based start-up afford to give away what its competitors value at \$50,000? Selling anonymized medical charts to research institutions could bring in two and a half times the revenue of traditional software sales. Thus far, Practice Fusion relies on funding, but the potential value of this type of data hasn't escaped the company. "Once it's de-identified, it becomes all of our data," says Matthew Douglass, vice president of engineering at Practice Fusion. "In 20 years, with all the aggregating and data mining that's taking place now, this information will be used to cure something. It's only a matter of time."

For now, we can begin answering a multitude of questions. Practice Fusion has contributed data to two forthcoming medical studies at Harvard, research that otherwise might have been cost-prohibitive. The company's internal research team published a Prescription Index, which breaks down the most common medications prescribed by specialty. Last year, the company also uploaded 15,000 anonymized records to Microsoft's Azure Marketplace, a one-stop shop for buying, selling, trading, and sharing large sets of data. In that case, Practice Fusion's data were made available for free as part of a public "hackathon" sponsored by the Department of Health and Human Services, which has recently been working to aggregate and make public huge swaths of health data on the websites [data.gov/health](http://data.gov/health) and [healthindicators.gov](http://healthindicators.gov). (These aren't personal records like Framingham, but more like smoking rates or hospital quality ratings.) By mashing up, say, obesity rates with the

## "WITH ALL THE AGGREGATING AND DATA MINING THAT'S TAKING PLACE NOW, THIS INFORMATION WILL BE USED TO CURE SOMETHING. IT'S ONLY A MATTER OF TIME."

number of fast-food restaurants per square mile, programmers can see correlations that could affect public policy.

Investigating potential epidemics in real time is another benefit to digitized data. Doctors taking part in ecosystems like the ones used by Kaiser and Practice Fusion can share charts and general observations to understand if, say, a new medication could be causing a wave of unanticipated side effects.

If the term "syndromic surveillance" sounds like a panoptic nightmare, it shouldn't. How often have you heard someone on Twitter or Facebook mention they have the flu?

Epidemiologists already routinely comb social media to track disease. In a study on influenza published in April, researchers determined that monitoring Twitter can "estimate disease activity" in real time up to two weeks quicker than the Centers for Disease Control's system. Of course, the use of social media is limited. Most people would never tweet about having an STD, for instance. But the larger point is obvious: Networked, digital databases are incredibly efficient. Imagine how many more people would report having the flu if it were just a matter of sending a private alert to their doctor. Now think bigger.

When the Framingham researchers got going in 1948, they didn't know what the data might lead to, so they locked away all that paperwork in filing cabinets just to be safe. Today, health care providers typically discard 90 percent of the data they generate, according to a recent report by the McKinsey Global Institute. What if some of what we're tossing—or still storing in filing cabinets—could eventually prove valuable? Ninety percent of doctors in the United States are still on paper.

It's easy for us to say, "That's their problem."

But consider Framingham. All that research, all those insights exist because average folks were not only willing to sign up, but went door to door to make sure their neighbors did, too. Today we have it so much easier. Saving our medical data is often as simple as checking a box. As long as the info is being logged, it could eventually become available for research. Maybe it never will, but at least there's a chance. ■

**BIO** Steven Leckart is a correspondent for *Wired* and has written for *Men's Health*, *TheAtlantic.com*, and *Men's Journal*.