
Emerging Biotechnologies to Promote Food Safety

JOHN BESSER

*Centers for Disease Control and Prevention
Atlanta, Georgia*

jtw8@cdc.gov

We have seen a massive explosion in technologies, especially in molecular biology, that started affecting public health in the 1990s and continues to accelerate. It's a good time to discuss the impact that biotechnology will have on food safety, especially in my specialty of foodborne-disease surveillance. Many technologies—MALDI-time¹ of flight mass spectroscopy, microarrays, sequencing, microfluidics, *etc.*,—are changing our concepts of microbial life, which is affecting how we detect and how we control microorganisms in their natural environments.

Each year, one out of every six Americans—48 million people—are thought to become sick with a foodborne illness, and 3,000 die. I'll provide background on foodborne-disease surveillance, and what it does for us, and on some of the limitations of surveillance and the impacts of technology.

The main points I will make are:

- Foodborne-disease surveillance is an important, but often overlooked, component of our food-safety system.
- How well it functions—or doesn't function—is vitally important to industry and to the public.
- The current system operates at only a fraction of its potential.
- New technology can exponentially magnify its effectiveness.

RECENT OUTBREAKS

2010 started out with *Salmonella* Typhimurium infections reported from forty-one states, caused by human contact with African dwarf water frogs. In the same year, widespread salmonella infections were associated with shell eggs, frozen meals, alfalfa sprouts, Romaine lettuce, and salami made with contaminated pepper, and *E. coli* O157 outbreaks were traced to beef and cookie dough. 2011 is shaping up to be another banner year for foodborne disease.

¹Matrix-assisted laser-desorption ionization.

PREVENTION

Much can and should be done to prevent foodborne illness, from farm to fork: good agricultural practices, good manufacturing practices and inspections, designing processes for safety, microbial monitoring, restaurant and food-store inspections, and consumer education. However, in spite of everything we do, foodborne illness will occur because we are imperfect beings. Some 356 billion pounds of food are consumed annually in the United States and it’s impossible to monitor it all. Contamination, which can occur anywhere along the food chain, can’t be seen and is unevenly dispersed within the affected product. Accordingly, detecting pathogens in food is an insensitive process. On the other hand, essentially all of the food consumed in the United States is, in a way, being tested because it is being eaten, and disease surveillance provides information on what can be done to reduce the burden of illness. Furthermore, surveillance can help limit ongoing illness by recalls, public notices, and publishing of guidelines.

PULSENET

Figure 1 lists US recalls—some of which have been massive—in which PulseNet played a role in detecting outbreaks and averting disease. The much more profound impact of disease surveillance is that it allows identification of underlying problems and their solution, providing feedback to industry, to regulators and to consumers about problems that would otherwise be unrecognized.

Year	Pathogen	Food	Amount recalled
2010	<i>Salmonella</i> Enteritidis	Shell eggs	>500,000,000 eggs
2010	<i>Salmonella</i> Montevideo	Ready-to-eat Italian sausage products/pepper	>1,263,754 lbs
2009	<i>E. coli</i> O157:H7	Non-intact steak and ground beef outbreaks	1,115,049 lbs
2009	<i>E. coli</i> O157:H7	Cookie dough	300,000 cases of product
2009	<i>Salmonella</i> Typhimurium	Peanut butter/peanut products	>3000 types of products
2008	<i>E. coli</i> O157:H7	Ground beef	5,300,000 lbs
2007	<i>E. coli</i> O157:H7	Frozen pizza	5,000,000 pizzas
2007	<i>E. coli</i> O157:H7	Ground beef (3 outbreaks)	35,400,000 lbs
2006	<i>Salmonella</i> Tennessee	Peanut butter	326,000,000 lbs
2004	<i>Salmonella</i> Enteritidis	Raw almonds	13,000,000 lbs
2003	<i>E. coli</i> O157:H7	Blade Tenderized Frozen Steak	750,000 lbs
2002	<i>Listeria monocytogenes</i>	Ready-to-eat poultry products	27,400,000 lbs
2002	<i>E. coli</i> O157:H7	Ground beef	18,600,000 lbs
2000	<i>Listeria monocytogenes</i>	Ready-to-eat poultry products	16,900,000 lbs
2000	<i>E. coli</i> O157:H7	Ground beef	1,100,000 lbs
1998	<i>Listeria monocytogenes</i>	Hot dogs, deli meats	35,000,000 lbs
1998 & 2008	<i>Salmonella</i> Agona	Toasted oats cereal	>3,000,000 lbs
1997	<i>E. coli</i> O157:H7	Frozen ground beef	25,000,000 lbs

Figure 1. Largest US food recalls in which PulseNet played a prominent role.

Figure 2 shows a few of the industrial processes that, over the years, have been changed in order to reduce the burden of illness. And Figure 3 shows the result of a study done by Rob Tauxe at the CDC on recent outbreaks, showing vehicles that were not formerly realized to be risky and weren't high on the "radar screens" at the FDA or USDA. Who would have thought that peanut butter would be a significant vehicle for salmonellosis, for instance, or that raw cookie dough could cause illnesses? These were picked up through our disease-surveillance system, allowing regulators and industry to direct their scarce resources towards where problems were actually occurring.



Figure 2. Addressing underlying problems.

Figure 4 shows some ingredient-driven outbreaks, which, formerly, would have been difficult to identify. To a certain extent, the recent situation in Germany was ingredient-driven; alfalfa sprouts are seldom eaten alone.

Figure 5 provides an illustration of the surveillance system. People become ill and visit their doctors who request stool samples and microbial cultures are sent to a laboratory. If a reportable pathogen is found, an isolate is sent to the Health Department for sub-typing. Representatives of the Health Department interview cases to find out what they ate and what they were exposed to. When the information is uploaded to PulseNet, it is reported to the CDC. FDA and USDA and other organizations are involved in tracking the cases of disease, using the information to try to minimize the impact. Other modes of finding information are used also. There is a system whereby state health departments are called up with clusters that are recognized by physicians or the public, but Figure 5 illustrates one of the more powerful methods that we have for discovering unrecognized problems in the food supply.

Figure 6 shows what PulseNet does. Every state has a laboratory in a large city where these pathogens are sub-typed. Each lane has a pulse-field gel electrophoresis (PFGE) pattern that is investigated in local databases, then clusters of cases with matching patterns are uploaded and we look at them on a national scale at the CDC database. The regulatory

2006	<i>E. coli</i> O157 and <u>bagged spinach</u>	2008	<i>E. coli</i> O157 and ground beef
2006	<i>E. coli</i> O157 and shredded lettuce (x2)	2008	<i>Salmonella</i> and <u>fresh produce items</u>
2006	Botulism and commercial <u>pasteurized carrot juice</u>	2009	<i>Salmonella</i> and <u>peanut butter containing foods</u>
2006	<i>Salmonella</i> and fresh tomatoes	2009	<i>Salmonella</i> and <u>imported white and black pepper</u>
2007	<i>E. coli</i> O157 and frozen pizza	2009	<i>Salmonella</i> and alfalfa sprouts
2007	<i>Salmonella</i> and <u>peanut butter</u>	2009	<i>E. coli</i> O157 and prepackaged <u>cookie dough</u>
2007	<i>Salmonella</i> and a <u>vegetarian snack food</u>	2009	Multidrug resistant <i>Salmonella</i> and ground beef (x2)
2007	<i>Salmonella</i> and <u>dry dog food</u>	2009	<i>E. coli</i> O157 and <u>blade tenderized steaks</u>
2007	<i>Salmonella</i> and <u>microwaveable pot pies</u>	2009	<i>Salmonella</i> and salami made with contaminated pepper
2007	<i>Salmonella</i> and <u>dry puffed breakfast cereal</u>	2010	<i>E. coli</i> O145 and romaine lettuce
2007	<i>E. coli</i> O157 and ground beef	2010	<i>Salmonella</i> and alfalfa sprouts
2007	Botulism and <u>canned chili sauce</u>	2010	<i>Salmonella</i> and <u>frozen meals</u>
2008	<i>Salmonella</i> and cantaloupe	2010	<i>Salmonella</i> and shell eggs

Figure 3. Selected recent multi-state outbreaks of foodborne infections (2006–2010): new food vehicles (underlined).

agencies also contribute from their food-monitoring programs. Data from FDA are directly uploaded and those from USDA come indirectly through a network called VetNet. A new network is being formed in industry called Voluntary Net; companies are keeping their own inventories of PFGE patterns for rapid early detection of potential problems.

PulseNet USA comprises all fifty states, and several large counties and cities have laboratories that are connected electronically (Figure 7). It started in 1996 in Minnesota, and was officially opened in 1998 by then Vice-President Gore. By 2001, it was present in all fifty states. Each year some 1,500 clusters are investigated at state and local health departments. About 250 multi-state clusters are examined by the CDC, of which ten to fifteen large, dispersed multi-state outbreaks are further scrutinized. At weekly meetings, we triage about fifty clusters and direct our resources accordingly.

PulseNet increases the sensitivity of cluster detection, strengthens the association between illness and exposure, and increases the speed of detection of outbreaks. It does this by amplifying the signal indicating ill cases. The number of patterns uploaded to PulseNet has stabilized at around 50,000 per year (Figure 8). The decrease in 2009 resulted from the emergence of novel H1N1; some states had insufficient resources to investigate both flu and foodborne disease.

2006	<i>E. coli</i> O157 and bagged spinach	2008	<i>E. coli</i> O157 and ground beef
2006	<i>E. coli</i> O157 and shredded lettuce (x2)	2008	<u><i>Salmonella</i> and fresh produce items</u>
2006	Botulism and commercial pasteurized carrot juice	2009	<u><i>Salmonella</i> and peanut butter containing foods</u>
2006	<i>Salmonella</i> and fresh tomatoes	2009	<u><i>Salmonella</i> and imported white and black pepper</u>
2007	<u><i>E. coli</i> O157 and frozen pizza</u>	2009	<i>Salmonella</i> and alfalfa sprouts
2007	<i>Salmonella</i> and peanut butter	2009	<u><i>E. coli</i> O157 and prepackaged cookie dough</u>
2007	<u><i>Salmonella</i> and a vegetarian snack food</u>	2009	Multidrug resistant <i>Salmonella</i> and ground beef (x2)
2007	<u><i>Salmonella</i> and dry dog food</u>	2009	<i>E. coli</i> O157 and blade tenderized steaks
2007	<u><i>Salmonella</i> and microwaveable pot pies</u>	2009	<u><i>Salmonella</i> and salami made with contaminated pepper</u>
2007	<i>Salmonella</i> and dry puffed breakfast cereal	2010	<i>E. coli</i> O145 and romaine lettuce
2007	<i>E. coli</i> O157 and ground beef	2010	<i>Salmonella</i> and alfalfa sprouts
2007	Botulism and canned chili sauce	2010	<u><i>Salmonella</i> and frozen meals</u>
2008	<i>Salmonella</i> and cantaloupe	2010	<i>Salmonella</i> and shell eggs

Figure 4. Selected recent multi-state outbreaks of foodborne infections (2006–2010): ingredient-driven (underlined).

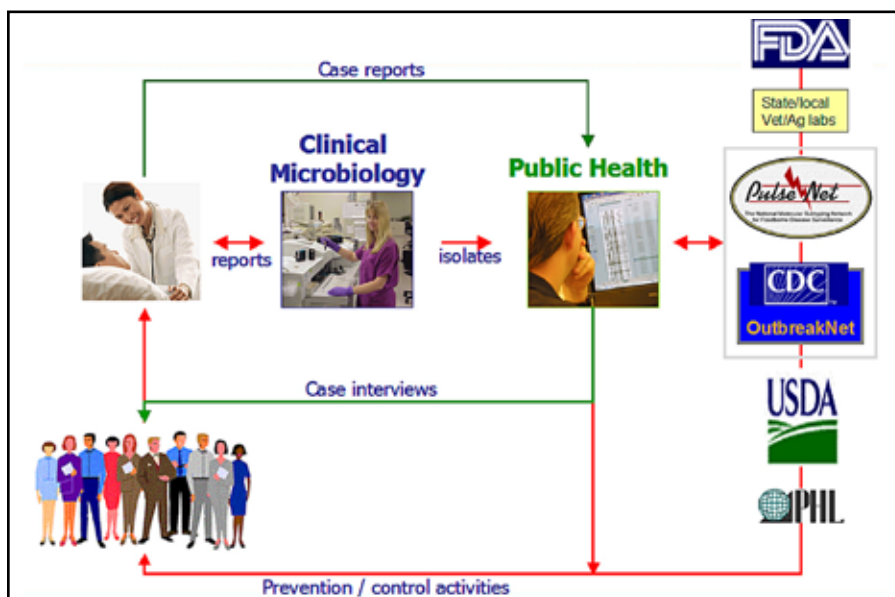


Figure 5. Pathogen-specific surveillance.

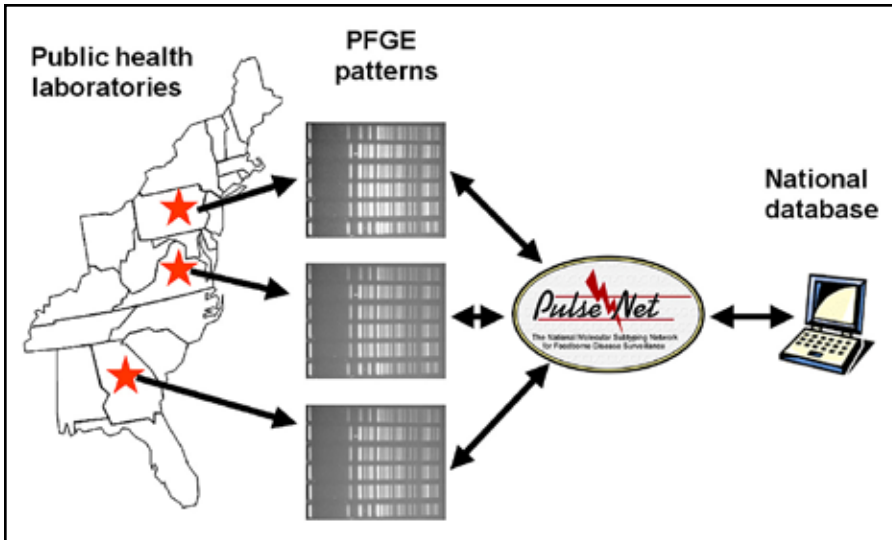


Figure 6. PulseNet electronic communication.



Figure 7. PulseNet USA.

Figure 9 provides an example of how it works. These are cases of *E. coli* O157H7 in Oregon in 2006. Interviews of all of these cases showed a variety of exposures, whereas a subset, sharing a common PFGE pattern, revealed that these individuals had consumed fresh, bagged spinach. It is safe to say that, in the absence of this system, this outbreak—199 cases in twenty-six states, three deaths, and thirty-one cases of hemolytic uremic

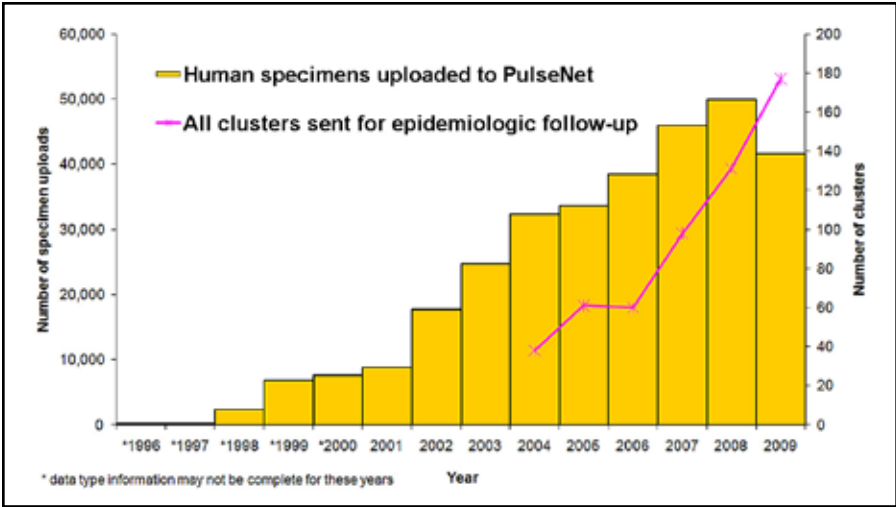


Figure 8. Human specimen isolates uploaded to PulseNet USA and identified clusters, 1996–2009.

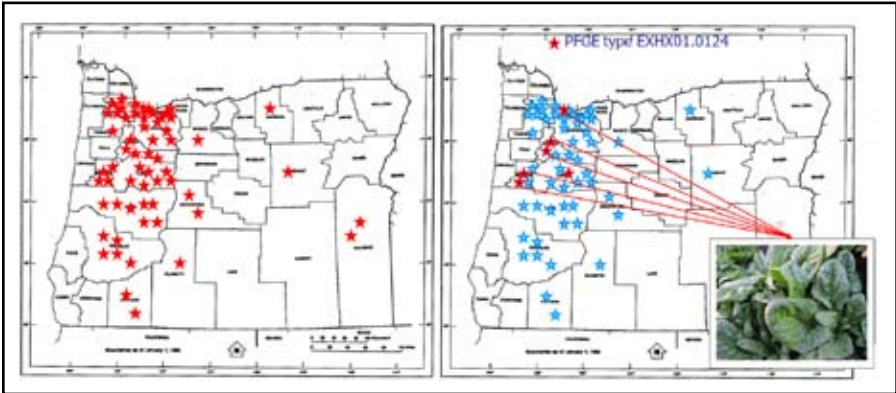


Figure 9. *E. coli* O157:H7, Oregon 6/1/2006–10/9/2006.

syndrome—would not have been detected. Each case of hemolytic uremic syndrome costs about a half-million dollars in medical expenses. A death has been costed at about \$6 million. By comparison, the 2011 sprout-associated outbreak in Germany resulted in 3,304 cases, thirty-eight deaths and 786 cases of hemolytic uremic syndrome. Adding the international cases associated with travel to Germany increases the cases of hemolytic uremic syndrome to 828, which is unprecedented.

Figure 10 provides another example of the signals received, this time for *Salmonella* Typhimurium over a 3-month period. Buried in these data were cases from around the United States that shared a PFGE pattern (Figure 11). They were traced to peanut products that led to 3,000 different items being recalled.

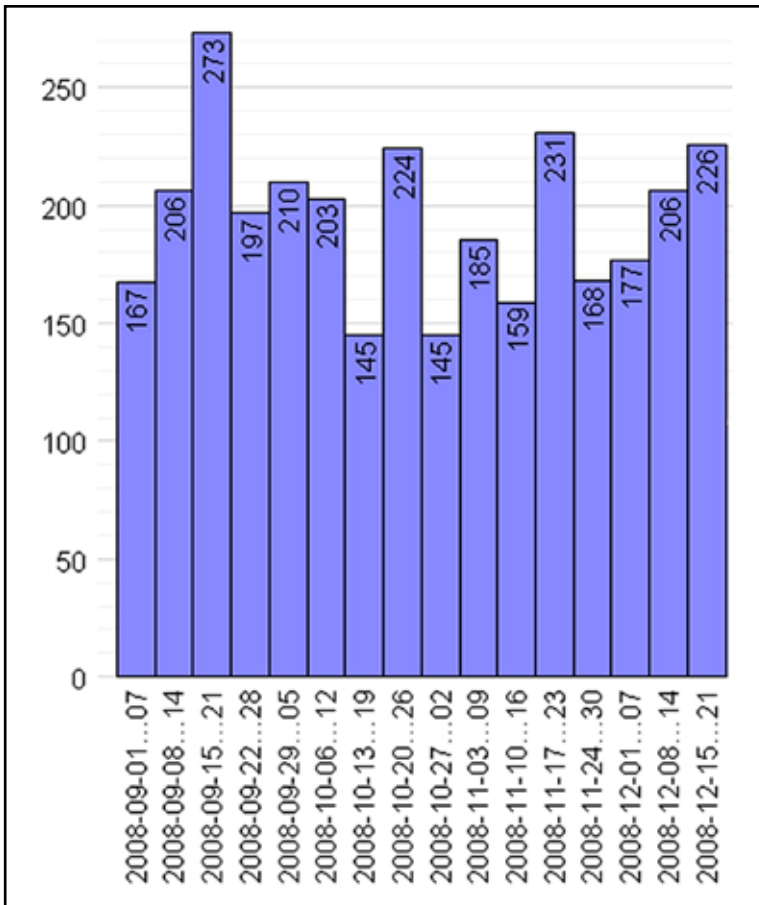


Figure 10. All *Salmonella* Typhimurium, 9/1/2008–12/15/2008, weekly.

“Before” and “after” pictures are shown in Figure 12. The upper “before” picture is the epidemiologic curve from the Jack In The Box outbreak of *E. coli* O157 in 1993, which took a long time to detect and resulted in many cases and four deaths. After seven weeks, 150,000 hamburger patties were withdrawn. The lower “after” pattern, of a 2002 outbreak of *E. coli* O157:H7 in Colorado, starts out looking similar whereas rapid detection led to early recall of hamburger meat and curtailment of the outbreak. This is an example of how PulseNet works and a hundred similar examples exist.

The theory underpinning PulseNet is that by detecting more outbreaks and curtailing them, future disease incidence will be reduced. We have seen this occur with listeria (Figure 13). Subsequent to the initiation of PulseNet for listeria in the 1990s, we detected more outbreaks, and, recognizing the roots of the problems, disease incidence fell. We have a long way to go with shiga-toxin-producing *E. coli* and salmonella, but every case of those diseases is potentially preventable; we need to work harder to reduce the burden of disease.

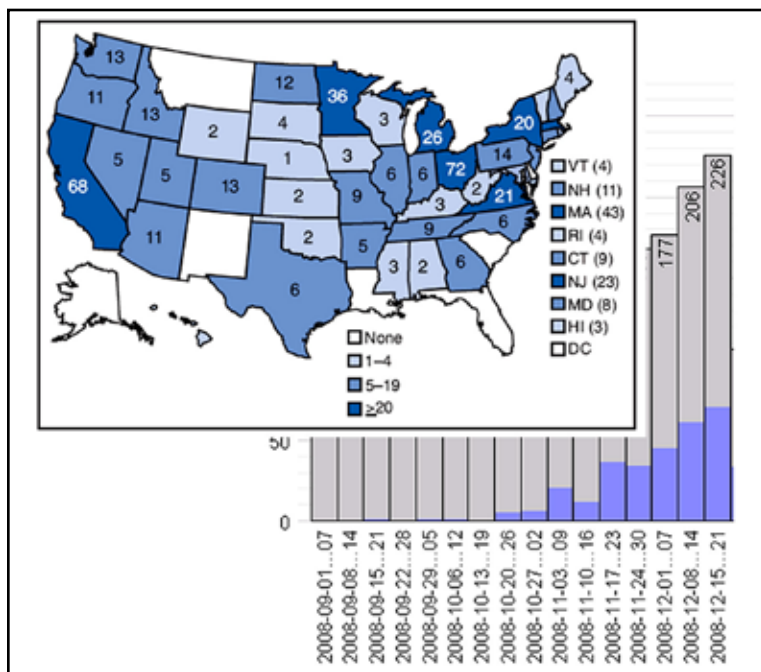


Figure 11. All *Salmonella* Typhimurium 9/1/2008–12/15/2008, weekly; JPXX01.1818, JPXX01.1825 and JPXX01.0459 highlighted.

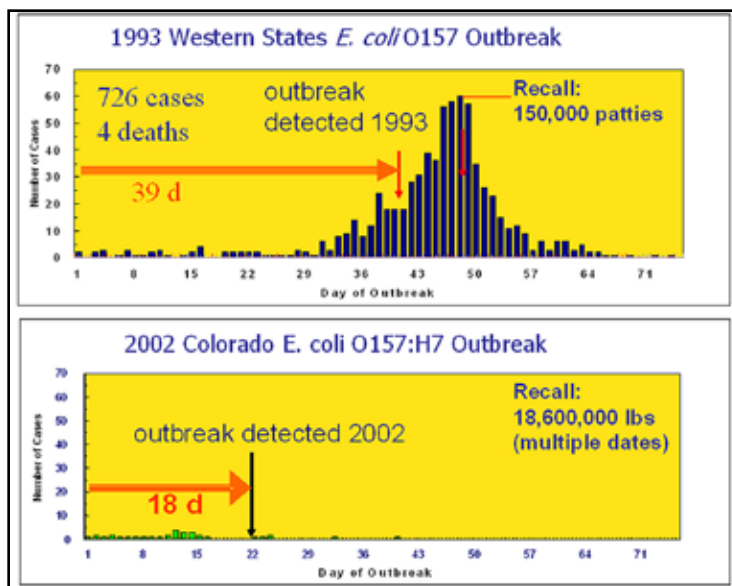


Figure 12. Foodborne outbreaks of disease caused by *E. coli*, before and after PulseNet.

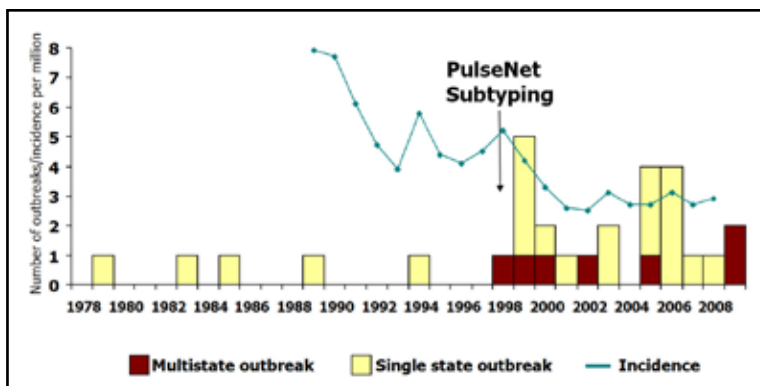


Figure 13. Listeriosis in the United States, 1978–2008.

BIOTERRORISM

The only major act of foodborne terrorism in the United States occurred in 1985, when—to influence an election—members of the Rajneeshi sect in Oregon contaminated the salad bars of ten local restaurants with salmonella, infecting 751 people, of whom forty-five received hospital treatment; all survived. The source of the infection took months to identify, whereas if it occurred today it would likely be detected and resolved quickly.

GLOBAL SURVEILLANCE

Our system is the most sensitive method for detecting unrecognized problems in our food supply—including from terrorism—with organisms that are under surveillance. PulseNet has been so successful in the United States that it has been adopted in many other countries. PulseNet International comprises eighty-four countries. The system in Canada is fully integrated with that in the United States. The Chinese have recognized the negative impacts that foodborne disease can have on trade and they are putting a lot of resources into PulseNet China. PulseNet Latin America and Caribbean is operational. However PulseNet Europe isn't fully integrated because some of the countries there prefer to operate autonomously, and many of the counties in Germany act like independent states. A benefit from the recent *E. coli* outbreak in Germany may be a refocusing of effort in Europe on disease surveillance.

Food is a global issue. Meat, and ingredients in processed meat products, consumed in the United States come from all over the world. The importation of fruits, vegetables, meats and grains has been increasing with our free trade agreements. Even the components of bread come from abroad. Clearly, foodborne disease is a problem of global scope and has to be solved in a global manner.

CURRENT LIMITATIONS

A number of limitations exist:

- We have minimal ability to control strain evolution.

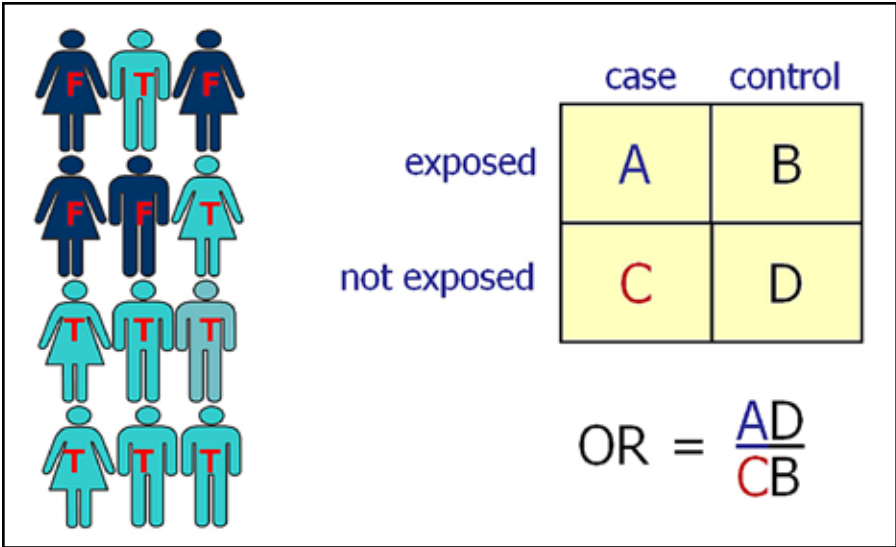


Figure 14. Case definitions for cluster detection and hypothesis generation/testing—1.

- The system is inherently slow and at every outbreak the media question why it takes so long to get information.
- Exposure information is difficult to obtain.
- Effective surveillance is limited to pathogens we know and can detect.

Strain Resolution

Figure 14 shows a group of cases with a particular disease, some of whom are truly associated (“T”) or falsely associated (“F”) with a particular product in an outbreak setting. Of course, when cases are reported to the public-health authorities, it is never known what they have been associated with. And on the right of the figure is a measure of association that is used in case-control studies when looking at what exposures ill people (“case”) had vs. people who are not ill (“control”). This produces a statistical measure, the odds ratio (OR), which I will use to illustrate how sub-typing and case classification help strengthen the association between illness and exposure. If we limit our study to individuals who are more likely to have a common association—in other words if they share a fingerprint pattern in their pathogen—we eliminate cases that are more likely to be falsely associated than truly associated (Figure 15). This improves the proportions in our statistical analysis, and increases the strength of association between illness and exposure. For a more-stringent case definition, we could use two PFGE enzymes instead of one (Figure 16A) and knock out some of the additional falsely included cases; however, we start knocking out truly associated cases as well. If we keep doing that and use, say, ten enzymes (Figure 16B), then fewer cases are left and eventually confidence in the results becomes smaller with smaller sample size. Eventually, with whole-genome sequencing every case would be

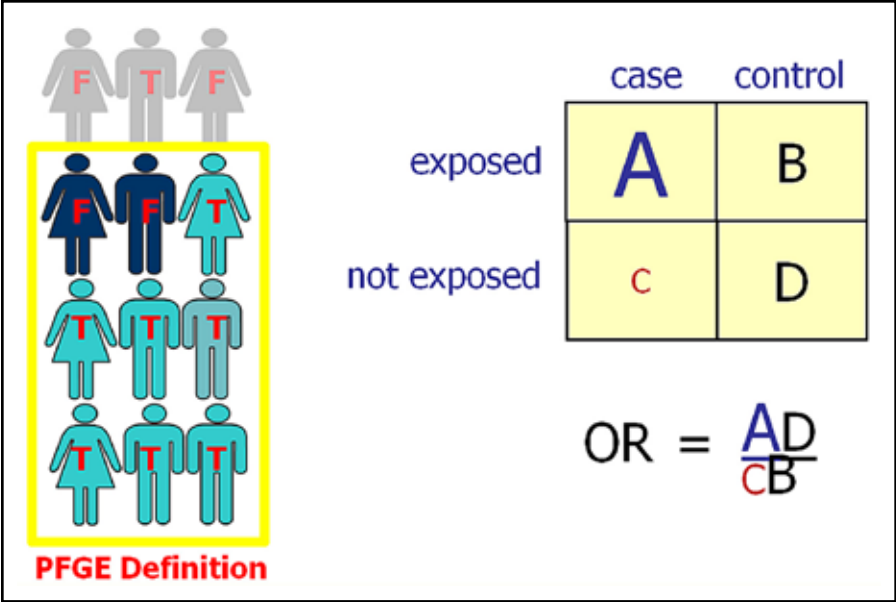


Figure 15. Case definitions for cluster detection and hypothesis generation/testing—2.

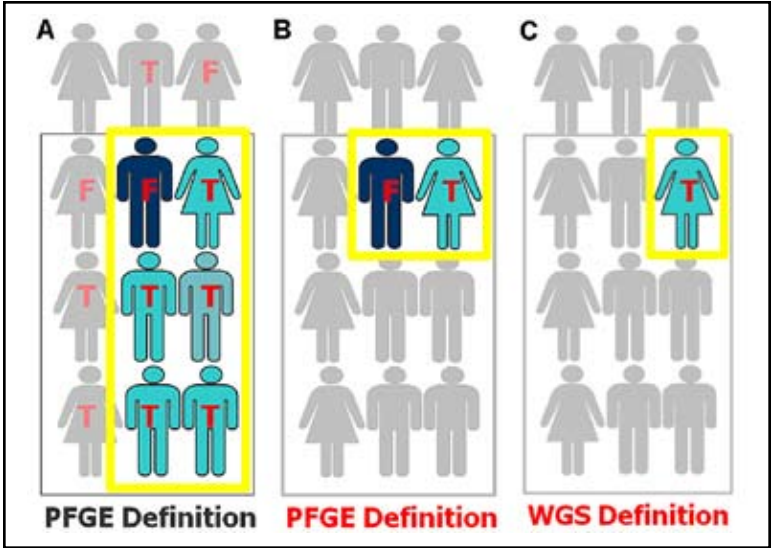


Figure 16. Case definitions for cluster detection and hypothesis generation/testing.
 A—two-enzyme PFGE case definition, B—ten-enzyme PFGE case definition,
 C—whole genome sequence case definition.

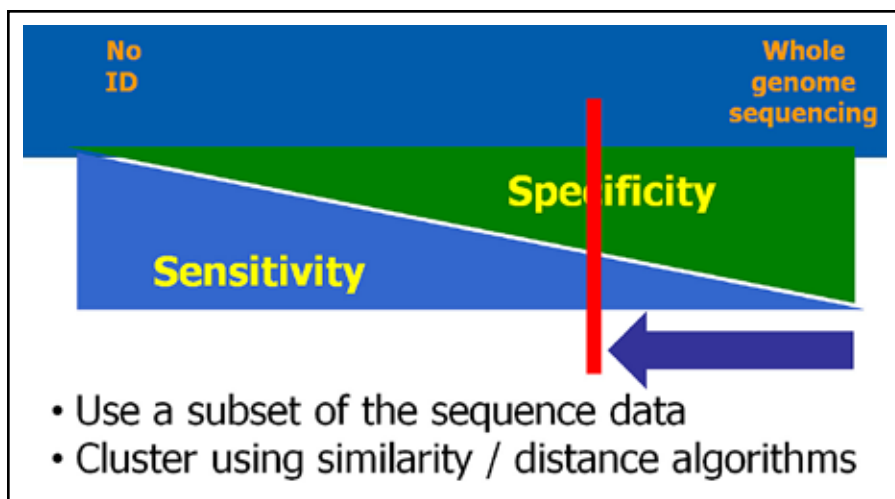


Figure 17. Relationship between sensitivity and specificity.

different from every other case and we would have 100% specificity and 0% sensitivity (Figure 16C), which would be as specific as you could have a case definition.

This relationship between sensitivity and specificity exists in all laboratory tests. At one end of the scale there is full specificity and zero sensitivity, whereas at the other end of the scale there is sensitivity with no specificity. Grouping together people who are sick, without knowing if they have salmonella or *E. coli*, would be a very inclusive case definition without specificity; it would be difficult to show an association between illness and exposure. We need to “move the bar” (Figure 17) to get a strong signal that’s neither too specific nor too sensitive. It has to be somewhere in the middle, which is achievable by using a subset of our data or clustering algorithms, like tuning a radio by maximizing the signal and minimizing the noise. One of the impacts of new technology is fine tuning our signals. When we layer upon that different time intervals and geography, we can look at demographics. These can be done simultaneously in an automated fashion to have multi-dimensional continuous analyses of surveillance data. This would not have been possible a few years ago because of the massive amounts of computing necessary. Soon we will be able to look at surveillance data exposure by exposure and ask the question, “Are any of these exposures potentially different from what we would expect?”

Slow System

Sick patients have to seek medical help and provide stool samples (Figure 18). The pathogens have to be cultured and identified. The cultures have to be shipped to the public-health laboratory. Each case has to get interviewed, and each culture has to be serotyped and sub-typed. This process can take anywhere from a few days to a few weeks. The most important part of the procedure is that interviewed cases must recall what was eaten approximately three weeks prior. It’s amazing that the system works as well as it

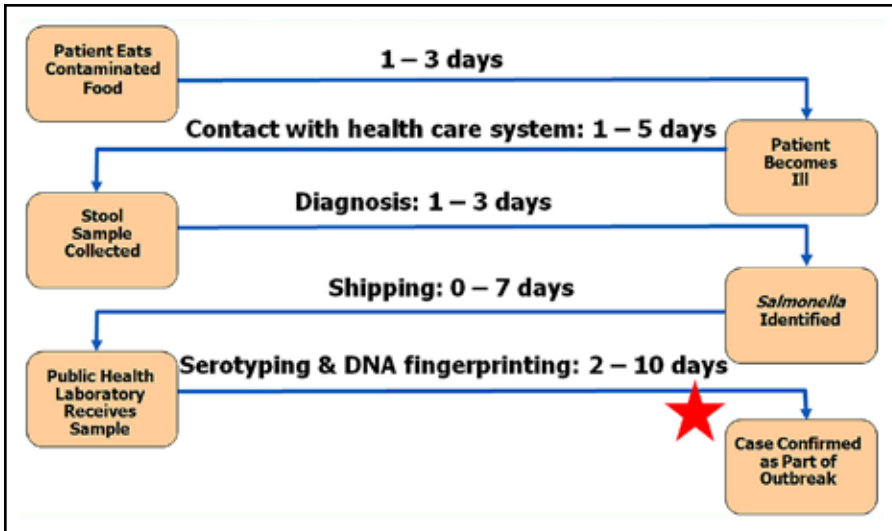


Figure 18. The surveillance process: laboratory reporting takes time

does. But recall drops off asymptotically with time. Ability to remember what was eaten three weeks ago is low, but it's orders of magnitude lower after four weeks, five weeks and six weeks, and at some point reaches zero.

There is potential to shave off a substantial amount of time by developing laboratory tests that can be done directly when the isolate is identified. Sub-typing could be done in doctors' offices and the results electronically communicated to PulseNet.

We are developing a rapid plate test for shiga-toxin-producing *E. coli* to simultaneously look at sero-type and virulence factors—whether it has shiga toxin, what type of shiga toxin, and whether other toxins are produced. And new tests are coming into clinical laboratories—where one would go to have an illness diagnosed—that are rapid and don't necessarily need stool samples. Accordingly, we need a crash research program to change from PFGE to something else. Although PFGE works well, we need alternative, more-rapid options. Certain micro-arrays can generate data directly from the stool, and we are looking at the possibility of single-cell sequencing of DNA. Experts from around the world will confer with us in Atlanta in November, 2011, to discuss technologies that will help us get at this problem.

Exposure Information

To identify an outbreak, we have two sources of information. The germs that made people sick and the interviews about what the people ate and what they did. We have discussed the technology that helps us get at the issue of the causal bacterium. Referring back to Figure 14, determining who is a "case" and who is a "control" is helped by the microbiological methods and PulseNet, but determining who was exposed and who was not exposed comes from the interview. Mathematically, from the 2x2 table (Figure 14), they are equally important and we are starting to focus on refining this issue. We've

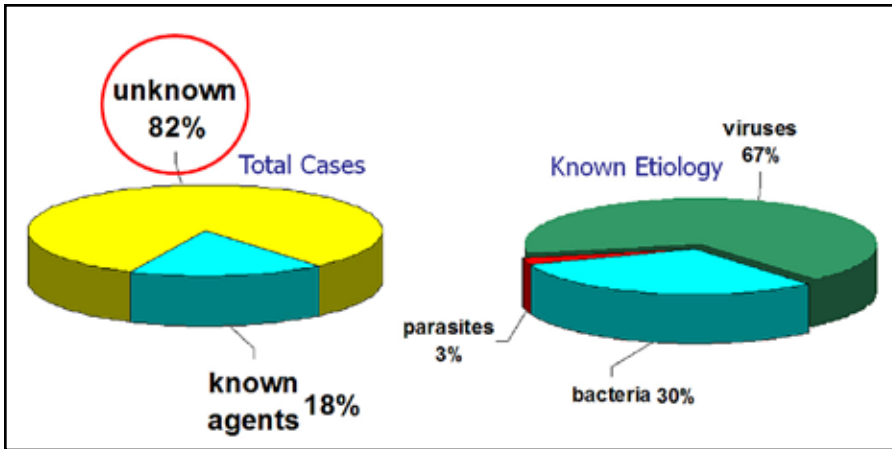


Figure 20. Etiology of acute gastroenteritis in the United States.

Pathogen Limitation

Surveillance is limited to pathogens that we know about. Data suggest that most foodborne pathogens fall into the “unknown” category (Figure 20). The bacteria that we track constitute only 30% of those that cause gastroenteritis. PulseNet and OutbreakNet activities cover only about 3.5% of all the cases of foodborne disease.

What interventions would be possible if we knew what pathogens cause the other 96.5% of cases? It’s hard to get at, but outbreaks of undetermined etiology present the possibility of finding out what’s actually making people sick. There are so many germs in the human gut, it would be very difficult to say which are causing disease. However, outbreaks provide a means for detecting pathogens and for triangulating them to the cause of the illness. When I worked in Minnesota, we did a national study of outbreaks of undetermined etiology and quickly found a number of new pathogens; this is worth pursuing nationally.

By employing metagenomic analyses, it is now possible to examine every single germ in the human gut of every single case. It’s not easy or cheap, but it’s possible and it will become less expensive. We then look at each germ as a risk factor for disease through our statistical analysis. As mentioned, these new methods are changing the way we view germs. Each time we sequence the genome of a germ, we find additional genes, with only about 3,000 genes stably present. It appears that, in nature, germs maintain only part of their genetic potential in their cells. The other genes are in the community and the cell can access different qualities as they need them. It’s an efficient way of evolving. This is exactly what we saw in the recent sprout-associated outbreak in Germany. The pathogen picked up new factors to help it adapt to a new niche. Not only will we be able to detect new pathogens, we will be able to detect the potential for outbreaks of disease like that in Germany by understanding not just the individual germ, but the whole system and its potential to cause harm to humans.

Scientists in approximately a hundred groups around the world are sequencing all of the strains from Germany and comparing those that cause hemolytic uremic syndrome to those that don't, with virulence studies in animals. It will be one of the most studied germs in history, thanks to new technology. There are now elegant new ways of looking for new types of germs. Using metagenomic techniques, Ian Lipkin found a putative cause for colony collapse disorder that affects honey bees. Handheld metagenomic devices now are coming onto the market that will allow us to identify new pathogens more quickly.



John Besser has served as the deputy chief of the Enteric Diseases Laboratory Branch at the Centers for Disease Control and Prevention since July 2009, where he is involved in national and global programs, to detect, characterize, and track enteric infectious diseases.

For two decades before joining the CDC, Dr. Besser managed the clinical laboratory at the Minnesota Department of Health where he was involved in the development of PulseNet and other innovative disease-surveillance programs. He received his BS, MS, and PhD degrees from the University of Minnesota.