

Introductory Paper

Molecular Subtyping and the Transformation of Public Health

ROBERT V. TAUXE

PUBLIC HEALTH SURVEILLANCE means monitoring specific infections diagnosed in a defined population. Surveillance is critical to define the burden of infections, to track the trends in their incidence as major control strategies are contemplated or implemented, and to detect anomalous surges or outbreaks. Outbreak detection and investigation are a critical part of the control strategy for many infections.

Prevention of foodborne infections is a complex art. Other than hepatitis A, there are no vaccines against them. Food can become contaminated by many possible pathways from the farm to the final consumer, that vary from pathogen to pathogen and over time. Preventing foodborne disease therefore depends in large part on engineering food production systems for safety. Detecting and investigating outbreaks is an important way to determine the pathways that are most problematic. The outbreak investigation is thus a fundamental part of the food safety process, as it is a major method by which new challenges and gaps in the food safety system are identified. Every foodborne outbreak is an opportunity to learn to improve our prevention methods and strategies, and drives the cycle of public health prevention (Fig. 1).

Each year, more than 1200 foodborne disease outbreaks are recognized, investigated, and reported by local and state health departments to the Centers for Disease Control and Prevention

(CDC, 2003). The foodborne outbreak is recognized when a group of people get the same disease that on investigation is linked to the same specific food. That detection event picks out the meaningful signal of the outbreak against the background noise of similar infections or illnesses that are sporadic and unconnected in nature. For *Salmonella*, serotyping has been an important first stage subtyping strategy that provided greater specificity than is provided by identifying the pathogen to the species level, and made possible the detection and investigation of many outbreaks (Olsen et al., 2001). Now routine molecular subtyping has further increased the specificity of surveillance, and the power of surveillance.

Molecular subtyping has had several major impacts on public health. First, it clearly increases the ability of surveillance to identify outbreaks that would otherwise be missed. This was well documented in Minnesota, during an evaluation of PulseNet for *Escherichia coli* O157:H7 (Bender et al., 1997). In 1994–5, the first years in which pulsed-field gel electrophoresis (PFGE) was routinely done on *E. coli* O157:H7 isolates in that state, 317 strains were typed, and 10 outbreaks were detected. Of these, four outbreaks, or 40% of the total, were thought unlikely to have been detected by standard surveillance. That means that the application of PulseNet within a single state was associated with a 67% increase in sensitivity. A similar

Foodborne and Diarrheal Diseases Branch, Division of Bacterial and Mycotic Diseases, National Center for Infectious Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia.

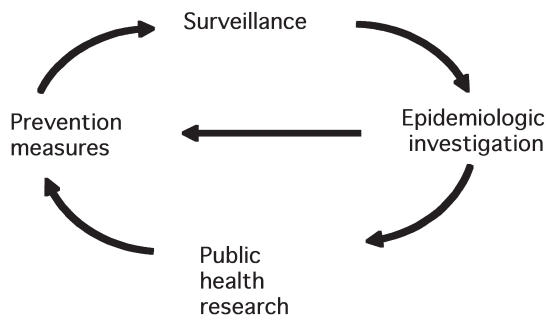


FIG. 1. The cycle of public health prevention.

assessment of routine pulsed-field gel electrophoresis (PFGE) for *Salmonella* Typhimurium showed a 33% increase in sensitivity (Bender et al., 2001). Expansion of the system beyond one state should lead to further increases of sensitivity, as it becomes possible to identify even more dispersed outbreaks. This general gain in sensitivity can be seen in the number of *E. coli* O157:H7 outbreaks reported nationwide before and after PulseNet was first deployed in 1996 (Rangel et al., 2005) (Fig. 2). This number was extremely low in the 1980s when few clinical labs cultured for the organism, and it accelerated sharply after the infection was made nationally notifiable in 1993 and then began to decline. A further quantum jump in reporting occurred around the time that PulseNet was deployed, rising from 17 outbreaks in 1997 to 44 outbreaks in 1998.

A second major impact has been on the specificity of the investigations themselves. The case definition used in an outbreak investigation is often as specific as possible, because the gain in information outweighs the loss in statistical power that follows the exclusion of some cases. Molecular subtyping increases the specificity of the definition, and hence of the findings. For example, in an outbreak of *E. coli* O157:H7 infections in Connecticut, the initial analysis suggested an association with a particular grocery store (Banatvala et al., 1996). The use of PFGE revealed a second unsuspected outbreak at the same time related to raw ground beef served at a wake in a neighboring town. After making PFGE part of the case definition, analysis now associated the store with illness, and illness could be specifically associated with food that had gone through the store's meat slicer. A recent *Campylobacter* outbreak illustrates this utility of PFGE in the

investigative phase, even though it did not play a role in the initial detection; the analysis is crisper with PFGE as part of the case definition, and the association with the vehicle clarified (Olsen et al., 2001; Fitzgerald et al., 2001).

For investigations of listeriosis, the specificity in definition now extends to the controls. Listeriosis primarily affects persons who are pregnant or seriously immunocompromised, and it can be difficult to find controls who are equally compromised. A new investigative strategy, first developed in France, and applied with recent success here, is to use other cases of listeriosis as the controls. This means interviewing all the listeriosis cases, and then using PFGE to divide them into outbreak cases and background controls (Graves et al., 2005; Olsen et al., 2005). Comparison of the reported exposures can reveal what is specific to the outbreak group. This method means that both cases and controls are defined by PFGE with equal specificity. By interviewing all cases of listeriosis in the country, and determining the PFGE patterns of all isolates, outbreak detection and investigation becomes a single standard process.

A third impact is that more outbreaks are detected and controlled at an earlier stage, and that fewer continue to a large size. The size of outbreaks is dropping, as a consequence of better surveillance, more direct intervention, and better prevention. For example, the average size of *E. coli* O157 outbreaks was 27 cases before 1992, and has decreased to <10 in recent years (Rangel et al., 2005). While it is difficult to know how much of this to attribute to PulseNet itself, faster detection and response is

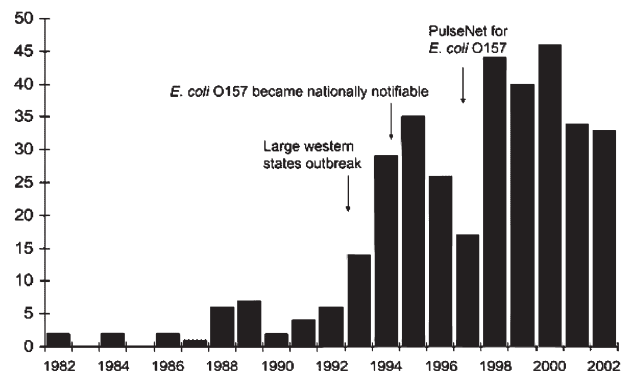


FIG. 2. *Escherichia coli* O157 outbreaks by year, 1982–2002 ($n = 350$) (Rangel et al., 2005).

likely to be part of the effect. This difference represents important evidence of prevention effectiveness. The cost of the program is easily recouped as a benefit to society. It has been estimated that preventing only five cases of *E. coli* O157:H7 infection annually pays the cost of PulseNet in a state (Elbasha et al., 2000).

A fourth major impact has been its use to develop clues about sources and to confirm a particular food as the culprit by subtyping pathogens from animals and from foods, collected as part of routine regulatory monitoring. The chance identification of strains of *Listeria monocytogenes* in sliced deli turkey meat that matched an ongoing cluster by PFGE was helpful early on in a recent listeriosis investigation (Olsen et al., 2005). Identifying a particular strain of *Salmonella* Typhimurium in a dead hamster led to a search of similar patterns nationwide, and led to the discovery of a highly dispersed outbreak related to these pets (CDC, 2005). Beyond hypothesis generation, finding a matching strain in an implicated food that is sampled because of a particular hypothesis can often provide overwhelmingly conclusive evidence of the source. Although regulatory action is often based on strong epidemiological information in the absence of sampling a food, the isolation of say, a matching isolate of *E. coli* O157 from ground beef associated with even a single case can lead to a meat recall and other regulatory actions. In the future, real time public health subtyping of strains from foods and animals, and comparison with the database of human isolates may provide earlier warning of contamination in the food chain.

The most profound impact has been on the kind of the outbreaks that have been detected. The nature of the outbreaks detected depends critically on the methods used. If outbreaks are only detected when concerned citizens, physicians or health care facilities report them, then only large and locally apparent outbreaks are likely to be found. These classic point source outbreaks affect a single group of people in a single town or city, following a single meal, with a substantial attack rate. Investigating this outbreak proceeds with local authority, and the foodhandling problems that are identified are most often local in scope. While important, these investigations may have few ramifications beyond the local scene (Fig. 3A).

The use of molecular subtyping to compare strains across many jurisdictions has revealed an entirely different kind of outbreak, in which a dispersed group of persons who do not know each other are affected at the same time with the same infecting organism in many states. In this scenario, no local listening post may perceive more than a few cases, and the local increase is often not apparent against the background of cases. Summed over many jurisdictions, the outbreak may in fact be very large, but dispersed. Investigating these dispersed scenario outbreaks requires the coordinated efforts of many health authorities acting in concert and pooling the information. Though difficult to detect and to investigate, the findings of these outbreaks can be of particular importance. The dispersion may well reflect a contamination event high in the food's chain of production, not just a problem in the final kitchen

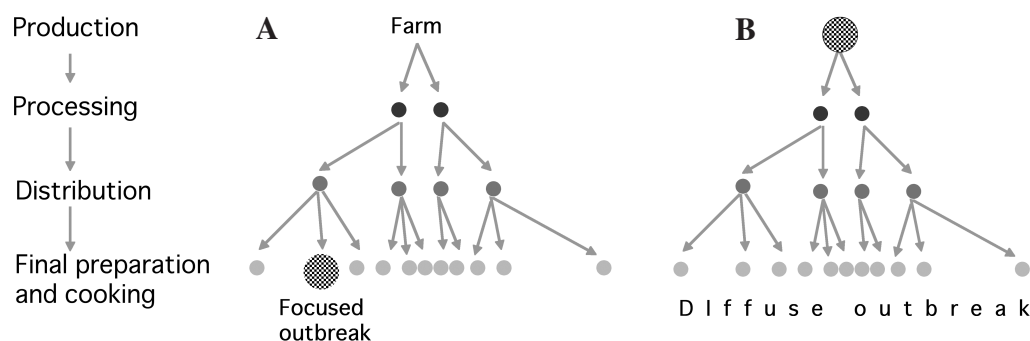


FIG. 3. The locus of the problem determines the shape of the outbreak. (A) A problem (indicated by checkerboard pattern) in one kitchen will usually lead to a focal outbreak. (B) A problem high in the food production chain can produce a highly dispersed outbreak.

(Fig. 3B). Identifying that event can instruct the industry and regulatory authorities about a flaw in the system that was previously unappreciated. Correcting it can lead to lasting and generalized protection.

This means that improved detection and investigation can serve to probe the safety of the food production system at several levels. These investigations, providing information about gaps in the food safety system, drive the cycle of prevention faster, and reduce the overall number of infections. The results of enhanced prevention can be seen in the recent declines in the incidence of listeriosis and *E. coli* O157 infections, the two pathogens tracked most intensively by PulseNet. Following the institution of PulseNet for *Listeria monocytogenes*, the number of outbreaks detected increased. Many of them were related to processed meats, focusing prevention efforts on that sector; incidence declined to an all time low of 2.7 per million in 2004, a drop of 40% since the baseline period 1996–1998 (CDC, 2005). The incidence of *E. coli* O157 infections began to decrease sharply after 2002, as the repeated investigations of PFGE clusters focused attention on more specific controls at the level of ground beef. By 2004, the incidence of *E. coli* O157 infections as measured in FoodNet had dropped 42% since the baseline period of 1996–1998, and was 0.9 per 100,000, below the health people goal for 2010 (CDC, 2005). It is doubtful that such progress would have been made without PulseNet.

We can also draw lessons of a more philosophic nature from PulseNet. The practical results confirm a central hypothesis of public health microbiology, that a single outbreak event is usually related to a single infecting strain. The philosophic underpinnings of subtype-based outbreak detection are rooted in the single strain postulate, and are validated by the empiric success of the strategy.

A second lesson is that enhancing surveillance creates a paradox, as it generally makes conditions look worse at first. The number of outbreaks detected increases, and some might conclude that the public health has deteriorated. It can take some years before the improved decision making that is the consequence of better surveillance and investigation drives the incidence of the disease downwards.

Third, the new dispersed outbreak blurs the line between individual sporadic cases and outbreaks. What appears to be a single sporadic case in one location, may in fact be part of an unrecognized cluster that affects many locations at once. Combining data across jurisdictions can bring this into focus. The new tack taken with listeriosis surveillance, described above, means we should consider all cases of listeriosis as potential members of some outbreak somewhere, and interview them accordingly. As PulseNet crosses international borders, the potential impact of investigating a single case, or a single outbreak, will be correspondingly magnified.

Many challenges remain to reach the full potential of subtype-based surveillance. The effort and resources needed to subtype in real time are not available in many jurisdictions, slowing the network, and delaying detection and investigations. The PFGE method includes some subjective judgments; DNA sequence-based methods may ultimately be more automated, and faster. Harmonized subtyping methods for more pathogens are needed. Already, applying sequence-based subtyping in network fashion to Norovirus and Hepatitis A infections has identified outbreaks and connections that were previously elusive (Fankhauser et al., 2002; Amon et al., 2005).

The great success of standardized subtyping in the public health laboratory has made information comparable across broad jurisdictions. A new challenge now is to do the same for the interviews of sporadic cases, that provide the complementary epidemiological information. The investigation of a dispersed cluster is unlikely to succeed without a joint effort based on the same interview questionnaire conducted in many jurisdictions at once. This need, combined with the blurring of the border between sporadic and outbreak-associated cases, means that even the single sporadic case of a high-priority infection may need to be investigated in a standardized way. This pressure to harmonize epidemiological methods means PulseNet is a profound affect on general public health practice.

Surveillance enhanced by molecular subtyping means that outbreaks are likely to be detected and investigated sooner, whether they

are naturally occurring or intentionally caused. Subtype-based surveillance, and the attendant general improvement in public health practice is critical to the preparedness of public health to meet the challenge of bioterror. Past intentional attacks with foodborne organisms were not announced as such, suggesting that an attack on the food supply with a foodborne pathogen will not initially be distinguishable from a natural event (Torok et al., 1997; Kolavic et al., 1997). That means that the public health investigation is likely to be the first response, just as the fire department responds to all fires, regardless of whether arson is suspected or not. Each of the 1200 foodborne outbreaks detected and investigated each year is a real exercise of the response system that will respond to an attack. Strengthening the existing systems for detecting and investigating foodborne outbreaks will mean that the public health is better protected against illnesses of all sorts, both natural and intentional.

REFERENCES

- Amon, J., R. Devasia, et al. 2005. Molecular epidemiology of foodborne hepatitis A outbreaks in the United States, 2003. *J. Infect. Dis.* 192:1323–1330.
- Banatvala, N., A.R. Magnano, et al. 1996. Meat grinders and molecular epidemiology—two supermarket outbreaks of *Escherichia coli* O157:H7 infection. *J. Infect. Dis.* 173:480–483.
- Bender, J., C. Hedberg, et al. 1997. Surveillance for *Escherichia coli* O157:H7 infections in Minnesota by molecular subtyping. *N. Engl. J. Med.* 337:388–394.
- Bender, J., C. Hedberg, et al. 2001. Use of molecular subtyping in surveillance for *Salmonella enterica* serotype Typhimurium. *N. Engl. J. Med.* 344:189–195.
- CDC (Centers for Disease Control and Prevention). 2003. Available at: www.cdc.gov/foodborneoutbreaks/outbreak_statistics/cdc_reported.htm. Accessed November 5, 2005.
- CDC. 2005. Outbreak of multidrug-resistant *Salmonella* Typhimurium associated with rodents purchased at retail pet stores—United States, December 2003–October 2004. *MMWR Morbid. Mortal. Wkly. Rep.* 54:429–433.
- CDC. 2005. Preliminary FoodNet data on the incidence of infection with pathogens transmitted commonly through food—10 sites, United States, 2004. *MMWR Morbid. Mortal. Wkly. Rep.* 54:352–356.
- Elbasha, E., T. Fitzsimmons, et al. 2000. Costs and benefits of a subtype-specific surveillance system for identifying *Escherichia coli* O157:H7 outbreaks. *Emerg. Infect. Dis.* 6:293–297.
- Fankhauser, R., S. Monroe, et al. 2002. Epidemiologic and molecular trends of “Norwalk-like viruses” associated with outbreaks of gastroenteritis in the United States. *J. Infect. Dis.* 186:1–7.
- Fitzgerald, C., L. Hessel, et al. 2001. Evaluation of methods for subtyping *Campylobacter jejuni* during an outbreak involving a food handler. *J. Clin. Microbiol.* 39:2386–2390.
- Graves, L., S. Hunter, et al. 2005. Microbiological aspects of the investigation that traced the 1998 outbreak of listeriosis in the United States to contaminated hot dogs and establishment of molecular subtyping-based surveillance for *Listeria monocytogenes* in the PulseNet network. *J. Clin. Microbiol.* 43:2350–2355.
- Kolavic, S., A. Kimura, et al. 1997. An outbreak of *Shigella dysenteriae* type 2 among laboratory workers due to intentional food contamination. *JAMA* 278:396–398.
- Olsen, S., R. Bishop, et al. 2001. The changing epidemiology of salmonella: trends in serotypes isolated from humans in the United States, 1987–1997. *J. Infect. Dis.* 183:753–761.
- Olsen, S., G. Hansen, et al. 2001. An outbreak of *Campylobacter jejuni* infections associated with food handler contamination: the use of pulsed-field gel electrophoresis. *J. Infect. Dis.* 183:164–167.
- Olsen, S., M. Patrick, et al. 2005. Multistate outbreak of *Listeria monocytogenes* infection linked to delicatessen turkey meat. *Clin. Infect. Dis.* 40:962–967.
- Rangel, J., P. Sparling, et al. 2005. Epidemiology of *Escherichia coli* O157:H7 outbreaks, United States, 1982–2002. *Emerg. Infect. Dis.* 11:603–609.
- Torok, T., R. Tauxe, et al. 1997. A large community outbreak of salmonellosis caused by intentional contamination of restaurant salad bars. *JAMA* 278:389–395.

Address reprint requests to:

Dr. Robert V. Tauxe
 Foodborne and Diarrheal Diseases Branch
 Division of Bacterial and Mycotic Diseases
 National Center for Infectious Diseases
 Centers for Disease Control and Prevention
 1600 Clifton Rd., Mailstop A-38
 Atlanta, GA 30333

E-mail: rvt1@cdc.gov