Published online: September 19, 2016

EMBO reports

Science & Society

Future challenges for tracking foodborne diseases

PulseNet, a 20-year-old US surveillance system for foodborne diseases, is expanding both globally and technologically

Efrain M Ribot & Kelley B Hise

Regardless of where they get their information from, Americans are very likely to learn almost instantly whenever there is an outbreak of bacterial pathogens—*Salmonella*, *Listeria*, or the "bad" *Escherichia coli*—from contaminated food products. This is a huge achievement and a great benefit for public health: The earlier this information reaches consumers, the less people will be affected and public health and other authorities have more time to identify and contain the source of the outbreak. However, despite its contribution to public health, most Americans are not aware that a little-known government program called "PulseNet USA" detects nearly all foodborne outbreaks of pathogenic bacteria. This is a bit odd because PulseNet has not only been very efficient in detecting foodborne disease but has thereby positively impacted public health and saved millions of dollars since it was founded 20 years ago. PulseNet is now undergoing profound changes as it both expands internationally to protect consumers in other countries and invests heavily—financially and scientifically—in new technologies such as next-generation sequencing (NGS) to further improve its capacity to detect food contaminations.

What is PulseNet

PulseNet is a national surveillance network based in Atlanta at the US Centers for Disease Control and Prevention to detect outbreaks of foodborne bacterial pathogens in real time [1,2]. Most of the detection itself is done at 83 accredited state, local, and federal laboratories that are connected with each other via an efficient communications network. PulseNet—both the center in Atlanta and individual laboratories—works closely with epidemiologists and other public health officials to investigate the source of an outbreak, establish appropriated public health measures, and assist federal agencies with improving the safety of the food supply. Simply stated, PulseNet's goal is to link information about people who have likely consumed the same contaminated food, even if they are in different parts of the country (Fig 1). The information from the partner laboratories is sent to the CDC where analysts constantly monitor the data to detect disease clusters at the national level.

"PulseNet has not only been very efficient in detecting foodborne disease but has thereby positively impacted public health and saved millions of dollars since it was founded 20 years ago"

Local clusters of foodborne illness, for instance after church potlucks or school picnics, are usually reported to and investigated by local public health agencies and are usually not very difficult to identify. Such clusters are typically caused by errors in food production—if a raw-egg mayonnaise has not been cooled properly or if raw chicken meat has been in contact with salad or vegetables. In contrast, PulseNet is particularly good at detecting outbreaks involving food that has been distributed across a large geographic area even if only a few cases are implicated.

PulseNet can help detect problems in food production and processing, and in the food distribution chain early on to help improve food safety. For example, after peanut butter was identified as the source of two major *Salmonella* outbreaks in 2006 and 2008, the food industry changed its production processes to make peanut butter products safer (http://www.cdc.gov/mmwr/PDF/wk/mm5621.pdf, http://www.cdc.gov/salmonella/2009/peanut-butter-2008-2009.html). Similar improvements were made for a wide range of other commodities, such as leafy green and vine vegetables, melons and fruits, poultry, beef, ready-to-eat foods, spices, tree nuts, and so on.

More than a hundred laboratories in 86 countries have adopted the PulseNet model. PulseNet International is comprised of seven regional networks—USA, Europe, Canada, Asia Pacific, Latin America and the Caribbean, Middle East, and Africa—that work together using standardized analysis methods for disease surveillance and outbreak response (Fig 2). Each region has a coordinating laboratory that is in charge of training and quality control, quality assurance programs, and organizes regular conference calls, meetings, and communication of epidemiological information. These laboratories use the same molecular subtyping protocols developed by PulseNet USA to ensure that laboratories compare the same data in a consistent manner and that information can be shared easily within and

Centers for Disease Control and Prevention, Atlanta, GA, USA. E-mail: eyr4@cdc.gov
DOI 10.15252/embr.201643128

Published online: September 19, 2016
EMBO reports

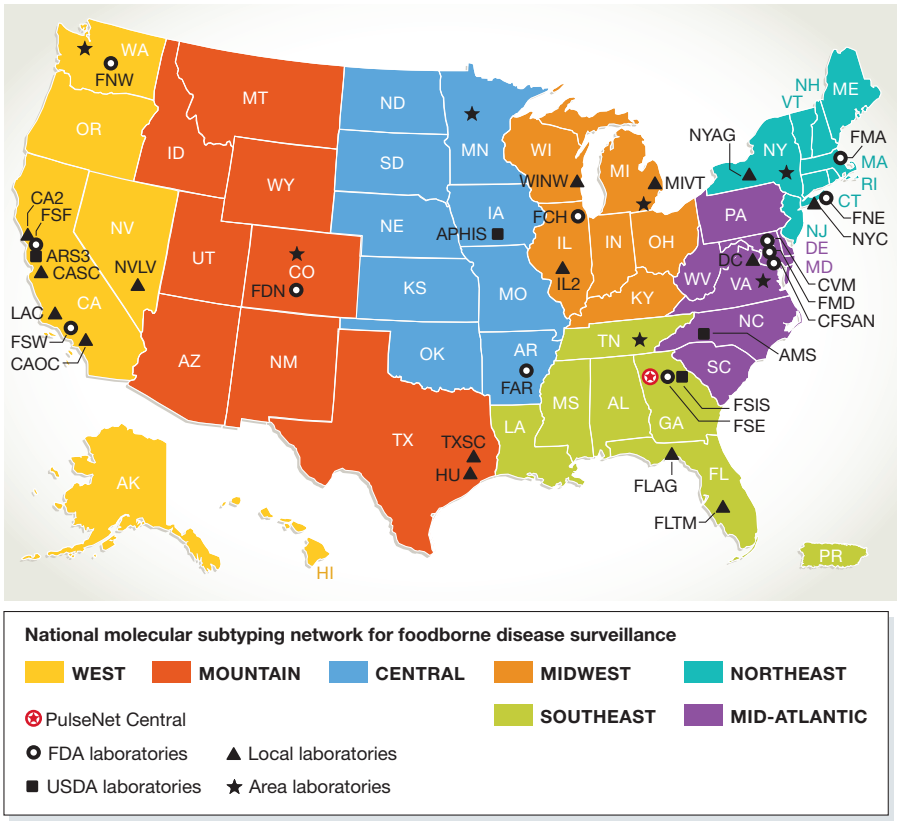

Figure 1. Map of PulseNet showing participating laboratories in the seven regions.

between regions. As with PulseNet USA, PulseNet International is improving its worldwide surveillance of foodborne bacterial pathogens by transitioning to whole genome sequencing (WGS) as a detection tool. The development and implementation of WGS methods and analytic processes in PulseNet USA will be influenced by the needs, ideas, and practical strategies arising from PulseNet laboratories and vice versa.

Earlier this year, PulseNet USA celebrated its 20th anniversary (Fig 3) and we believe that it is one of those "things that get better with age" such as wine, scotch, or well-crafted cheese. Most of these fine products are the result of careful planning and flawless execution and, in our mind, PulseNet fits the bill perfectly. Here, we try to describe the main ingredients of the recipe that has made PulseNet such a successful public health program.

Impact on foodborne diseases

When PulseNet was established, there were an estimated 76 million foodborne illnesses in the USA each year [3]. Today, foodborne illnesses affect one in six Americans annually, which translates into an estimated 48 million annual cases, 128,000 hospitalizations, and 3,000 deaths [4]. Globally, the number of diarrheal illnesses is around 2.8 billion [5,6]. These numbers should provide enough incentive to convince governments and public health officials that foodborne illness and food safety is a serious problem that requires attention. However, the impact of foodborne illnesses goes far beyond the number of individual cases. The US Department of Agriculture's Economic Research Service estimates that the cost of foodborne illness to the US economy is more than US $15 billion each year, of which approximately US$9 billion can be attributed to bacterial pathogens tracked by PulseNet [7]. These estimates do not take into consideration the cost to the food industry as a result of loss of consumer confidence, recalls, lawsuits, or other expenses incurred by local, state, and federal agencies as part of their outbreak response.

Another recent study measuring the impact and benefits the network provides estimated that PulseNet surveillance prevents more than 270,000 foodborne illnesses in the USA each year and, in the process, saves approximately US$500 million in medical costs and loss of productivity [8]. These cost savings are significant considering that the annual cost of running the network was approximately US$7 million dollars when the study was conducted. If there is a punch line, it is that "PulseNet saves lives and money".

PulseNet laboratories perform molecular subtyping analysis in as close to real time as possible on all bacterial pathogens that are currently tracked by the network. They submit subtyping data along with minimal metadata (state, species, serotype, source type, isolation date, received date, age, and sex of the patient) to the national databases at the CDC. In 2015, the 83 laboratories submitted more than 89,000 pulsed field gel electrophoresis (PFGE) profiles to the national databases representing close to 76,000 isolates (Figs 4 and 5). The total number of PFGE profiles in the PulseNet national databases reached 900,000 in 2015 and is projected to surpass the one million mark in 2017. Each year, approximately 1,500 clusters of indistinguishable or highly similar isolates are identified by state, local health, and federal laboratories, of which more than 275 represent multistate outbreaks. PulseNet scientists and epidemiologists track between 15 and 40 clusters of human illnesses per *week* compared to the approximately 20 clusters per *year* before PulseNet was established. As impressive as these numbers are, the profiles are only part of the picture: Additional epidemiological data are always required to define and solve outbreaks.

A decentralized network

From the start, it was clear that the early detection of outbreaks at the national level would only be possible if public health laboratories (PHLs) were capable of detecting "local" events in as close to real time as possible. This principle, not the technical tools, shaped the operational structure of the network. By decentralizing, subtyping activities surveillance could be carried out in "closer" proximity to the patient, reduce delays in transporting isolates to distant central laboratories, and communicate results to the epidemiologist closest to where the event may have occurred. Other laboratories are then able to access the information

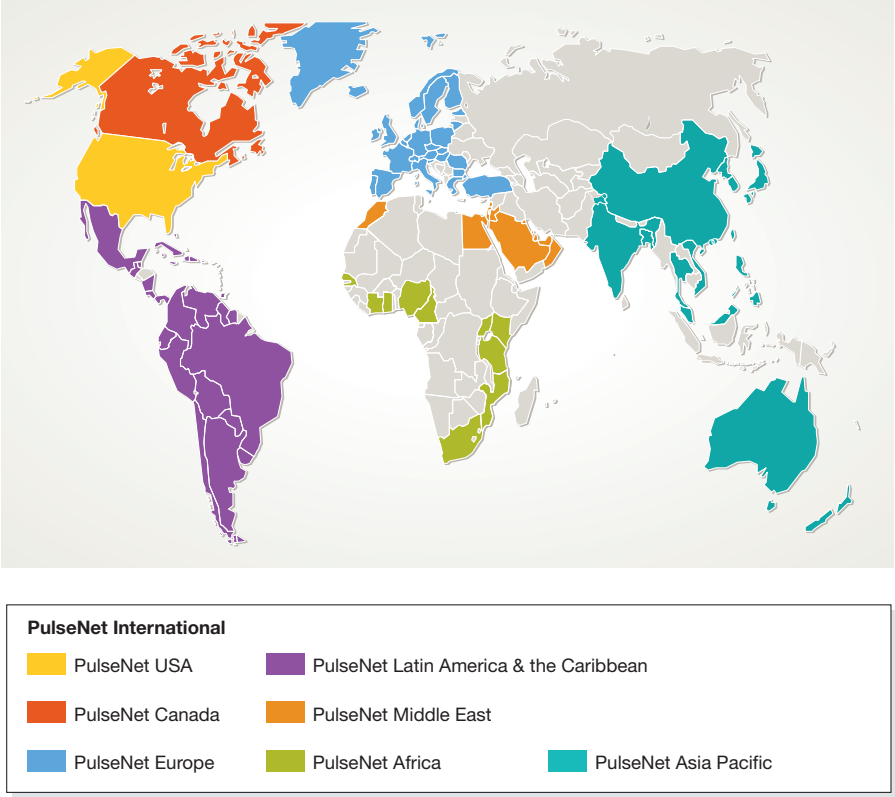

Figure 2. PulseNet International networks.

and provide support if necessary. Pursuing this model may seem obvious today, but 20 years ago it was a complex undertaking that required a clear and focused vision, not to mention a strategic plan. Most importantly, it required cooperation from public health partners including state and local public health laboratories, the Association of Public Health Laboratories (APHL), the Council of State and Territorial Epidemiologists (CSTE), the US Food and Drug Administration (FDA), the US Department of Agriculture (USDA), and international public health groups and collaborators [1,2]. It was clear that in order for PulseNet to succeed, the CDC could not act solitarily without the perspective and counsel of its partners.

Figure 3. PulseNet 20th Anniversary Logo.

To develop the original vision into an operational network, PulseNet had to implement a business model that was in line with the capacity and needs of the PHLs, including infrastructure, personnel, training, and data exchange and communication platforms for conducting real-time surveillance. A part of this was the establishment of the PulseNet Steering Committee that is comprised of members of the public health laboratories and institutions that participate in the network [2]. The steering committee not only provides guidance on what the network needs to focus on, but also brings a much broader and practical perspective on how to address challenges along with identifying new opportunities to improve PulseNet in the future. Other key factors include the establishment of standardized molecular tools for generating subtyping data including an external quality assurance system; a well-structured strategy for the exchange and analysis of data; a communication platform where cluster and outbreak information could be communicated in real time among networkparticipants and epidemiologists; and a sustainable, practical, and stable nomenclature system that allows for consistent naming of individual DNA profiles. Each of these functions takes on a higher level of complexity when viewed from the perspective of a network as opposed to a single laboratory.

PulseNet also requires robust standardized protocols that allow different laboratories with varying degrees of expertise and capacity to generate reproducible data. Training workshops are conducted every year to ensure that laboratorians are familiar with protocols and procedures. In addition, PulseNet has a quality assurance and quality control system [1,2]. This system is an active process that requires a significant investment of resources such as a stable core of subject matter experts and the proper infrastructure to rapidly address any issues—troubleshooting, surge capacity, retraining, and so on—that laboratories may run into.

Data management and communication

Data management and analysis is a process that is much more than the sum of its parts, which must be designed and executed with extreme care as errors could have a devastating impact on public health or the food industry. The informative value of the data increases as the number of entries in the national databases increases, but this is only true if the analytic processes are stable and the quality of data it relies on is high and consistent. It was therefore important to establish quality parameters that clearly define what data are acceptable to submit to the PulseNet national databases and, in the case of WGS, repositories such as the US National Center for Biotechnology Information's GenBank (http://www.ncbi.nlm.nih.gov/genbank/) or the European Molecular Biology Laboratory's European Nucleotides Archive (http://www.ebi.ac.uk/ena).

Open and effective communication is perhaps the most underestimated of all the PulseNet's activities. Any diagnostic method that produces epidemiologically relevant data falls short of its goal if the information is not communicated clearly and rapidly. Communicating the results or information related to laboratory data effectively at the scale of PulseNet, with more than 80,000 profiles being submitted to the national

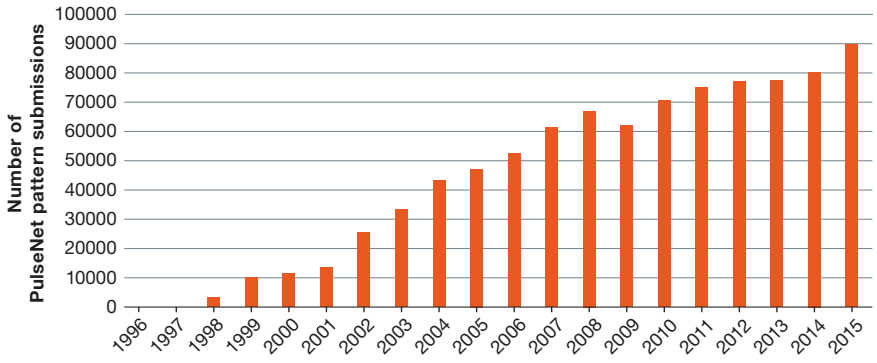

Figure 4. Total number of PFGE pattern submissions to PulseNet 1996–2015.

databases annually, requires a stable nomenclature system with unique and unambiguous naming of DNA fingerprints or sequence profiles. It also requires a systematic approach for assigning cluster names (groups of two or more highly related or indistinguishable isolates) and outbreak codes. As new subtyping approaches are being implemented, such as WGS, developing a stable nomenclature system will be even more challenging than the current approach, especially when one considers the sensitivity of WGS to detect differences in genome sequences.

Several international groups are working on the development of a WGS nomenclature system that reflects the needs of public health, academia, and the global sequencing community. From the surveillance point of view, the development of a practical nomenclature system that has more to do with how it is implemented than to the complexities of genome sequences (http://ecdc.europa.eu/en/publications/Publications/food-and-water-borne-diseases-next-generation-typing-methods.pdf). It means that PulseNet must implement processes, including nomenclature strategies, that are self-contained (not solely dependent on external entities) and self-sustained (stable funding sources), so new challenges and needs can be addressed locally and within expected timelines. In short, the implementation of a common language to communicate results is critical for surveillance and outbreak investigations.

Implementing WGS

The technological developments in nucleic acid sequencing represent both great challenges and opportunities for PulseNet. The advent of "bench top" sequencers coupled with a drop in the cost of sequencing has prompted PulseNet to replace the current gold-standard method, PFGE, with NGS to enhance its surveillance and investigation capacity. The versatility of NGS is already transforming clinical microbiology by replacing traditional reference characterization—serotype, pathotype identification, virulence profile, and antimicrobial resistance—with one efficient WGS-based workflow. WGS provides greater discriminatory power and precision than PFGE and, unlike PFGE, is also phylogenetically relevant; that is, isolates that are very similar likely share a common ancestor, which in epidemiological terms may translate to a common source. WGS clusters are therefore defined by a degree of similarity between organisms and outbreak scenarios, whereas PFGE clusters typically encompass one pattern. By focusing on one PFGE pattern, one excludes closely related strains, which has at least two effects: Smaller clusters might be missed, or it may take longer to generate a clear signal for larger clusters. In other words, WGS allows for the "fine tuning" of case definitions, which makes it possible to detect and solve clusters faster and with

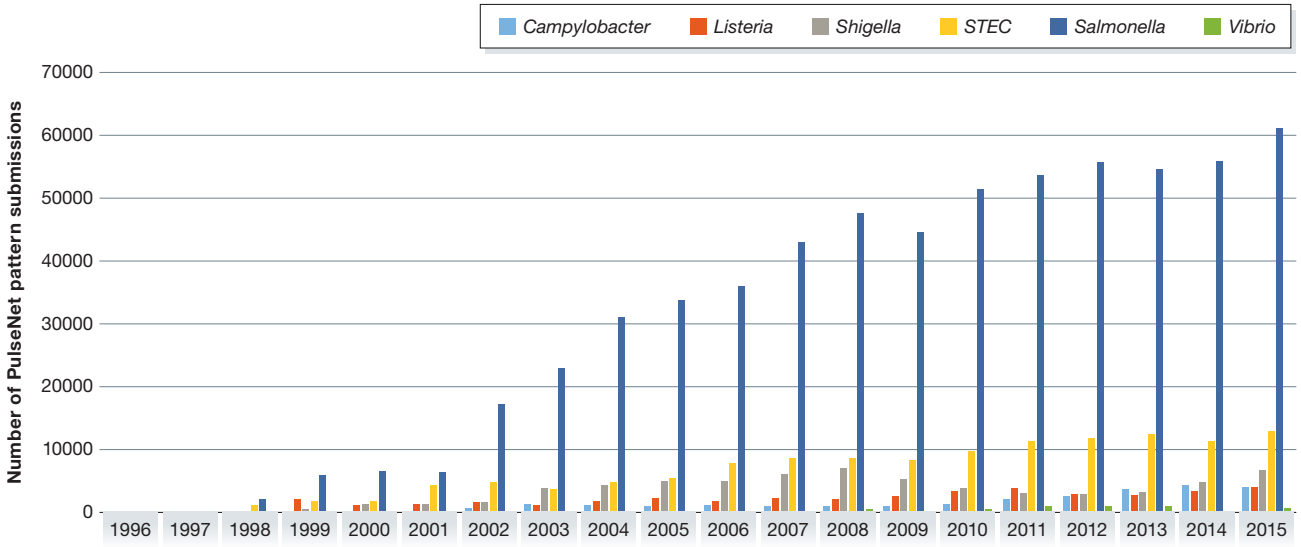

Figure 5. PulseNet pattern submissions by organism 1996–2015.

fewer cases. The implementation of WGS is already transforming the way foodborne outbreak investigations are carried out in the USA.

The implementation of radically different methodologies into an established network like PulseNet takes time and is costly, and any misstep could cripple the network's ability to conduct real-time surveillance. That is, the methods used by participating laboratories must produce the same type of standardized data and use the same analytical approaches. Even in a few years when WGS becomes the new standard for subtyping, PulseNet Central will continue to maintain PFGE capacity and expertise to provide support for late adopters of the technology. Other issues to consider when adopting a new technology include its impact on processes such as certification and proficiency testing, or external quality assessment requirements the laboratories may have to comply with state or federal regulations.

The transition from PFGE to WGS will not only transform how PulseNet laboratories perform their analytical work, but will also provide the basis for PHLs to expand this capacity beyond the realm of foodborne pathogens. CDC and PHLs are already moving toward the implementation of NGS approaches in areas outside food safety in parallel with the transformation of PulseNet. Full WGS implementation is well underway in PHLs under their own initiative with support provided by the CDC through Food Safety, Antimicrobial Resistance, and Advanced Molecular Detection (AMD) initiatives. These initiatives currently provide the resources for developing the analytical tools and the infrastructure in state PHLs and are expected to lead to collaborations and increased efficiency of public health activities involving bacterial, viral, and parasitic agents. In addition, combining the power of WGS with data generated by other surveillance and monitoring systems, such as the Foodborne Diseases Active Surveillance Network (FoodNet) and the National Antimicrobial Resistance Monitoring System (NARMS), it will be possible to establish clear links between sporadic cases of foodborne diseases and the source of infections in ways that PFGE never could.

The switch to WGS began more than 5 years ago with the sequencing of the *Vibrio cholerae* outbreak strain from Haiti, but it was not until 2013 that efforts kicked into a higher gear with the launch of the WGS *Listeria* pilot study, which, in combination with other efforts at the CDC and elsewhere, has led to the development of robust workflows for generating and analyzing WGS data. PulseNet teamed up with the FDA-sponsored GenomeTrakr network, which provided sequence data for isolates from food or food production environments. Interpretation of WGS *Listeria* data for surveillance would have been impossible without the epidemiological context from the *Listeria* Initiative, an enhanced surveillance system established by the CDC in 2005. Its main objective is to conduct patient interviews of laboratory-confirmed cases to reduce the time it takes for implementing public health interventions before clusters or outbreaks are detected (http://www.cdc.gov/nationalsurveillance/listeria_surveillance.html).

By combining the epidemiological data from the *Listeria* Initiative with WGS data generated by state public health and food regulatory laboratories, we can get a better understanding of how sequence data can be used and interpreted by laboratorians, public health scientist, and epidemiologists. In fact, this Initiative was one of the main reasons why *Listeria* was used as the pilot organism in the implementation of WGS in PulseNet. The outcome of this study has been extremely positive and confirms that WGS analysis provides more precision and better discrimination compared with PFGE and in turn has led to the discovery of more outbreaks, better identification of the source, and better detection of smaller clusters/outbreaks (Figs 6 and 7).

Future challenges

PulseNet is faced with another challenge that, if left unattended, has the potential to cripple laboratory-based surveillance in the USA and abroad: the loss of isolates as clinical laboratories are implementing new culture-independent diagnostic tests (CIDTs) [9,10]. CIDTs provide physicians and patients with an advantage over culture-based methods because they are faster at identifying the cause of illness, which means that treatment can be started earlier. They present new opportunities for tracking and controlling diseases, especially those diseases for which practical diagnostic tests were not formerly available, such as enterotoxigenic *E. coli* or ETEC. However, given that the adoption of CIDTs is expected to continue at an accelerated pace in the near future, less isolates may be available to PHLs and the CDC, compromising their ability to detect and investigate dispersed foodborne disease outbreaks. Other side effects include diminished capacity to closely monitor the efficacy of control measures, which in turn will likely weaken the safety of the food supply. A number of strategies have

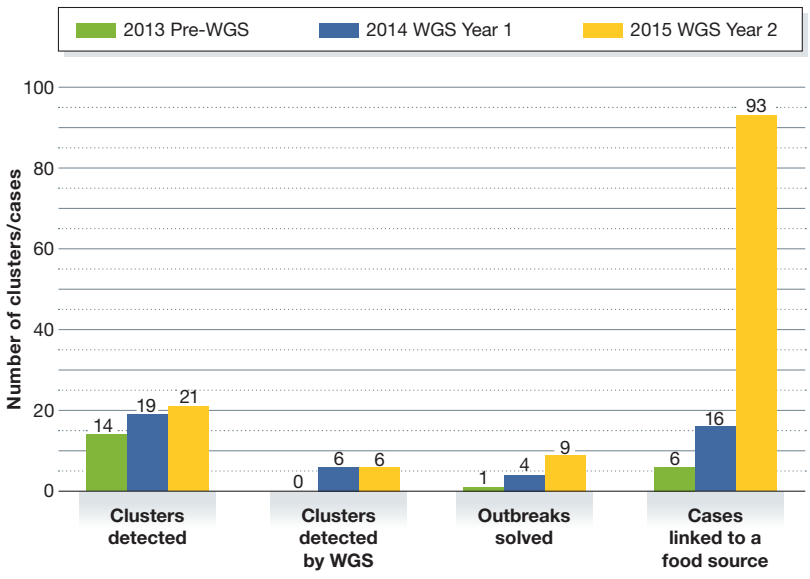

Figure 6. Listeria surveillance pre and post implementation of WGS.

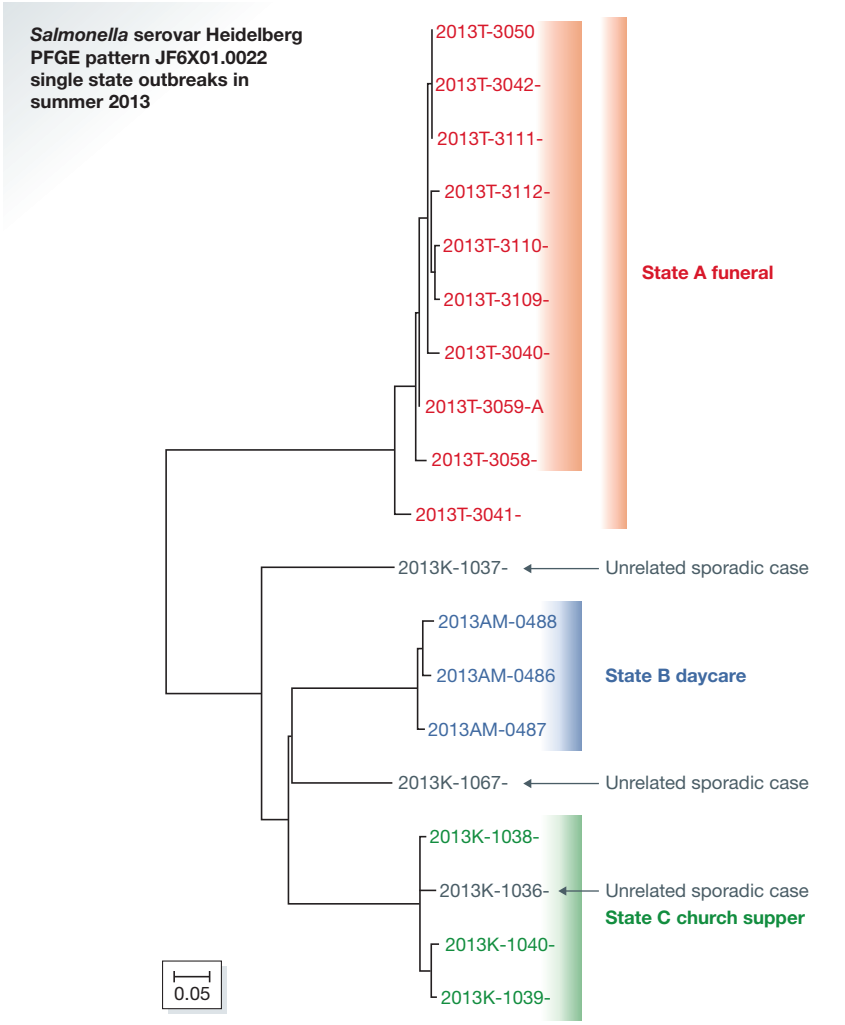

Figure 7. Comparison of *Salmonella* serovar Heidelberg isolates from outbreaks in three different states that were indistinguishable by PFGE. The higher discriminatory power of WGS proofed to be critical in clarifying the epidemiological picture.

been considered while addressing the impact of CIDTs on PulseNet in the short term and long term. One focuses on educating CIDT device manufacturers about the impact their tests are likely to have on public health along with suggestions of how to ensure the survival of the pathogen after the specimen is processed. This approach is currently being used by several state laboratories in the USA, but, because of its cost, it is likely not sustainable in the long term.

Other approaches used to address the CIDT issue focus on the development and implementation of molecular tools for the identification and subtyping of pathogens directly from complex matrices such as stool samples without culture. A number of projects are being conducted at the CDC to combine the analytic power of bioinformatics with the empirical validation that only laboratory testing can provide. For instance, PulseNet, in collaboration with other groups, is using bioinformatics tools to interrogate whole genome sequences in order to identify targets that are unique and conserved for organisms such as Shiga-toxin-producing *Escherichia coli* (STEC) and *Salmonella enterica*.

Another, more complex approach uses metagenomics to recover and analyze genetic material from environmental or complex samples. Metagenomics opens a new world of possibilities for public health and food safety. In addition to helping us mitigate the impact of CIDTs, it also has incredible potential to help us recognize etiologies of illnesses that are still unknown. It has the potential to identify multiple pathogens at once, including ones that we might not be looking for without prior knowledge of their existence in the sample. One of the most significant challenges of this approach is how to differentiate between pathogens that are genetically similar to the commensal flora such as STEC and commensal *E. coli*. It will require coming up with effective ways to eliminate unwanted genetic material, maximize the harvest of desired genetic target samples, and link all the markers originating from a single genome and differentiate these from other organisms with similar genetic backgrounds.

A new decade has begun for PulseNet as the network replaces a useful but blunt tool, PFGE, with a more surgical instrument, WGS, to identify clusters of foodborne bacterial illness with a precision not possible before. In addition, the role PulseNet plays in public health will continue to change. Perhaps the biggest transformation will occur as PulseNet continues to work with international partners to implement tools, practices, and nomenclature schemes that take into consideration the needs and realities of PulseNet and individual countries or regions. To this end, one of the most exciting changes PulseNet is experiencing is a shift from a closed system, in which the data are managed and contained in a national repository accessible only to properly accredited network members, to a model where the WGS data are deposited in an open system in as close to real time as possible. The new approach represents a quantum leap forward toward achieving open data exchange and true globalization of WGS-based surveillance of the pathogens tracked by PulseNet. The expected outcome of this transformational project is faster and more efficient national and international surveillance, better connectivity with the global community, and a safer food supply. And all of this provides additional evidence that PulseNet is, indeed, still getting better with age.

Acknowledgements

The authors thank Ms. Brittany Behm and Dr. Jenifer Concepción-Acevedo for their helpful comments on the manuscript.

Conflict of interest

The authors declare that they have no conflict of interests.

References

1. Swaminathan B, Barrett TJ, Hunter SB, Tauxe RV (2001) PulseNet: the molecular subtyping network for foodborne bacterial disease surveillance, United States. *Emerg Infect Dis* 7: 382 – 389
2. Gerner-Smidt P, Hise K, Kincaid J, Hunter S, Rolando S, Hyytiä-Trees E, Ribot EM, Swaminathan B (2006) PulseNet USA: a five-year update. *Foodborne Pathog Dis* 3: 9 – 19
3. Mead PS, Slutsker L, Dietz V, McCaig LF, Bresse JS, Shapiro C, Griffin PM, Tauxe RV (1990) Food-Related Illness and Death in the United States. *Emerg Infect Dis* 5: 607 – 625
4. Scallan E, Hoekstra RM, Angulo FJ, Tauxe RV, Widdowson M, Roy SL, Jones JL, Griffin PM (2011) Foodborne Illness Acquired in the United States—Major Pathogens. *Emerg Infect Dis* 17: 7 – 15
5. Majowicz SE, Musto J, Scallan E, Angulo FJ, Kirk M, O'Brien SJ, Jones TF, Fazil A, Hoekstra RM, and for the International Collaboration on Enteric Disease "Burden of Illness" Studies (2010) The Global Burden of Nontyphoidal *Salmonella* Gastroenteritis. *Clin Infect Dis* 50: 882 – 889
6. Scallan E, Kirk M, Griffin PM (2013) Estimates of disease burden associated with contaminated food in the United States and globally. In *Foodborne Infections and Intoxications*, Morris GJ, Potter M (eds) pp 3 – 18, San Diego: Elsevier BV
7. Flynn D (2014) U.S. Foodborne Illnesses Cost More Than $15.6 Billion Annually. Food Safety News, October 9, 2014
8. Scharff RL, Besser J, Sharp DJ, Jones TF, Gerner-Smidt P, Hedberg CW (2016) An economic evaluation of PulseNet, a network for foodborne disease surveillance. *Am J Prev Med* 50(5 Suppl 1): S66 – S73
9. Cronquist AB, Mody RK, Atkinson R, Besser J, D'Angelo MT, Hurd S, Robinson T, Nicholson C, Mahon BE (2012) Impacts of culture-independent diagnostic practices on public health surveillance for bacterial enteric pathogens. *Clin Infect Dis* 54: S432 – S439
10. Jones TF, Gerner-Smidt P (2012) Nonculture diagnostic tests for enteric diseases. *Emerg Infect Dis* 18: 513 – 514