

## Epidemiologic profiling: evaluating foodborne outbreaks for which no pathogen was isolated by routine laboratory testing: United States, 1982–9

J. A. HALL<sup>1</sup>\*, J. S. GOULDING<sup>2</sup>, N. H. BEAN<sup>2</sup>, R. V. TAUXE<sup>2</sup>  
AND C. W. HEDBERG<sup>1</sup>

<sup>1</sup> Minnesota Department of Health, Minneapolis, MN, USA

<sup>2</sup> Centers for Disease Control and Prevention, Atlanta, GA, USA

(Accepted 29 June 2001)

### SUMMARY

The objective was to evaluate foodborne outbreaks of undetermined aetiology by comparing them to pathogen-specific epidemiologic profiles of laboratory-confirmed foodborne outbreaks. National foodborne outbreak data reported to CDC during 1982–9 were categorized by clinico-epidemiologic profiles based on incubation, duration, percent vomiting, fever and vomiting to fever ratio. From the pathogen-specific profiles, five syndromes were developed: a vomiting-toxin syndrome resembling *Bacillus cereus* and *Staphylococcus aureus*; a diarrhoea-toxin syndrome characteristic of *Clostridium perfringens*, a diarrhoeagenic *Escherichia coli* syndrome, a Norwalk-like virus syndrome, and a salmonella like syndrome. Of 712 outbreaks, 624 (87.6%) matched one of five syndromes; 340 (47.8%) matched the Norwalk-like syndrome and 83 (11.7%) matched the salmonella-like syndrome. After combining information on known pathogens and epidemiologic profiles, only 88 (12.4%) outbreaks remained unclassified. Norwalk-like virus outbreaks appear as common as salmonella-like outbreaks. We conclude that profiling can help classify outbreaks, guide investigations and direct laboratory testing to help detect new and emerging pathogens.

### INTRODUCTION

In the most recently published summary of surveillance for foodborne disease outbreaks in the United States, the Centers for Disease Control and Prevention (CDC) reported the occurrence of 2751 outbreaks during 1993–7. These outbreaks were defined by the occurrence of a similar illness among two or more persons, which an investigation linked to consumption of a common meal or food item. The reported outbreaks involved 29 known pathogens. A reported 86058 persons became ill, of which 29 died [1]. These reported figures represent a small fraction of all foodborne illness. Recently, CDC estimated that 76 million cases of foodborne illness and 5000 deaths occur each year in the United States [2]. With

increased international travel, the growing demand for imported fresh fruits and vegetables, the threat of antibiotic resistance and the emergence and recognition of new pathogens, foodborne disease continues to be a substantial public health and economic burden in the United States with annual cost estimates reaching five billion dollars [3, 4].

CDC's national foodborne disease surveillance categorizes outbreaks by laboratory-confirmed aetiological agent [1]. For example, between 1993 and 1997, 878 (32%) of the 2751 outbreaks reported to CDC had a laboratory-confirmed aetiology. Of those, 75% were caused by a bacterial agent, 17% were chemical, 6% were viral and 2% were parasitic in origin. Thus, 68 percent of foodborne outbreaks were categorized as having an undetermined aetiology [1]. Outbreaks may be classified as undetermined aetiology because (1) an appropriate specimen for testing was not

\* Author for correspondence: 12125 Technology Drive, MN002-0258, Eden Prairie, MN 55344 USA.

collected or (2) the specimen was negative for all pathogens tested for in the laboratory.

While results of microbiologic testing may definitively identify the aetiologic agent, there are many foodborne pathogens that are not routinely tested for in clinical or reference laboratories in the United States. For example, only a few state public health departments currently test for Norwalk-like caliciviruses. Outbreaks tend to have pathogen-specific patterns of symptoms and other epidemiologic characteristics that are amenable to epidemiologic analysis [5, 6]. The use of clinico-epidemiologic profiles can help to define better the potential public health importance of categories of foodborne pathogens. In addition, before results of routine laboratory tests are available, epidemiologic profiles can help guide decisions about the outbreak investigation and potential control strategies, and indicate the need for specialized laboratory tests.

The 1997 Food Safety Initiative called for the development of an early warning system to detect foodborne outbreaks and identify emerging foodborne pathogens. Coupling of laboratory testing with epidemiologic analysis of an outbreak to identify the aetiologic agent and source of an outbreak is critical to the successful investigation of any outbreak. Using available information efficiently could improve the responsiveness of the food safety system.

The purpose of this study was to evaluate the use of epidemiologic profiles to classify outbreaks originally classified as 'undetermined aetiology' based on the absence of a laboratory-confirmed agent.

## METHODS

### Design

This was a retrospective analysis of existing epidemiologic and laboratory data gathered by state and local health departments during the investigation of foodborne outbreaks reported to CDC. These data represent all foodborne outbreaks that were reported to CDC between 1982 and 1989. This time period was chosen because the data had been collected and reported to CDC in a consistent manner and were available in an electronic format. The variables in the data set included incubation period, duration of illness and the number of persons reporting diarrhoea, cramps, vomiting, nausea and fever. Also provided were the numbers of persons ill, the number hospitalized, the number of deaths and the state, month and year in which an outbreak occurred.

### Data set used for analysis

Initially, the data were divided into two categories: (1) outbreaks with a laboratory-confirmed aetiologic agent and (2) outbreaks in which no agent was confirmed (Fig. 1). Each group was subsequently pared down until the final data set used in the analysis included only those outbreaks with (1) summarized clinical data from at least five case histories, and (2) complete data reported for the following six fields: incubation, duration, number reporting diarrhoea, number reporting vomiting, number reporting fever and the number of histories taken during the outbreak.

### Epidemiologic profiles

The pathogens responsible for a large proportion of outbreaks during the 8 year period were selected for closer study. The eight pathogens chosen for further examination included *Bacillus cereus*, *Campylobacter* spp., *Clostridium perfringens*, *Escherichia coli*, Norwalk virus, *Staphylococcus aureus*, *Salmonella* spp. and *Shigella* spp. The distinction between ETEC and STEC *E. coli* outbreaks was not possible with the level of detail provided. All laboratory-confirmed outbreaks for these eight pathogens were characterized by incubation, duration, percentage of cases reporting vomiting, percentage reporting fever, percentage reporting diarrhoea, percentage reporting cramps and the vomiting to fever ratio. The range of values between the first and third quartile around the median was used for each of the seven variables to create a clinico-epidemiologic profile for each of the eight pathogens. The profiles were created with the help of SPSS statistical software version 6.1.3.

Because of the small number of confirmed Norwalk virus outbreaks available for generating epidemiologic profiles, and because epidemiologic criteria for evaluating outbreaks of viral gastroenteritis have been published, the published criteria were used to supplement the Norwalk-like virus profile created from the data set. Kaplan's criteria for outbreaks of Norwalk-like viral gastroenteritis include; incubation period between 24 and 48 h, duration of illness between 12 and 60 h, and  $\geq 50\%$  cases reporting vomiting [5]. Kaplan's criteria were modified to include  $\geq 50\%$  of cases reporting vomiting or a vomiting to fever ratio greater than one [7].

The eight pathogen profiles were further condensed into five syndromes. The pathogen profiles for *Bacillus cereus* and *Staphylococcus aureus* overlapped. Out-

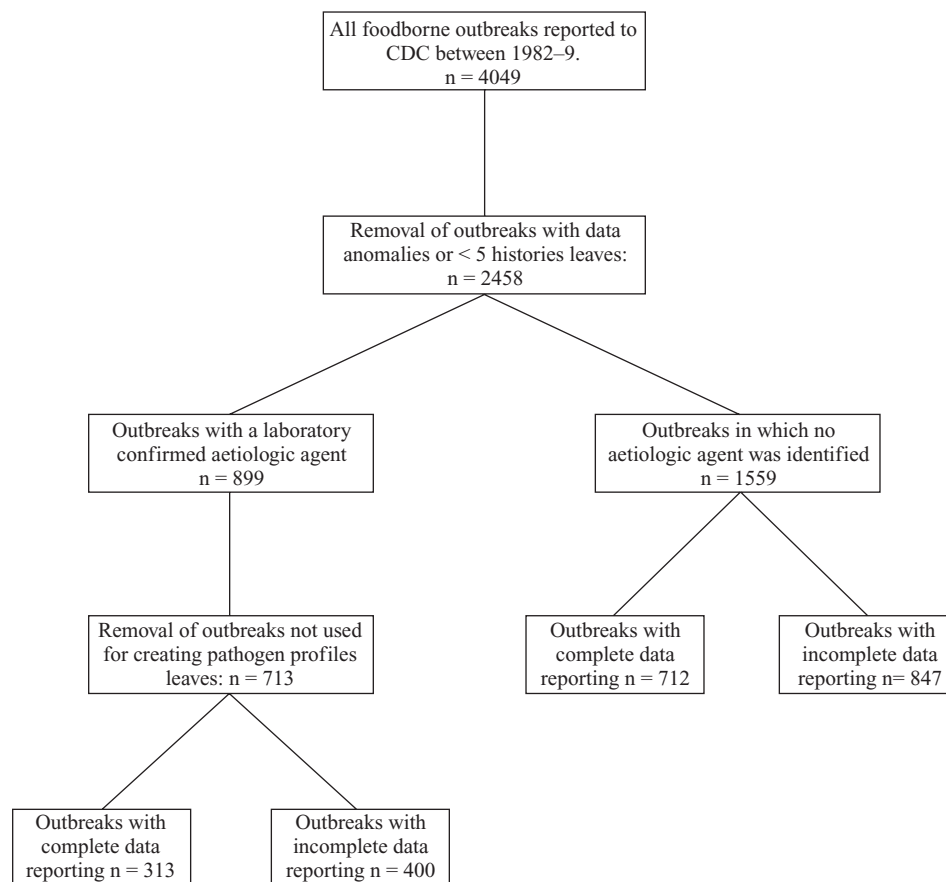


Fig. 1. Use of foodborne outbreak data. United States: 1982-9.

Table 1. *Distinct pathogen syndromes*

Syndrome	Incubation (h)	Duration (h)	Vomiting (%)	Fever (%)	Vom/Fev ratio*
Vomiting-toxin	1.5-9.5	6.3-24	50-100	0-28	0.4-3
Diarrhea-toxin	10-13.0	12.0-24	3.6-20	2.3-10	0.40-1.3
<i>E. coli</i>	48-120	104-185	3.1-37	13-25.3	0.25-1.1
Norwalk-like virus	34.5-38.5	33-47	54-70.2	37-63	0.70-1.7
Salmonella-like	18.0-88.5	63-144	8.9-51	31-81	0.20-1.0

\* Ratio of proportion vomiting to proportion with fever.

breaks of both are caused by pre-formed toxins, have very short incubation periods and a large percentage of case experience vomiting, but not fever. As a result, these two profiles were combined to produce the vomiting-toxin syndrome. Similarly, the profiles for *Salmonella* spp., *Shigella* spp. and *Campylobacter* spp. overlapped. Outbreaks of all three have similar incubation periods, and cases generally experience more fever than vomiting [8]. Therefore, these three profiles were combined to form the salmonella like syndrome. This left five distinct syndromes: (1) the vomiting-toxin syndrome, (2) the diarrhoea-toxin syndrome, (3) the Norwalk-like virus syndrome,

(4) the salmonella-like syndrome, and (5) the *E. coli* syndrome (Table 1). The cramps and diarrhoea variables were subsequently dropped as they provided little distinction between pathogens. Outbreaks from each syndrome reported upwards of 80-100% of cases with cramps and diarrhoea.

#### Application of pathogen syndromes to outbreaks of confirmed aetiology

A series of nine programmes were created to determine how well laboratory-confirmed outbreaks fit the five syndromes that were created. Each programme was

run against the pool of laboratory-confirmed outbreaks to see which programme could most often classify the outbreaks into their correct pathogen syndrome. Programme 1 matched an outbreak to a syndrome based on incubation period. For example, using programme 1, an outbreak with an incubation period between 1.5 and 9.5 h would be counted as a match for the vomiting-toxin syndrome. Programme 2 matched on incubation and either vomiting or fever. Programme 3 matched on vomiting and fever without incubation. Programme 4 matched on incubation, vomiting and fever. Programme 5 matched on all seven variables. Therefore, using programme 5, an outbreak would not be counted as a match for any syndrome unless all seven variables matched a particular syndrome. This was the most rigorous approach. For programmes 6–9, a different approach was used. The number of variable matches was summed. The pathogen syndrome with the greatest number of variable matches was selected as the correct syndrome. Programme 6 considered the sum of matches on vomiting, fever and the vomiting to fever ratio. Programme 7 summed matches on all variables except incubation period. Programme 8 summed matches on all variables except diarrhoea. Finally, programme 9 summed matches on all seven variables. For example, using programme 9, an outbreak with 50–100% of cases reporting vomiting, zero to 28% of cases reporting fever, and an incubation period between 1.5 and 9.5 h would receive a total of three matches out of a possible seven for the vomiting-toxin syndrome. The final step was to take the five pathogen syndromes and the programme best able to classify outbreaks into their correct pathogen syndrome and apply the programme to the group of outbreaks in which an aetiologic agent was not laboratory-confirmed.

## RESULTS

Between 1982 and 1989, 4049 outbreaks were reported to CDC's national foodborne disease outbreak surveillance system. Of these, 1049 (25.9%) outbreaks included at least five case histories and had complete information for the six variables of interest. A laboratory-confirmed aetiologic agent was identified in 337 outbreaks (32.1%) and 712 outbreaks (67.9%) were classified as having an undetermined aetiology. Among outbreaks with a laboratory-confirmed aetiology, *Salmonella* spp. was the most frequently reported pathogen, accounting for 196 (58.2%)

confirmed outbreaks. Combined, the eight pathogens chosen for further analysis ( $n = 313$ ) accounted for 92.9% of confirmed outbreaks.

The programme best able to match an outbreak to the correct pathogen syndrome was the programme that summed matches on all variables excluding diarrhoea. Because the cramps and diarrhoea variables were subsequently dropped from analysis, a perfect match was five out of five variables. This programme correctly classified 225 (71.9%) laboratory-confirmed outbreaks and incorrectly classified 44 (14.1%) laboratory-confirmed outbreaks. The remaining 44 (14.1%) laboratory-confirmed outbreaks matched more than one pathogen syndrome (Table 2). Four outbreaks that matched more than one syndrome, a tiebreaker programme that used incubation period was considered. While this programme was able to classify an additional 9.0% of outbreaks, the margin of error outweighed its benefits and the programme was therefore dropped.

The incubation periods of the vomiting-toxin and diarrhoea-toxin syndrome were not significantly different. This was expected as outbreaks caused by *B. cereus* may have a dual clinical presentation. One illness is characterized by a short incubation period with more vomiting than fever. The other is characterized by a longer incubation period and more diarrhoea [8]. Outbreaks with the longer incubation period would most likely fit into the diarrhoea-toxin syndrome. For the duration variable, the vomiting-toxin syndrome and Norwalk-like syndrome were not significantly different. For the fever variable, there was no difference between the *E. coli* syndrome and the vomiting-toxin and salmonella-like syndromes. None of the confirmed *B. cereus*, *C. perfringens*, *Campylobacter* spp., *E. coli*, *S. aureus*, or *Shigella* spp. outbreaks matched the published Norwalk-like virus criteria used in the Norwalk syndrome. However, 3.6% of salmonella spp. outbreaks did. Thus, the specificity of the Norwalk-like syndrome was 98% among the outbreaks with laboratory-confirmed agents. Five (71.4%) of seven laboratory-confirmed Norwalk virus outbreaks matched the Norwalk-like syndrome.

When the outbreak-matching programme, supplemented with the published criteria for outbreaks of Norwalk-like virus, was applied to the undetermined outbreaks, 340 (47.8%) outbreaks fitted the Norwalk-like syndrome, 155 (21.8%) outbreaks fitted the vomiting-toxin syndrome, 83 (11.7%) outbreaks fitted the salmonella-like syndrome, 38 (5.3%) outbreaks

Table 2. Classification of all complete outbreaks with  $\geq 5$  histories: 1982–9, United States

Pathogen	% Outbreaks assigned to each syndrome						Total
	Vomiting toxin	Diarrhoea toxin	<i>E. coli</i>	Norwalk-like virus	Salmonella-like	Ties	
<i>B. cereus</i>	5 (71.4)	2 (28.6)	0 (0)	0 (0)	0 (0)	0 (0)	7
<i>S. aureus</i>	35 (100)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	35
<i>C. perfringens</i>	9 (37.5)	6 (25.0)	0 (0)	0 (0)	0 (0)	9 (37.5)	24
<i>Campylobacter</i> spp.	0 (0)	0 (0)	2 (12.5)	0 (0)	9 (56.3)	5 (31.3)	16
<i>E. coli</i>	1 (20.0)	0 (0)	3 (60.0)	0 (0)	0 (0)	1 (20.0)	5
Norwalk virus	0 (0)	0 (0)	0 (0)	5 (71.4)	2 (28.6)	0 (0)	7
<i>Salmonella</i> spp.	11 (5.6)	0 (0)	8 (4.1)	7 (3.6)	147 (75.0)	23 (11.7)	196
<i>Shigella</i> spp.	0 (0)	0 (0)	2 (8.7)	0 (0)	15 (65.2)	6 (26.1)	23
Undetermined	155 (21.8)	38 (5.3)	8 (1.1)	340 (47.8)	83 (11.7)	88 (12.4)	712

Table 3. Summary of all reported foodborne outbreaks, USA, 1982–9, after application of pathogen-specific profiles

Pathogen profile	Aetiology confirmed by lab	Aetiology assigned through profiling	Total (%) outbreaks
Norwalk virus	7	340	347 (33.1)
Norwalk-like outbreaks			
Salmonella, campylobacter, shigella, salmonella-like outbreaks	235	83	318 (30.3)
<i>B. cereus</i> , <i>S. aureus</i>	42	155	197 (18.8)
Vomiting toxin outbreaks			
<i>C. perfringens</i>	24	38	62 (5.9)
Diarrhoea toxin outbreaks			
<i>E. coli</i>	5	8	13 (1.2)
<i>E. coli</i> -like outbreaks			
Remaining undetermined			88 (8.4)
Total			1,049*

\* There are 24 additional outbreaks with a laboratory-confirmed aetiology that were not chosen for further analysis as described in the Methods section.

fitted the diarrhoea-toxin syndrome, and 8 (1.1%) outbreaks fitted the *E. coli* syndrome. When information on known pathogens and the clinico-epidemiologic profiles of outbreaks with undetermined aetiology were combined, only 88 outbreaks of undetermined aetiology (12.4%) remained unclassified (Table 3). Considering the outbreaks with aetiology determined by these profiles, Norwalk virus-like outbreaks ( $n = 347$ , 33.1%) appear to be as common as salmonella-like outbreaks ( $n = 318$ , 30.3%) among all reported outbreaks ( $n = 1049$ ).

## DISCUSSION

Epidemiologic profiles can be used to create pathogen syndromes that classify outbreaks in which no pathogen was recovered by laboratory testing. The

use of epidemiologic profiles provides a more detailed description of the likely causes of foodborne disease outbreaks than does classifying outbreaks of undetermined aetiology together. For example, Dalton et al. [6] recently demonstrated the ability of clinical and epidemiologic profiles to distinguish between outbreaks of enterotoxigenic *E. coli* and Norwalk-like viral gastroenteritis. Although Norwalk viruses were confirmed in approximately 1% of outbreaks in this study, 33.1% of all outbreaks had epidemiologic characteristics of Norwalk-like viruses. Given recent experience at CDC in retrospectively confirming the presence of Norwalk-like viruses in 90% of a selected group of outbreaks of acute non-bacterial gastroenteritis, it is likely that most of the outbreaks that were epidemiologically classified as such, were caused by Norwalk-like viruses [9].

By identifying likely groups of pathogens, epidemiologic profiles are useful as an outbreak investigation tool. Information on signs and symptoms experienced by persons during an outbreak can be collected and quickly analysed, in many cases before appropriate stool specimens can be collected and tested by clinical or reference laboratories. This information may be useful to help guide the investigation to determine the source of the outbreak and initiate control measures. It may also be useful to direct laboratory testing towards likely aetiologies, as well as studies to help detect new and emerging pathogens. For example, in 1998, recognition that two Minnesota outbreaks had epidemiologic characteristics of enterotoxigenic *E. coli* led to the identification of an ETEC outbreak associated with parsley imported from Mexico [10]. Similar methods were used to identify a foodborne outbreak caused by *E. coli* O39, an agent that does not fit into the existing scheme for classifying diarrhoeogenic *E. coli* [11]. Outbreaks such as these represent known and emerging pathogens that are not being accounted for because no effective surveillance exists for them in the United States. Epidemiologic profiles may provide a framework for improving our surveillance of these pathogens.

As in other uses of surveillance data, reporting must be systematic and complete for epidemiologic profiles to achieve a high degree of accuracy and usefulness. Although the profiles associated with individual pathogens are too close to distinguish, the validity of the profiles will increase when a larger, more complete pool of information about clinical symptoms is made available. Examination of the data set by year revealed that completeness of data reporting has increased between 1982 and 1989. The percentage of outbreaks with all six variables reported increased 10% during this time period. This is a step in the right direction, as these data are critical in conducting retrospective studies, and provide the basis for prospective analysis of agents, such as Norwalk virus and diarrhoeogenic *E. coli*. Standardization of collection methods and inclusion of epidemiologic information along with laboratory analysis is essential to improve the interpretability.

A limitation of this data set was the lack of complete data reporting. Only one quarter of reported outbreaks contributed to this study, and there were several pathogens included in this data set that could not be summarized due to missing data. A follow-up study on more current data could be performed to validate our observations on outbreaks of known

aetiology in order to extend them to the group of outbreaks for which no pathogen was identified in the laboratory. Finally, outbreaks that were consistent with two or more pathogen syndromes were unable to be assigned to one syndrome due to a high margin of error with the tiebreaker programme. Further studies could improve upon the programmes presented here to develop a more accurate hierarchy when faced with outbreaks matching more than one pathogen syndrome. In the future, additional data such as subtype, age and gender could be incorporated into the analysis. Significant differences in burden of illness may exist between age groups and subtypes. A variable indicating the proportion with bloody diarrhoea would also be helpful in distinguishing between the otherwise similar clinical presentation of *E. coli* O157:H7 and ETEC outbreaks. As most of the *E. coli* outbreaks in this data set were probably O157:H7, the inability to distinguish between ETEC and STEC outbreaks was probably not significant. With more complete reporting of data on laboratory-confirmed outbreaks, future studies could use epidemiologic profiles to create other useful pathogen syndromes in addition to the five presented here. Results of this study suggest that routine use of epidemiologic profiles should improve the quality of outbreak investigations. This will, in turn, improve outbreak surveillance and help provide better estimates of the overall occurrence of foodborne diseases.

#### ACKNOWLEDGEMENTS

We thank Alan Lifson, at the University of Minnesota, School of Public Health, Minneapolis, MN, for his guidance throughout the course of this study.

#### REFERENCES

1. Olson SJ, MacKinnon LC, Goulding JS, Bean NH, Slutsker L. Surveillance for foodborne disease outbreaks – United States, 1993–1997. *MMWR* 2000; **49** Suppl 1: 1–62.
2. Mead PS, Slutsker L, Dietz V, et al. Food-related illness and death in the United States. *Emerg Infect Dis* 1999; **5**: 607–25.
3. Hedberg CW, MacDonald KL, Osterholm MT. Changing epidemiology of foodborne disease: a Minnesota perspective. *Clin Infect Dis* 1994; **18**: 671–82.
4. Slutsker L, Altekruze SF, Swerdlow DL. Foodborne diseases: emerging pathogens and trends. *Emerg Infect Dis* 1998; **12**: 199–211.
5. Kaplan JF, Feldman R, Campbell DS, et al. The

- frequency of Norwalk-like pattern of illness in outbreaks of acute gastroenteritis. *Am J Pub Hlth* 1982; **72**: 1329–31.
6. Centers for Disease Control and Prevention. Guidelines for confirmation of foodborne disease outbreaks. *MMWR* 1996; **45** Suppl 5: 57–66.
  7. Hedberg CW, Osterholm MT. Outbreaks of foodborne and waterborne viral gastroenteritis. *Clin Microbiol Rev* 1993; **6**: 199–210.
  8. Dalton CB, Mintz ED, Wells JG, Bopp CA, Tauxe RV. Outbreaks of enterotoxigenic *Escherichia coli* infection in American adults: a clinical and epidemiologic profile. *Epidemiol Infect* 1999; **123**: 9–16.
  9. Frankhauser RL, Noel JS, Monroe SS, et al. Molecular epidemiology of 'Norwalk-like viruses' in outbreaks of gastroenteritis in the United States. *J Infect Dis* 1998; **178**: 1571–8.
  10. Naimi T, Kassenborg H, Hedberg C, et al. Epidemiologic and molecular aspects of linked enterotoxigenic *Escherichia coli* outbreaks, Minnesota – 1998. In: Program and abstracts of the 39th Annual Interscience Conference on Antimicrobial Agents and Chemotherapy. San Francisco: 2212.
  11. Hedberg CW, Savarino SJ, Besser JM, et al. An outbreak of foodborne illness caused by *Escherichia coli* O39:NM, an agent no fitting into the existing scheme for classifying diarrheogenic *E. coli*. *J Infect Dis* 1997; **176**: 1625–8.