Epidemiol. Infect. (1998), **120**, 37–42. Printed in the United Kingdom

Surveillance of outbreaks of waterborne infectious disease: categorizing levels of evidence

H. E. TILLETT, J. DE LOUVOIS* AND P. G. WALL

Public Health Laboratory Service Communicable Disease Surveillance Centre, 61 Colindale Avenue, London, NW9 5EQ, UK

(*Accepted 2 October 1997*)

SUMMARY

Public health surveillance requires the monitoring of waterborne disease, but sensitive and specific detection of relevant incidents is difficult.

The Communicable Disease Surveillance Centre receives information from various sources about clusters of cases of illness in England and Wales. The reporter may suspect that water consumption or recreational water exposure is the route of infection, or subsequent investigation may raise the hypothesis that water is associated with illness.

It is difficult to prove beyond reasonable doubt that such a hypothesis is correct. Water samples from the time of exposure are seldom available, some organisms are difficult to detect and almost everyone has some exposure to water. Therefore, we have developed a method of categorizing the degree of evidence used to implicate water. The categories take into account the epidemiology, microbiology and water quality information. Thus outbreaks are classified as being associated with water either 'strongly', 'probably' or 'possibly'.

This system allows a broad database for monitoring possible effects of water and is not confined to the few outbreaks which have been intensively investigated or have positive environmental microbiology. Thus, for reported incidents, the sensitivity of classifying it as water associated should be high but this may be at the expense of specificity, especially with the 'possible' association.

INTRODUCTION

Contaminated water has the potential to cause extensive outbreaks of illness due to the size of populations served by some distribution systems and the large numbers of people who use some recreational water facilities. The biggest documented drinking water outbreak occurred in Milwaukee, USA where an estimated 403000 residents developed cryptosporidiosis [1]. The community then experienced a period of secondary infections [2].

The surveillance of outbreaks of waterborne infection, which are usually of gastrointestinal (GI) illness, requires the detection of increased illness in a population and evidence to demonstrate that water was the route of transmission. An outbreak can be defined as an incident in which two or more people thought to have a common exposure experience a similar illness, or proven infection where at least one is ill. There are established mechanisms for collecting information and for investigating outbreaks in general which have been tried and tested in national surveillance centres [3]. However, if the outbreak is associated with water there may be difficulties.

This paper describes problems with assessing waterborne outbreaks and gives a description of the current approach to surveillance in England and Wales.

* Author for correspondence.

https://doi.org/10.1017/S0950268897008431 Published online by Cambridge University Press

DETECTING AND INVESTIGATING OUTBREAKS

Problems with identifying waterborne outbreaks

Detecting cases

Increases in numbers of clinical cases may be difficult to detect against the fluctuating background of gastroenteritis. If illness is severe enough for patients to seek medical advice then an increase in morbidity may be noticed within General Practice or through a morbidity reporting scheme. If faeces specimens are taken then the diagnostic laboratories may note an increase in the numbers of referrals. If the specimens yield positive microbiology then there is the potential to spot increases in cases of that diagnosis, either in the laboratory or in local public health departments or in regional or national centres where laboratory data are collected. This includes the PHLS Communicable Disease surveillance Centre (CDSC) where there is an automated screening process to look for unusual increases in weekly reports of certain organisms, including cryptosporidia, giardia and campylobacter [4].

Monitoring disease patterns using laboratory reports is a useful way of looking for major trends but is prone to major biases due to under-detection, under-reporting and changes in laboratory methods [5]. These biases affect geographical comparisons especially, because of varying practices in GP referral [5, 6].

Many of the organisms which have been associated with water ingestion are also spread by other routes of transmission. Thus observed increases may be non-waterborne, even if they are genuine outbreaks rather than random fluctuation in the disease pattern. Outbreaks of disease which are investigated microbiologically but where no causative organism is identified will not be detected through a laboratory reporting scheme.

The number of outbreaks of GI illness which go undetected either nationally or locally is unknown. Only a minority of laboratory reports of GI pathogens arise from identified outbreaks. [7].

Water microbiology

Occasionally a water contamination incident may be reported which will lead to enhanced surveillance of illness. More often the quality of the water is only questioned after a cluster of clinical cases has come to light and water is thought to be a possible route of transmission. By the time such an outbreak is detected and investigated the relevant body of water may have gone, although every opportunity should be taken to find archived water or environmental evidence (e.g. ice, reservoirs, filters). With drinking water incidents the supply may only have been contaminated transiently, with recreational water incidents the flow of natural water or the changeover in artificial pools may leave no evidence as to the water quality at the relevant time. This contrasts with outbreaks transmitted from person-to-person where contacts can be traced and tested, or with food poisoning outbreaks where leftover food or ingredients can often be recovered. Although drinking water and some recreational waters are monitored in accordance with regulations this is unlikely to help because samples are collected infrequently and from relatively few sites. These routine samples are examined for indicator organisms such as total coliform organisms and seldom for pathogens. The samples cannot be stored for repeat testing.

If contamination at the water source is still continuing at the time of an outbreak investigation then the pathogen may be detectable, although some are hard to recover. Failure to detect a pathogen in the implicated water cannot be interpreted as absolving it.

Epidemiology

Because of the problems associated with the microbiological examination of water it is likely that much of the investigation will rely on epidemiological evidence.

Descriptive epidemiology

The aim of the investigator will be to establish *who* was ill, *when* and *where*. It will then become apparent that the primary cases have been exposed to the same water supply or recreational source. This fact could, however, be coincidental due to the patients all living or working in the same area and thus having other activities in common, such as eating at the same restaurant or shopping at the same grocers. Thus the apparent correlation with water exposure could be an indirect rather than a direct association (i.e. a statistical confounding). If the descriptive epidemiology is to be accepted as suggestive of a waterborne outbreak then attempts must be made to exclude other routes of transmission appropriate for the causative organism. For example in a cryptosporidium outbreak the data collection should include a check on whether the cases had been on a farm visit,

https://doi.org/10.1017/S0950268897008431 Published online by Cambridge University Press

travelled abroad, etc. Lists of checks should be expanded whenever new knowledge about routes of transmission is published. Particular attention needs to be given to the plausibility of the timing, not only to reflect the incubation of the organism concerned but the course of the suspected body of water, which may have been contaminated for a single episode or longer-term.

Descriptive epidemiology does not furnish direct confirmation of the route of transmission but may be useful in raising hypotheses about the cause of the outbreak which are then tested by an appropriately designed observational study [8]. Such a study should also ask questions about other plausible routes of transmission in case the descriptive epidemiology failed to expose their possible role.

Analytical epidemiology

Observational studies, whether case-control or cohort in design, are subjected to statistical analysis to test the hypothesis of the waterborne route. The investigator will use a specially designed questionnaire incorporating standard features to collect data from patients and from well people who might have been exposed. Success of the study in illustrating whether or not water was associated with illness depends partly on good statistical power. This is achieved when there are sufficient numbers interviewed in the two groups, cases and non-cases. Waterborne outbreaks may yield few laboratory confirmed cases which lessens the power of analytical studies. A clinical case definition may be preferable and more speedy and, indeed, is essential when no pathogen has been identified.

The power of a study depends on good differentiation in exposure between the ill and the well. Poor statistical power is sometimes encountered in food poisoning outbreak investigations. If the exposure was a single function where guests were offered a set menu then almost everyone will have eaten the same food items. Thus, although all the cases will have eaten the contaminated food item, so too will most of the non-cases and there will be no apparent statistical association between consumption and illness. This should not be regarded as 'not statistically significant' but as inconclusive because the statistical power is low and the confidence intervals of the estimated risk will be very wide. Functions offering a wide choice of buffet items are easier to investigate successfully.

Drinking water outbreaks, as with food poisoning from a set menu, may have poor power because most people questioned will have had some exposure to unboiled water even if it is only for teeth cleaning or food preparation. If the outbreak and therefore the study are sufficiently large then an analysis of amounts of water consumed and risk of illness may yield a significant association. This is sometimes referred to as a 'dose-response' analysis although it involves the dose of the potential source of contamination rather than dose of organisms themselves [9]. A statistical analysis of trend to look for an increased risk of illness with increasing average consumption of water was first proposed by Armitage [10]. It can be partly implemented using EPI INFO software, although this only reports on the linear component of the trend. It is important to check for non-linearity because non-monotonic increments in risk may be less convincing.

The absence of a significant association with water in an observational study, using the statistical convention of a probability of no association being $\leqslant 0{\cdot}05$, should be interpreted in the context of the power of the study. If power is low then the water may not be absolved.

The success of the study also depends on the elimination of confounding effects which could bias the interpretation, such as age group, sex and consumption of other drinks. In an ideal situation the cases and the controls will be interviewed without their knowing the hypothesized route of infection. With food poisoning outbreaks of brand-named products the question can be buried in among many other brand names and a rapid telephone case-control study can provide enough evidence prior to publicity and withdrawal of the product from sale. Suspected drinking water incidents tend to receive considerable publicity before or during the study period which may result in biased responses.

The interpretation of all epidemiological results must be made with acknowledgement to possible sources of bias and confounding. This evidence is then evaluated in conjunction with the microbiology and water quality information. Conclusions from the outbreak investigation should follow a decision making process [8]. This compilation of evidences is the basis for the proposed classification scheme.

CLASSIFICATION OF RESULTS

The CDSC surveillance scheme encourages reporting of all outbreaks of infectious intestinal disease, including those where water is suspected to be the vehicle of infection [11]. Information about the outbreak is reviewed as it accumulates and the report is added to the data base on waterborne outbreaks if, and only if, there is some supportive evidence. Anecdotal reports are not accepted but, because of the problems indicated in the introduction, acceptance is not restricted to outbreaks with a lucky investigation leading to positive water microbiology and significant statistical epidemiology. Outbreak investigations will be expected to have covered water and clinical microbiology, epidemiology and information on water treatment and quality.

The following set of definitions which rank the strength of association between water exposure and illness were introduced in the Communicable Disease Report (CDR) together with 6-monthly reviews of water associated illness and environmental water microbiology for England and Wales [12].

Definitions for use in outbreak investigations are expressed in logic combinations:

1. *Strongly associated with water*. Evidence from an analytical epidemiological study demonstrates association between water and illness *and* pathogen identified in clinical cases is also found in water

OR

descriptive epidemiology suggests that the outbreak is water related and excludes obvious alternative explanations *and* pathogen identified in clinical cases is also found in water

OR

evidence from an analytical epidemiological study demonstrates association between water and illness *and* water quality failure and/or water treatment problems of relevance are recorded but pathogen not detected in water.

2. *Probably Associated with water*. Descriptive epidemiology suggests that the outbreak is water related and excludes obvious alternative explanations *and* water quality failure and/or water treatment problems of relevance are recorded but pathogen not detected in water

OR

evidence from an analytical epidemiological study demonstrates association between water and illness *and* supporting water microbiology is absent

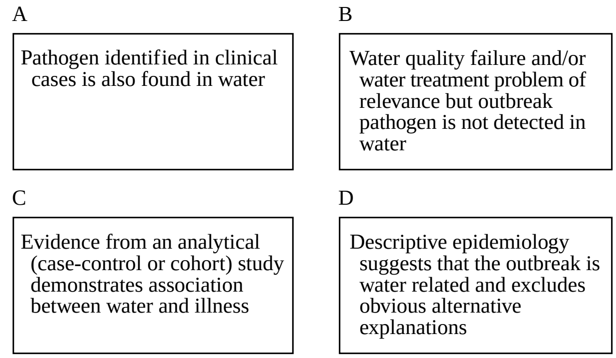
OR

pathogen identified in clinical cases is also found in water *and* supporting epidemiological evidence is absent.

3. *Possibly associated water*. Water quality failure and/or water treatment problems of relevance are recorded but pathogen not detected in water *and* supporting epidemiological evidence is absent

OR

descriptive epidemiology suggests that the outbreak is water related and excludes obvious alternative explanations *and* supporting microbiological evidence is absent.



strongly associated **if** (A + C) *or* (A + D) *or* (B + C).
probably associated **if** (B + D) *or* C only *or* A only.
possibly associated **if** B only *or* D only.

Fig. 1. Classifications.

These defined strengths of association of a cluster of cases with water as the route of transmission are summarized schematically in Figure 1. Boxes A and B are hierarchical (i.e. A takes precedence if both A and B apply) as are C and D.

The categorizations were tested retrospectively on two sets of outbreaks reported to CDSC. The first set are six outbreaks which were reported in journal publications and they are summarized in Table 1. They all involved cryptosporidiosis and one also involved campylobacter infections; five were associated with drinking water supplies and one with a swimming pool. In three [14, 16, 17] analytical epidemiology, using special observational studies, showed a statistically significant association with quantity of water drunk or with head immersion in the pool incident. In three outbreaks [13–15] oocysts were recovered from treated water. In the private water supply outbreak [18] pathogens were not detected but there were high counts of *E. coli* indicating faecal contamination.

Thus in four of the six outbreak investigations water microbiology gave some supportive evidence to the hypothesis of water association. In the other two there were relevant observations from water treatment records. In one outbreak [16] a part-time treatment works had been kept in operation longer than usual during which time water was abstracted which was of higher than normal turbidity. The relevance of the timing and the likelihood that turbidity correlated with cryptosporidium contamination remain debat-

Table 1. *Summary of six published outbreak reports*

Reference	Date	Number of cases	Organism	Summary	Association
13	4/88	27	Cryptosporidium	Public water supply; descriptive epidemiology showed association with water from treatment works; oocysts found in treated water	Strong [A+D]
14	8/88	67	Cryptosporidium	Swimming pool; case-control study showed association with immersion; oocysts found in water	Strong [A+C]
15	12/88	516	Cryptosporidium	Public water supply; descriptive epidemiology showed association with water from treatment works; oocysts found in treated water	Strong [A+D]
16	12/90	47	Cryptosporidium	Public water supply; case-control study showed association with water consumption; water microbiology negative but increased turbidity in raw water	Probable [C only]
17	11/92	47	Cryptosporidium	Public water supply; case-control study showed association with water consumption; water draining from grazing fields had bypassed filtration; pathogen not found in treated water	Strong [B+C]
18	5/93	43	Campylobacter and Cryptosporidium	Private water supply; descriptive epidemiology showed association; high levels of indicator organisms found in water, but no pathogens	Probable [B+D]

Table 2. *Size and strength of association of outbreaks reported for 1992–5*

Strength of association	Number of outbreaks	Number of reported cases	
		Median	Range
Strong	11	53	8–575
Probable	9	40	4–108
Possible	6	19	5–42

able and so this outbreak was arbitrarily classified as a ‘probable’ association. In the other outbreak [17] the conditions had led to water from a grazed pasture bypassing the usual, natural sandstone filtration. This was considered to be a ‘water treatment problem of relevance’ and the association was classified as ‘strong’.

These published outbreak reports involve large outbreaks where there was substantial evidence supporting the water route. For a wider test of the classification system it has also been applied to the surveillance database held at CDSC, looking at all outbreaks from 1992–5 where water was suspected [11]. The 26 outbreaks were all classifiable. There were 11 outbreaks where the association was rated as strong, 9 as probable and 6 as possible. The numbers of cases involved are summarized in Table 2. The average size diminishes with strength of evidence, from a median of 53 with ‘strong’, 40 with ‘probable’ to 19 with ‘possible’.

DISCUSSION

A surveillance database of information on waterborne outbreaks requires a consistent approach to assessing the evidence available. The evidence may be difficult to collect, as described in the introduction. The proposed categorization system sets guidelines for the minimum amount of evidence which is required to justify recording the outbreak as potentially relevant for public health study. This should encourage reporters to send in information even from outbreaks which have been difficult to investigate, perhaps because of small numbers of confirmed cases, or where all the microbiological and epidemiological findings are negative. Initial reports should include all episodes where water is a plausible explanation.

The value of this wide set of classifications is that public health lessons may be learned even from the outbreaks with minimal evidence. The classifications may not be reliably specific in that some outbreaks accepted onto the surveillance database may have been caused by another route. Every attempt is made to minimise this by applying the categorizations as consistently as possible and reviewing the database whenever new information comes to light, thus allowing reclassification or deletion if applicable. The final attribution of the strength of association will depend on the combination of evidence. Inevitably there is some arbitrariness involved. An analytical epidemiological study may just fail to achieve conventional statistical significance. The interpretation of what is a relevant water treatment problem may be a matter of opinion as illustrated in Table 1, reference 16. Although the underlying aim will be to assess reports consistently over time, it is realized that new information about water microbiology and treatment is emerging and so modifications will need to be made to the way in which these classifications are applied. However, reproducibility in applying the classifying technique is paramount, and modifications should be documented. Successful outbreak surveillance may help feed this process. Published outbreak reports have highlighted water treatment problems (Table 1) and here the associations were strong or probable. It may be that a well-publicized surveillance scheme which accepts possible associations will help to highlight any new water quality problems or unusual pathogens emerging in the future.

Regular and timely feed back of information to reporters should help encourage investigation of outbreaks and completeness of reporting to the surveillance centre. This feedback is made through the CDR publication with routine sections every 6 months and *ad hoc* coverage of current news items. The waterborne events and lessons to be learned from them thus reach a wide readership of public health professionals. The publication of these 6-monthly reviews does not preclude more detailed publication of outbreak reports elsewhere.

The surveillance scheme should allow data to be monitored over time. Incompleteness in ascertaining and in reporting outbreaks makes the data potentially biased, as is the case with all routine surveillance. The approach to monitoring described here proposes methods which should reduce that bias although some incompleteness is inevitable.

Terminology for these classifications has been devised in a public health context and may not transfer into other contexts, such as legal applications.

ACKNOWLEDGEMENTS

We thank Drs Mary O'Mahony, Norman Begg and James Stuart for encouragement and support.

REFERENCES

1. MacKenkie WR, Hoxie NJ, Proctor ME, et al. A massive waterborne outbreak of *Cryptosporidium* infection transmitted through a public water supply. N Engl J Med 1994; **331**: 161–7.
2. Osewe P, Addiss DG, Blair KA, Hightower A, Kamb ML, Davis JP. Cryptosporidiosis in Wisconsin: a case-control study of post-outbreak transmission. Epidemiol Infect 1996; **117**: 297–304.
3. Galbraith NS, Palmer S. General epidemiology. In: Topley and Wilson's Principles of bacteriology, virology and immunity, vol 3. London: Edward Arnold, 1990.
4. Farrington CP, Beale AD. Computer-aided detection of temporal clusters of organisms reported to the Communicable Disease Surveillance Centre. CDR 1993; **3**: R78–R82.
5. Tillett HE, Thomas MEM. Monitoring infectious diseases using routine microbiology data. J Hyg 1981; **86**: 49–69.
6. Palmer S, Houston H, Lervy B, Ribeiro D, Thomas P. Problems in the diagnosis of foodborne infection in general practice. Epidemiol Infect 1996; **117**: 479–84.
7. Wall PG, de Louvois J, Gilbert RJ, Rowe B. Food poisoning: notifications, laboratory reports and outbreaks. Where do the statistics come from and what do they mean? CDR Rev 1996; **7**: R93–R100.
8. Tillett HE, Carpenter RG. Statistical methods applied in microbiology and epidemiology. Epidemiol Infect 1991; **107**: 467–78.
9. Tillett HE. Statistical analysis of case-control studies of communicable diseases. Int J Epidemiol 1986; **15**: 126–33.
10. Armitage P. Tests for linear trends in proportions and frequencies. Biometrics 1955; **11**: 375–86.
11. Furtado C, Adak GK, Stuart J, Wall PG, Evans HS, Casemore DP. Outbreaks of waterborne infectious intestinal disease in England and Wales, 1992–1995. Epidemiol Infect. Submitted.
12. CDSC. Strength of association between human illness and water: revised definitions for use in outbreak investigations. CDR 1996; **6**: 65, 68.
13. Smith HV, Patterson WJ, Hardie R, et al. An outbreak of waterborne cryptosporidiosis caused by post-treatment contamination. Epidemiol Infect 1989; **103**: 703–15.
14. Joce RE, Bruce J, Kiely D, et al. An outbreak of cryptosporidiosis associated with a swimming pool. Epidemiol Infect 1991; **107**: 497–508.
15. Richardson AJ, Frankenberg RA, Buck AC, et al. An outbreak of cryptosporidiosis in Swindon and Oxfordshire. Epidemiol Infect 1991; **107**: 485–95.
16. Joseph C, Hamilton G, O'Connor M, et al. Cryptosporidiosis in the Isle of Thanet; an outbreak associated with local drinking water. Epidemiol Infect 1991; **107**: 509–19.
17. Bridgman SA, Robertson RMP, Syed Q, Speed N, Andrews N, Hunter PR. Outbreak of cryptosporidiosis associated with a disinfected groundwater supply. Epidemiol Infect 1995; **115**: 555–66.
18. Duke LA, Breathnach AS, Jenkins DR, Hatkis BA, Codd AW. A mixed outbreak of cryptosporidium and campylobacter infection associated with a private water supply. Epidemiol Infect 1996; **116**: 303–8.