Epidemiol. Infect. (2002), **128**, 433–438.
DOI: 10.1017/S0950268802006805 Printed in the United Kingdom

A community survey of self-reported gastroenteritis undertaken during an outbreak of cryptosporidiosis strongly associated with drinking water after much press interest

P. R. HUNTER[1]* AND Q. SYED[2]

[1] *School of Medicine, Health Policy and Practice, University of East Anglia, Norwich NR4 7TJ, UK*
[2] *Communicable Disease Surveillance Centre – North West, Vernon Pritchard Court, 57a Upper Northgate Street, Chester CH1 4EF, UK*

(*Accepted 3 January 2002*)

SUMMARY

We took the opportunity provided by a large outbreak of cryptosporidiosis in the North West of England to investigate the potential impact of recall bias on strength of association and on estimates of outbreak size. We conducted a community-based survey of 4 towns within the outbreak area and 4 control towns. A postal questionnaire was sent to 120 homes, chosen at random from the local telephone directory, in each of the 8 towns. Although not statistically significant, the prevalence of self-reported diarrhoeal disease was marginally higher in the control towns than in the outbreak towns. There was a very strong association between self-reported diarrhoea and drinking water consumption in both control and outbreak areas. The impact of recall bias in outbreak investigations is much greater than previously thought. Identification of the cause of outbreaks should not be based solely on case-control studies conducted after the press has reported the outbreak. Such evidence is likely to be unreliable and give falsely significant associations between water consumption and disease.

INTRODUCTION

Cryptosporidium was first identified as a cause of waterborne disease after an outbreak in Texas in 1984 [1]. Since then it has become recognized as the most commonly detected cause of outbreaks associated with mains drinking water in the United Kingdom [2, 3]. Although not the most common cause of water-borne outbreaks in the United States, an outbreak in Milwaukee in the spring of 1993 has become the largest recorded single outbreak of any waterborne disease, with an estimated 403 000 cases [4]. However, only some 739 of these cases were identified by positive laboratory examinations. The estimated size of the outbreak was based on a telephone-based community survey conducted some time after the outbreak. A recent study conducted in the United Kingdom has suggested that retrospective surveys of subjective diarrhoeal disease may substantially overestimate the true population level of illness [5]. If this is the case, then the estimated size of the Milwaukee outbreak is also an overestimate. Even so, it is clear that only a proportion of cases of cryptosporidiosis have stools taken for examination. Consequently, there remains the question of how large outbreaks of waterborne cryptosporidiosis really are. We took the opportunity provided by an outbreak of cryptosporidiosis in the North West of England to conduct a population-based survey in an attempt to estimate the true size of the outbreak.

* Author for correspondence.

The outbreak

The full details of the outbreak of cryptosporidiosis in the North West of England during April–May 1999 are presented elsewhere [6]. During April and May there were some 498 cases of cryptosporidiosis reported by laboratories to one Communicable Disease Surveillance Centre – North West, of which some 308 (62 %) were resident in one of just four Health Authorities. These four authorities receive most of their drinking water from Thirlmere reservoir in the English Lake District. Within the distribution of this single source there was a strong correlation between attack rate and proportion of water received from the implicated reservoir (χ^2 for linear trend = 76·8, $P < 0{\cdot}0000001$). The outbreak was preceded by the detection of a high count of cryptosporidial oocysts (34) in a 10 l random sample, which was reported on 23 April (2 weeks before the first outbreak meeting), having been taken 3 days earlier on 21 April. This water was taken from a water treatment works that treated water from Thirlmere Reservoir. Subsequent samples were either negative or contained very low counts. The sample with the highest count was reported before the outbreak was detected, but in retrospect the outbreak had started even before the sample was taken. The outbreak was judged to be strongly associated with drinking water on the basis of the detection of oocysts in the water and descriptive epidemiology that supported a waterborne hypothesis and excluded possible alternative hypotheses.

A feature of this outbreak was the large amount of media interest. Regional and local television, radio and press featured this outbreak on numerous occasions, clearly linking the outbreak to the water supply.

Community survey

We identified four towns in the area affected by the outbreak which were within the distribution of drinking water from Thirlmere. Four towns that were close to the outbreak area, but did not receive any water from Thirlmere were chosen as controls. For each town, 120 names and addresses were selected randomly from up-to-date telephone directories. A postal questionnaire was sent to each household asking about all members of the family. To increase the response rate the questionnaire was kept very simple and asked for the age and sex of each member of the household, whether they drank unboiled tap water and if so, how much they drank each day, whether they had had diarrhoea since the 15 April and if so, how long had the illness lasted, had they visited their doctor and had they lost any time from work. The questionnaire was sent out during the first week of June. Those who had not responded were sent a second questionnaire 2 weeks later and if they had not responded to that, within a week, a contact was made by telephone with an offer to complete the questionnaire over the telephone. All completed questionnaires were entered onto a database and analysed using StatsDirect [7].

Table 1. *Age and sex distribution of respondents from the outbreak and control towns*

	Outbreak towns	Control towns
Sex distribution		
Male	408	384
Female	405	410
Unknown	2	4
Age group/years		
0–9	106	89
10–19	87	83
20–29	99	107
30–39	112	109
40–49	103	80
50–59	127	121
60–69	87	109
70–79	61	69
80–89	23	22
90–99	3	4
Unknown	7	5
Total	815	798

RESULTS

Overall some 701 questionnaires covering 1613 individuals were able to be included in the study. This represents a response rate of 73 %. Response rates varied between towns with Blackburn having the lowest at 62 % and Ormskirk the highest at 88 %. The response rate from the outbreak area was marginally lower (71 %) compared to the control area (75 %) but this was not statistically significant. The age and sex distribution between respondents from the outbreak and control towns is shown in Table 1. There was no significant difference between the outbreak and control towns for either age or sex of respondent.

The attack rates for each town, for all outbreak and control areas and all towns are shown in Table 2. Overall there the attack rate of self-reported gastroenteritis was slightly greater in the control areas but this was not statistically significant.

To obtain an indication of the severity of any illness we asked about the duration, whether the individual

Table 2. *Population covered by respondents and diarrhoeal attack rates for each town, for all towns within and without the outbreak area and for all towns*

Town	Part of outbreak	Total responders	Total reporting diarrhoea	Attack rate/% (95% CI)
Blackburn	No	159	28	18 (12–24)
Chorley	Yes	227	26	12 (8–16)
Ormskirk	No	255	22	9 (6–13)
Preston	Yes	196	17	9 (5–14)
Rochdale	No	193	20	10 (6–16)
Salford	Yes	171	28	16 (11–23)
Warrington	No	191	38	20 (15–26)
Wigan	Yes	221	33	15 (11–20)
Towns in outbreak area		815	104	13 (11–15)
Control towns		798	108	14 (11–16)
All towns		1613	212	13 (12–15)

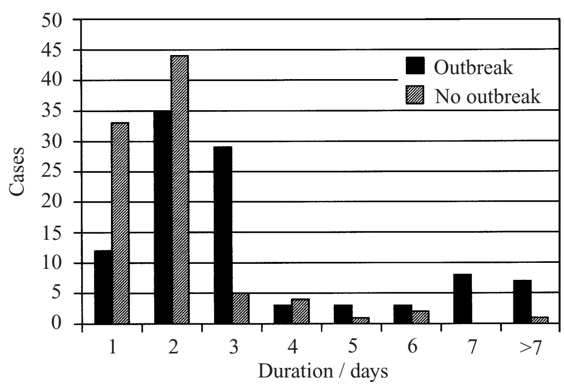

Fig. 1. Duration of self-reported diarrhoeal illness in outbreak and non-outbreak areas.

attended the doctor because of any diarrhoea and how long the diarrhoea lasted. In the outbreak area only 6 of 104 people reported visiting their doctor compared to 13 of 108 in the control areas. Some 12 people reported losing time (25 person days) off work or school because of the diarrhoea in the outbreak area and 7 persons a total of 21 person days in the control areas. There was, however, a difference in the duration of illness reported between the two areas (Fig. 1). The mean duration of diarrhoea was highest in the outbreak areas 3·45 days compared to the control areas 2·28 ($t = 3{\cdot}186$, D.F. $= 192$, $P = 0{\cdot}0017$).

The relationship between self-reported diarrhoea and water consumption behaviour is shown in Table 3. In the outbreak areas there is a very strong correlation between unboiled tap water consumption and diarrhoea, including the presence of a dose–response effect. There is also almost as strong a correlation in the control areas.

DISCUSSION

The primary aim of this study was to validate community surveys as a means of estimating the total size of outbreaks of cryptosporidiosis. The study reported here is essentially the same as that done in Milwaukee, though it differed in one respect. In our study we used a postal questionnaire and telephoned respondents only if they had not replied to our initial two mailings. In Milwaukee, respondents were initially contacted by phone. In our study we also did not specify that diarrhoea should be 'watery'. The reason for this was that a high proportion of our laboratory confirmed cases reported having had bloody diarrhoea [6]. Cryptosporidiosis is not generally thought to be associated with bloody diarrhoea. However, our observation of bloody diarrhoea in this outbreak is consistent with observations in about 10% of cases in an ongoing study of cryptosporidiosis. The biological mechanisms underlying this finding are not clear.

The finding in our study that respondents in the control areas reported an even greater incidence of diarrhoea than those in the outbreak areas effectively invalidates the community-based survey as a tool for estimating outbreak size. If we had not sampled the control towns and used the Milwaukee approach to estimating outbreak size we would have estimated a total size of 12·8% of 1 260 160 (the population served by Thirlmere Reservoir water), i.e. 161 300 cases, less an estimate of the background rate of 0·5% per

https://doi.org/10.1017/S0950268802006805 Published online by Cambridge University Press

Table 3. *Association between diarrhoea and tap water consumption in the outbreak and non-outbreak towns*

	No diarrhoea	Diarrhoea	Relative risk	*P* value
Outbreak towns				
No tap water	200	8	4·12 (2·04–8·33)	0·00001
Tap water	510	96		
Amount tap water/pints				
0	203	8	1	
< 1	189	16	2·08 (0·90–4·71)	
> = 1 and < 3	253	59	4·99 (2·43–10·22)	
> = 3 and < 5	33	13	7·45 (3·28–16·94)	
> = 5	6	2	6·59 (1·66–26·19)	0·00000
Control towns				
No tap water	216	13	2·94 (1·68–5·14)	0·00004
Some tap water	474	95		
Amount tap water/pints				
0	218	13	1	
< 1	128	24	2·81 (1·47–5·34)	
> = 1 and < 3	270	52	2·87 (1·60–5·14)	
> = 3 and < 5	38	12	4·26 (2·07–8·78)	
> = 5	6	3	5·92 (2·04–17·17)	0·00003

month would give a figure of 9500 to give 151 800 cases. This estimate of 0·5% per month background diarrhoeal incidence is a gross underestimate. Using a retrospective survey of self-reported symptoms a recent US FoodNet survey estimated the annual prevalence of diarrhoea in America at 1·4 episodes per person per year or 11·7% of people per month [8]. In a recent large retrospective study in the United Kingdom, it was suggested that 6·7% of people developed diarrhoea per month [5]. Using the UK figure of 6·7% in our population we would expect that 84 430 people would have had diarrhoea, unrelated to the outbreak, so the total estimated outbreak size would be 76 870 cases. However, given that the self-reported incidence of diarrhoea in the control towns was marginally higher than in the outbreak towns, it would seem to us that even this estimate is far too high. Surveys of self-reported symptoms are very poor indicators of outbreak size. Recent evidence of a very high rate of sero-conversion from children resident in the outbreak area of Milwaukee [9] does not invalidate this conclusion as the relationship between infection, development of symptoms and sero-conversion is still far from clear, especially in adults.

The problem with retrospective surveys in estimating outbreak size is likely to be due to recall bias [10, 11]. As already discussed, the control towns were chosen because of their close proximity to the outbreak area. Although we cannot be certain, we suggest that many of the respondents in the control areas will have believed themselves to live in areas that were at risk from waterborne illness. This would, in our view, increase the risk of recall bias in the control populations.

Perhaps the most interesting finding from this study is that there was a very strong association between illness and water consumption behaviour in both the outbreak and control towns. We would suggest that there are three possible explanations of this finding; the outbreak of waterborne cryptosporidiosis was more extensive than previously appreciated, there were other waterborne infections in the region at the time, and recall bias was affecting individuals' reporting of water consumption behaviour.

It could be the case that many of the respondents with diarrhoea in both outbreak and control towns were in fact suffering from cryptosporidiosis. Indeed, one of the control towns had increased reporting of cryptosporidiosis at that time. That cryptosporidiosis may be present in respondents from the outbreak towns is indicated by the higher proportion of cases reporting prolonged symptoms. It is known that cryptosporidiosis causes more prolonged symptoms than most other causes of gastroenteritis [2]. To investigate this, we excluded all cases of diarrhoea lasting 3 or more days. There is still a strong association between diarrhoea and tap water consumption in control towns (RR = 4·05, 95% CI 1·89 −8·67) and outbreak towns (RR = 15·95, CI 2·12 −110·39). We can, therefore, exclude cryptosporidiosis as the primary cause of this association.

The possibility that control and outbreak towns were affected by a further, unidentified, cause of waterborne gastroenteritis cannot be excluded. We would, however, suggest that this is unlikely as the control towns took their water from very different supplies. Furthermore, most of the other causes of waterborne gastroenteritis are fully sensitive to chlorine and one would not expect chlorinated mains water to be a risk factor [2]. The most likely cause for the association between self-reported diarrhoea and water consumption is recall bias. Nevertheless, the possibility that this represents a real affect cannot be excluded. Work by various researchers has suggested that a proportion of cases of diarrhoeal disease is due to drinking water, even when that water meets currently accepted quality standards [10–15]. However, some of these studies are, themselves, open to criticism because of potential recall bias. Furthermore, the attributable risk due to drinking water in our outbreak towns (70%, 95% CI 50–90) and control towns (58%, 95% CI 37–79) are far higher than any of these previous studies would suggest.

We have suggested that recall bias has affected the results of this study in two ways; firstly causing an increase in the estimate of diarrhoea and secondly causing an erroneously high estimate of the relationship between self-reported diarrhoea and water consumption. That recall bias may affect the outcome of epidemiological studies is well known [18–23]. Unfortunately, this is an issue that is seldom, if ever, addressed during outbreak investigation. It is likely that recall bias will be more marked when cases are aware of the hypothesis under investigation. One of the features of waterborne outbreaks of cryptosporidiosis, at least in the United Kingdom, is that they invariably generate considerable press interest. The result of this interest is that most cases would be aware that an outbreak has happened and that it is thought to be waterborne by the time they are interviewed. It is doubtful whether any of our respondents were unaware of the fact that a large waterborne outbreak had occurred in the region.

The results of this study have important implications for the epidemiological investigation of future outbreaks, whether or not they are thought to be waterborne. In the United States and United Kingdom outbreaks of waterborne disease are classified based on the strength of evidence in favour of an association [24, 25]. Both systems give considerable weight to the results of case-control studies. These systems of weighting evidence need to be reassessed in the light of our findings.

We would further suggest there is an urgent need for research to identify the impact of recall bias on the epidemiological investigation of infectious disease outbreaks and on ways of controlling for such bias, either in the designing questionnaires to attempt to quantify bias or in the choice of controls. McCarthy and Giesecke have recently suggested that the use of case–case comparisons would help overcome the problem of bias [18]. A form of this latter approach that has been used in the investigation of waterborne cryptosporidiosis is when controls are chosen from patients who have had a specimen submitted from the laboratory on the same day as the control, but from whom no pathogen was detected [26].

These conclusions raise concerns about other types of epidemiological investigation. In particular, there must be doubts about the impact of recall bias on intervention studies where volunteers are aware of their exposure history [12]. There is a very real danger that just because the news media believe that an environmental factor is responsible for disease, epidemiological studies will support this belief whatever the true cause [27].

REFERENCES

1. Antonio RG, Winn RE, Taylor JP, et al. A waterborne outbreak of cryptosporidiosis in normal hosts. Ann Intern Med 1985; **103**: 886–8.
2. Hunter, PR. Water-borne disease: epidemiology and ecology. Chichester: Wiley, 1997.
3. Furtado C, Adak GK, Stuart JM, Wall PG, Evans HS, Casemore DP. Outbreaks of waterborne infectious intestinal disease in England and Wales, 1992–5. Epidemiol Infect 1998; **121**: 109–19.
4. MacKenzie WR, Hoxie NJ, Proctor ME, et al. A massive outbreak in Milwaukee of cryptosporidium infection transmitted through the public water supply. N Engl J Med 1994; **331**: 161–7.
5. Wheeler JG, Sethi D, Cowden JM, et al. Study of infectious intestinal disease in England: rates in the community, presenting to general practice, and reported in national surveillance. B M J 1999; **318**: 1046–50.
6. CDSC NW. Report of an outbreak of cryptosporidiosis in the Northwest Region – April and May 1999. Liverpool: CDSC – NW, 2000.
7. StatsDirect Buchan IE. StatsDirect Statistical Software version 1.605. 2000. http://www.statsdirect.com

8. CDC/USDA/FDA Foodborne Diseases Active Surveillance Network and CDC's Emerging Infections Program. *FoodNet 1997 Annual Report*. CDC, Atlanta 1998. http://www.cdc.gov/ncidod/dbmd/foodnet/ANNUAL/97_surv.htm
9. McDonald AC, Mac Kenzie WR, Addiss DG, et al. Cryptosporidium parvum-specific antibody responses among children residing in Milwaukee during the 1993 waterborne outbreak. J Infect Dis 2001; **183**: 1373–9.
10. Hennekens CH, Buring JE. Epidemiology in medicine. Boston: Little, Brown and Co., 1987: 272–86.
11. Greenland S. Concepts of validity in epidemiological research. In: Detels R, Holland WW, McEwen J, Omenn GS, eds. Oxford textbook of public health, 3rd edn. Oxford: Oxford University Press, 1997: 597–615.
12. Payment P, Richardson L, Siemiatycki J, Dewar R, Edwardes M, Franco E. A randomized trial to evaluate the risk of gastrointestinal disease due to the consumption of drinking water meeting currently accepted microbiological standards. Am J Publ Hlth 1991; **81**: 703–8.
13. Payment P, Siemiatycki J, Richardson L, Renaud G, Franco E, Prévost M. A prospective epidemiological study of gastrointestinal health effects due to the consumption of drinking water. Int J Environ Hlth Res 1997; **7**: 5–31.
14. Beaudeau P, Payment P, Bourderont D, et al. A time series study of anti-diarrheal drug sales and tap-water quality. Int J Environ Hlth Res 1999; **9**: 293–312.
15. Morris RD, Naumova EN, Levin R, Munasinghe RL. Temporal variation in drinking water turbidity and diagnosed gastro-enteritis in Milwaukee. Am J Publ Hlth 1996; **86**: 237–9.
16. Schwartz J, Levin R, Hodge K. Drinking water turbidity and pediatric hospital use for gastrointestinal illness in Philadelphia. Epidemiol 1997; **8**: 615–20.
17. Zmirou D, Rey S, Courtois X, et al. Residual microbiological risk after simple chlorine treatment of drinking ground water in small community systems. Eur J Publ Hlth 1995; **5**: 75–81.
18. McCarthy N, Giesecke J. Case–case comparisons to study causation of common infectious diseases. Int J Epidemiol 1999; **28**: 764–8.
19. Basso O, Olsen J, Bisanti L, Karmaus W. The performance of several indicators in detecting recall bias. European Study Group on Infertility and Subfecundity. Epidemiol 1997; **8**: 269–74.
20. Chouinard E, Walter S. Recall bias in case-control studies: an empirical analysis and theoretical framework. J Clin Epidemiol 1995; **48**: 45–54.
21. Pastides H, Goldberg RJ. Case-control studies in pediatric epidemiology: parent surrogates and potential pitfalls of inaccurate and selective recall. Sozial- und Praventivmedizin 1992; **37**: 22–6.
22. Drews CD, Greeland S. The impact of differential recall on the results of case-control studies. Int J Epidemiol 1990; **19**: 1107–12.
23. Neugebauer R, Ng S. Differential recall as a source of bias in epidemiologic research. J Clin Epidemiol 1990; **43**: 1337–41.
24. Levy DA, Bens MS, Craun GF, Calderon RL, Herwaldt BL. Surveillance for waterborne-disease outbreaks – United States 1995–1996. MMWR 1998; **47**(SS-5): 1–33.
25. Tillett HE, de Louvois J, Wall PG. Surveillance of outbreaks of waterborne infectious disease: categorizing levels of evidence. Epidemiol Infect 1998; **120**: 37–42.
26. Bridgman S, 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.
27. Hunter PR. Medicine, postmodernism, and the end of certainty. Studies of environmental risk must not be subject to bias from pre-existing beliefs. B M J 1997; **314**: 1045.