Rainfall and outbreaks of drinking water related disease and in England and Wales

Gordon Nichols, Chris Lane, Nima Asgari, Neville Q. Verlander and Andre Charlett

ABSTRACT

A case-crossover study compared rainfall in the 4 weeks before drinking water related outbreaks with that in the five previous control years. This included public and private drinking water related outbreaks in England and Wales from 1910 to 1999. Of 111 outbreaks, 89 met inclusion criteria and the implicated pathogens included *Giardia*, *Cryptosporidium*, *E. coli*, *S.* Typhi, *S.* Paratyphi, *Campylobacter* and *Streptobacillus moniliformis*. Weather data was derived from the British Atmospheric Data Centre There was a significant association between excess cumulative rainfall in the previous 7 days and outbreaks ($p = 0.001$). There was an excess of rainfall below 20 mm for the three weeks previous to this in outbreak compared to control weeks ($p = 0.002$). Cumulative rainfall exceedances were associated with outbreak years. This study provides evidence that both low rainfall and heavy rain precede many drinking water outbreaks and assessing the health impacts of climate change should examine both.

Key words | climate change, drinking, outbreaks, rainfall, waterborne, water safety plans

Gordon Nichols* (corresponding author)
Chris Lane
Environmental and Enteric Diseases Department,
Health Protection Agency Centre for Infections,
61, Colindale Avenue, London NW9 5EQ,
UK
*Also at:
School of Medicine,
Health Policy and Practice,
University of East Anglia,
Norwich NR4 7TJ,
UK
Tel.: +44(0)208 327 6036
Fax: +44(0)208 327 6036
E-mail: *gordon.nichols@hpa.org.uk*

Nima Asgari
Consultant in Communicable Disease Control,
North East and Central London Health Protection
    Unit, 68 Leadenhall Street, London EC3A 3DH,
UK

Neville Q. Verlander
Andre Charlett
Statistics Department,
Health Protection Agency Centre for Infections,
61, Colindale Avenue, London NW9 5EQ,
UK

INTRODUCTION

There have been 111 outbreaks of disease in England and Wales that were thought to be transmitted through the consumption of drinking water over the 20th century. While outbreaks in the nineteenth century were caused by typhoid, paratyphoid and cholera and in the first half of the 20th by typhoid and paratyphoid, these outbreaks were uncommon in the second half of the 20th century and in the last 20 years outbreaks have been dominated by the more newly recognised pathogens *Cryptosporidium* and *Campylobacter* (Galbraith *et al.* 1987; Furtado *et al.* 1998; Smith *et al.* 2006). There is circumstantial evidence that on some occasions an outbreak was preceded by heavy rainfall (Atherton *et al.* 1995; Bridgman *et al.* 1995) or rainfall following a period of dry weather (Willocks *et al.* 1998). In Walkerton, Canada an outbreak of *Escherichia coli*

O157 and Campylobacter was preceded by a one in 60 year extreme weather event (heavy rainfall) (Auld *et al.* 2004).

A study of the relationship between rainfall and waterborne diseases in the US used 548 reported outbreaks between 1948 and 1994 and precipitation data of the National Climatic Data Centre (Curriero *et al.* 2001). They found that waterborne disease outbreaks were preceded by precipitation events above the 90th percentile, and 68% by events above the 80th percentile, with outbreaks due to surface water contamination showing the strongest association with extreme precipitation during the month of the outbreak. Another study examined extreme rainfall and snowmelt in association with 92 Canadian waterborne disease outbreaks between 1975 and 2001 (Thomas *et al.* 2006). A case-crossover methodology examined rainfall,

air temperature and peak stream flow to determine the relationship between high impact weather events and waterborne disease outbreaks. Total maximum degree-days above 0°C and accumulated rainfall percentile were associated with outbreak occurrence. For rainfall events above the 93rd percentile the relative odds of an outbreak increased by a factor of 2.283 (95% [CI] = 1.216–4.285). For each degree-day above 0°C the relative odds of an outbreak increased by a factor of 1.007 (95% confidence interval [CI] = 1.002–1.012).

This study has used data from a variety of sources to create a database of outbreaks in the 20th Century that were caused by public or private drinking water supplies in England and Wales to examine the relationship between rainfall and outbreaks of drinking water related disease.

MATERIALS AND METHODS

Outbreaks

This study used a case-crossover design with the 90 days before the outbreak representing "cases" and the same 90 day period in the previous five years representing "controls". The cases are those water-borne outbreaks which have been reported during the years 1910–1999 where the location and time of the outbreak is known. The sources used included Medline, Communicable Disease Reports, unpublished reports held by the HPA Centre for Infections and published papers (Anonymous 1938; Galbraith *et al.* 1987; Furtado *et al.* 1998) and these were cross referenced to ensure that there were not any duplicate entries. The Medline search was restricted to human waterborne outbreaks in England and Wales in English publications in the 20th Century (983). Outbreaks were defined as two or more cases of infectious disease associated with a common potable water source. Evidence that an outbreak was waterborne was based on the previously published CDSC assessment sytem (Tillett *et al.* 1998) but this classification was not used in the analysis. For each outbreak information was collected on the geographical location (easting's and northing's), average incubation period of the pathogen ($< = 6$ days, 6 to 12 days, $> = 12$ days), season (spring [March, April, May], summer [June, July, August], autumn [September, October, November], winter [December, January, February]), water supply (private water supply, mains), water source (surface, ground), rainfall implicated (yes and no), and whether the exact date of first symptoms in infected individuals was known or estimated. Outbreaks were excluded where the source of infection was attributed to food, recreational use of water, ships berthed in English or Welsh ports or with *Legionella* outbreaks. Outbreaks were excluded from the analysis if the location and date the outbreak occurred were not known or if there was no relevant rainfall data available for the 90 day period before the outbreak.

Rainfall

Daily precipitation data were acquired on-line from the British Atmospheric Data Centre (BADC), the Natural Environment Research Council's (NERC) designated data centre for the Atmospheric Sciences by pre-arranged agreement. Each outbreak location was cross-referenced against a list of available rainfall stations for the United Kingdom, and the closest, representative station with data for the five-year period ending with the outbreak year was chosen. In the event that more than one outbreak occurred at the same location within the 5 year period, additional control years were added to replace those years with outbreaks listed in order to remove any bias caused by what may have been unusual rainfall years.

Where outbreak events were reported with a given date (start of outbreak, or onset of symptoms of first case), station rainfall data for the preceding 90-day period were collated with denominator data represented by rainfall data from the same time period in the preceding 5 years. Where no specific date was available for the outbreak, but a month was indicated, data were collated for the 90 days preceding the last day of the indicated month. Where a season was indicated, (summer) the last day of central month (July) was chosen. In all outbreaks where an estimated date for the start of the outbreak was assumed, the event was recorded as "estimated" in the data-set. In the event of the outbreak report only giving a year, the event was not analysed.

Two approaches were used to examine the relationship between outbreaks and rainfall. The rationale was to

examine the relative worth of different approaches that focus on total rainfall in a particular period and excessive rainfall events. The second approach was complicated by the diverse ways in which exceedence can be determined.

First method of classifying rainfall; cumulative rainfall

In four time periods prior to each outbreak (1–7, 8–14, 15–21, 22–28 days prior to the outbreak), the cumulative rainfall was determined for the outbreak year and the self-control period by averaging over the five non-outbreak years. Due to non-linear associations these were categorised into four groups (0 to 10 mm, >10 to < = 20 mm, >20 to < = 40 mm and >40 mm). An additional analysis for the period eight to twenty-eight days was performed by summing the rainfall in this period.

Due to the paired nature of the case and control data, a conditional logistic regression analysis was performed to estimate the strength of association between the cumulative rainfall categories and waterborne outbreaks. Other factors of interest were considered as possible effect modifiers and their impact was assessed by examining their interaction with the cumulative rainfall categories. An indicator variable was included to assess whether there was any effect modification in those outbreaks where the onset date was uncertain.

Second method of classifying rainfall; excessive rainfall event

An alternative approach was used where an excessive rainfall event was defined as rainfall on a day preceding an outbreak, exceeding the upper limit of the 95% reference range. The upper limit of the 95% reference range was estimated as the mean $+ 1.96 *$ standard deviation, with the mean and standard deviation being calculated from the rainfall for that day and the day preceding and after, in the four years prior to outbreak and control (year before outbreak) years. (e.g. for the 1st June 2005, the mean rainfall was calculated from that falling on the 31st May and 1st and 2nd June in 2000, 2001, 2002 and 2003 (12 days). The rainfall was first subjected to a square root transformation to remove the positive skew in the data, the resultant upper limit was back transformed. The number of excessive events was determined for both the outbreak year and the control year prior to the outbreak. The total number of days in which the rainfall exceeded the upper limits and whether there was at least one day with excessive rainfall since day two until a particular day of interest (the cumulative exceedance), up to ninety days prior to the outbreak, were determined. These were then used as predictor variables in a conditional logistic regression analysis. All analyses were performed using STATA 9.2.

RESULTS

There were 111 identified outbreaks associated with consumption of drinking water between 1910 and 1999 in England and Wales (Table 1). The location was not known for 13 outbreaks, the date was not known for seven and there was no available rainfall data for two outbreaks. Following these 22 exclusions the number of outbreaks included in the study was restricted to 89 (Figure 1).

Table 1 │ Outbreaks linked to drinking water, England and Wales 1900–1999

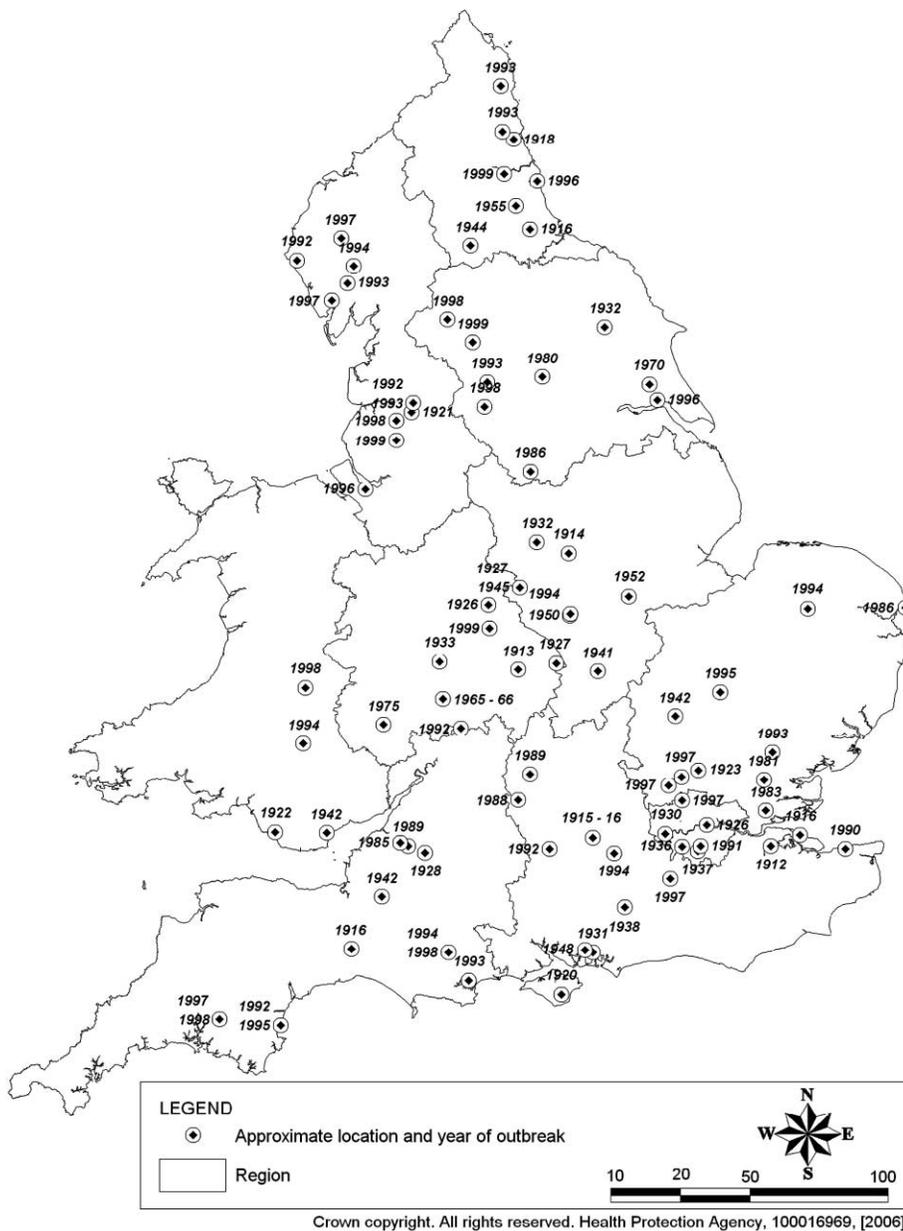| Infective agent	1910s	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	Total
Campylobacter								5	9	14
Cryptosporidium								7	30	37
Giardia								1	2	3
Mixed	2		2		1	1			3	9
Other	1			1	1		1	4	2	10
Typhoid/Paratyphoid	6	9	4	8	2		2			31
Unknown		2	2						3	7
Total	9	11	8	9	4	1	3	17	49	111

Figure 1 │ Site and year of drinking water related outbreaks 1900–1999.

The exact date of onset (day of month) was not known for 18 of the 89 outbreaks.

There was a significant association between rainfall in outbreak as compared to the control years for all four one week periods before the outbreak date and a significant association between cumulative rainfall of over 40mm in the previous 7 days (days 1–7) and outbreaks ($p = 0.001$)

(Table 2). For all four one week periods there was an excess of outbreaks compared to controls in the >40 mm rainfall group. For the days 8 to 28 there was a significant excess of low weekly rainfall (less than 20 mm per week) in the outbreak compared to the control years ($p = 0.002$). There was an association between the cumulative rainfall and outbreaks, with cumulative rainfall greater than ten and less

Table 2 │ Number of case and control years showing the cumulative rainfall in four categories ($n = 89$) (First method)

Period before outbreak (days)	Rainfall category (mm)				p value[*]
	0 to10	>10 to < = 20	>20 to < = 40	>40	
1–7 (outbreak)[†]	34	25	17	13	
1–7 (control)[†]	36	33	18	2	
Odds Ratio	1.00	0.75	1.14	NE	
95% CI		0.35, 1.16	0.49, 2.65	NE	$p = 0.001$
8–14 (outbreak)[†]	41	20	21	7	
8–14 (control)[†]	26	45	16	2	
Odds Ratio	1.00	0.33	0.94	2.88	
95% CI		0.15, 0.70	0.34, 2.56	0.29, 28.1	$p = 0.006$
15–21(outbreak)[†]	47	15	20	7	
15–21 (control)[†]	34	39	16	0	
Odds Ratio	1.00	0.29	1.08	NE	
95% CI		0.13, 0.63	0.44, 2.65	NE	$p < 0.001$
22–28 (outbreak)[†]	49	21	10	9	
22–28 (control)[†]	29	44	14	2	
Odds Ratio	1.00	0.25	0.46	2.56	
95% CI		0.11, 0.55	0.17, 1.27	0.51, 12.9	$p < 0.001$
8–28 (outbreak)[‡]	8	19	20	42	
8–28 (control)[‡]	2	7	39	41	
Odds Ratio	1.00	0.59	0.11	0.20	
95% CI		0.06, 6.39	0.01, 0.95	0.03, 1.68	$p = 0.002$

[*]Significance measured by conditional logistic regression.
[†]Rainfall measured over seven days.
[‡]Rainfall measured over three weeks.
NE: Not estimated due to small numbers.

than or equal to 20 millimetres having lowest risk and greater than forty millimetres the greatest risk, irrespective of the time period.

For the periods 8–14, 15–21 and 22–28 days prior to the outbreak, those outbreaks with a known start date had lower risk for cumulative rainfall greater than ten and less than or equal to twenty millimetres compared to those outbreaks with unknown start dates. There was a greater risk when the source was groundwater, for rainfall greater than twenty and less than or equal to forty millimetres for the periods 15–21 and 22–28 days prior to the outbreak compared to source of water which was surface derived. There was a higher risk in spring compared to other seasons and between private and mains water supply for the 15–21 and 22–28 days prior to the outbreak. However, including these significant interactions together in a model for the

respective time periods prior to the outbreak resulted in all of the interactions being highly non-significant ($p > = 0.2$), suggesting that there is confounding. Hence it can be concluded that there is no significant effect modification occurring in this study, including from outbreaks where the exact date of onset was not known.

The rainfall exceedance during the period before the start of outbreaks was calculated (Figure 2(a)). The odds ratio comparison of rainfall exceedance between matched outbreak and non outbreak years was not statistically significant for any single day before the outbreak (where lower 95% CI exceeded 1). However, the cumulative odds ratio was above 1 for all days up to day 30 and almost reached statistical significance for days 6 and 10 (Figure 2(b)). Retesting using 8 days as the starting point for the cumulative analysis indicated that although the results were all
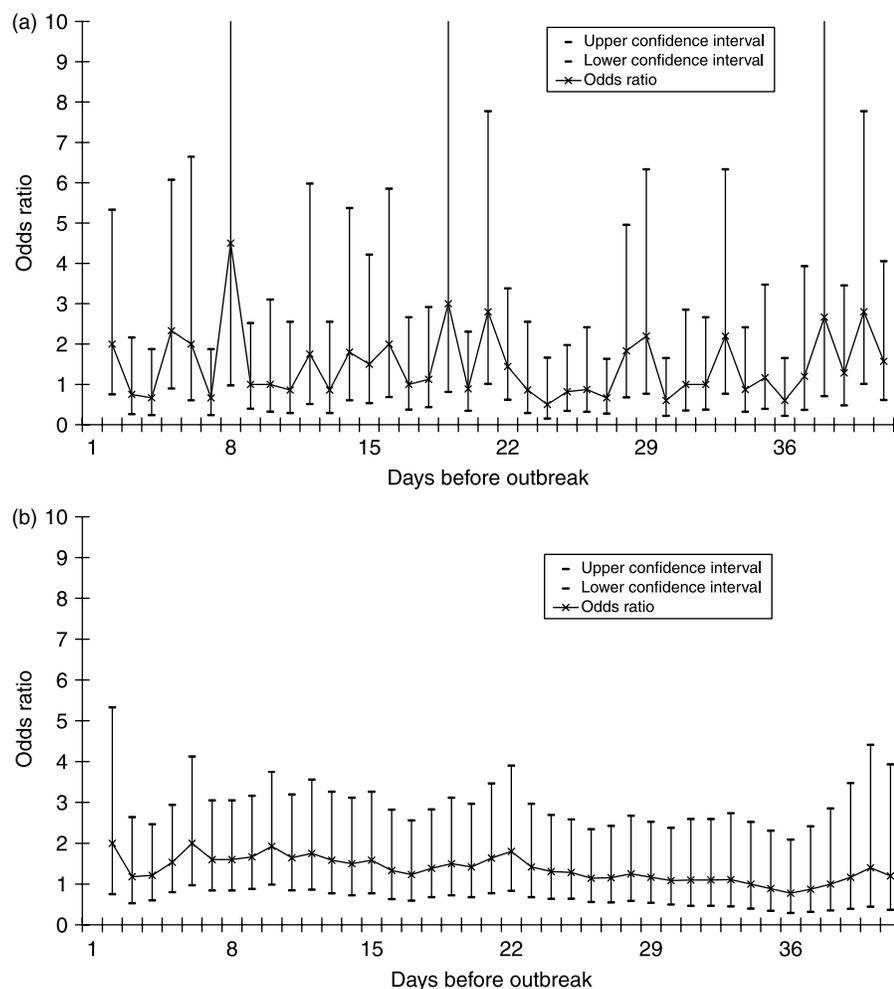
Figure 2 | Estimated odds ratios and their 95% confidence intervals for exceedance event by day before outbreak. (Second method) (a) matched comparison of case and control years. (b) cumulative exceedence based on matched case and control pairs.

above 1 for days up to 25 the main exceedance over control years occurred within the first ten days.

DISCUSSION

Outbreaks associated with contaminated drinking water can be as a result of exceptional weather conditions (Auld *et al.* 2004; Schuster *et al.* 2005). The results from this study show that 30.3% of outbreaks had less than 20 mm rainfall in the three weeks prior to the week before the outbreak, compared to 10.1% in control years. In addition 14.6% of outbreaks had a period of rainfall in excess of 40 mm compared to 2.2% of control weeks. These data imply that

the attributable fraction of outbreaks associated with a sustained period of low rainfall is 20% compared to a period of heavy rainfall of 10%. Because the dataset used in this study is an historical one the results may not reflect current risks, because water companies may have adopted treatment strategies which limit the problems associated with heavy rainfall.

The mechanisms whereby low rainfall might contribute to outbreaks through increased contamination of source waters are straightforward (Nichols *et al.* 2006). These include an increased percentage of sewage effluent in rivers as rainfall decreases, the opening up of water flow channels as the water table drops, allowing groundwater to become contaminated with surface water, the wash out of storm

drains, the wash of animal faeces from fields, cracked ground reducing the filtering effect of earth and the low water content of soil making run-off more likely. Excess extreme rainfall events result from some predictions of climate change (Hunter 2003), and the association with low rainfall followed by high may be more significant.

In testing whether waterborne outbreaks are related to prior rainfall, methodology is important. In preparation for the multiplicity of options for analysis it was important to avoid multiple testing which would increase the likelihood of a statistically significant result occurring by chance. Comparing rainfall in the 90 days before the outbreak with that measured for the same 90 day period for the previous 5 years was a useful approach, because it controls for geographic and seasonal differences in rainfall. However, seasonal differences in rainfall are not regular and comparing a day's rainfall with that from the same day in previous years shows marked differences. It was not clear in designing the protocol what the best criterion to use as a measure of rainfall was. Would an association be more likely when extreme rainfall events were recorded and what cut off point would best represent such events? Would daily, cumulative weekly, average or median rainfall be better to test? One of the approaches used was to create an exceedance based on data for rainfall for specific sites (e.g. using one year's worth of control data) and adjust the criteria for an exceedance for each site. The perceived advantage of this approach was that it controlled for geographical and temporal differences across the country in general and higher rainfall in the North West in particular The negative side of this approach was that noise could be introduced into the analysis through using too small a dataset on which to base the exceedance (only 89 outbreaks with data). Another approach counted the exceedances per week. In practice neither of these approaches offered an advantage over weekly total rainfall data split into four groups, which gave results that seemed realistic, robust and useful. Although there was a cumulative exceedance odds ratio of greater than one for all days from two to 30, it was not a useful way of measuring low rainfall.

Heavy rainfall has been associated with public supply and small system waterborne outbreaks (Willocks *et al.* 1998; Curriero *et al.* 2001; Miettinen *et al.* 2001; Hunter 2003;

Said *et al.* 2003; Auld *et al.* 2004; Thomas *et al.* 2006) and changes in climate may influence future risks (Rose *et al.* 2001; Charron *et al.* 2004). For drinking water providers the results emphasise that specific climatic conditions have increased the risk of outbreaks occurring and suggest that interventions focussed on periods of low rainfall as well as following heavy rain might be useful in formulating Water Safety Plans.

This case-crossover study of drinking water related outbreaks in England and Wales during the 20th century found a significant association with excess rainfall over the previous week and low rainfall in the three weeks before the week of the outbreak.

ACKNOWLEDGEMENTS

REFERENCES

Anonymous, Ministry of Health 1938 Water supplies and waterborne diseases. *Annual Report of Chief Medical Officer of the Ministry of Health for the year 1937*.

Atherton, F., Newman, C. P. & Casemore, D. P. 1995 An outbreak of waterborne cryptosporidiosis associated with a public water supply in the UK. *Epidemiol. Infect.* **115**, 123–131, Aug.

Auld, H., MacIver, D. & Klaassen, J. 2004 Heavy rainfall and waterborne disease outbreaks: the Walkerton example. *J. Toxicol. Environ. Health A.* **67**, 1879–1887.

Bridgman, S. A., Robertson, R. M., Syed, Q., Speed, N., Andrews, N. & Hunter, P. R. 1995 Outbreak of cryptosporidiosis associated with a disinfected groundwater supply. *Epidemiol. Infect.* **115**, 555–566.

Charron, D., Thomas, M., Waltner-Toews, D., Aramini, J., Edge, T., Kent, R., Maarouf, A. & Wilson, J. 2004 Vulnerability of waterborne diseases to climate change in Canada: a review. *J. Toxicol. Environ. Health A.* **67**, 1667–1677.

Curriero, F. C., Patz, J. A., Rose, J. B. & Lele, S. 2001 The association between extreme precipitation and waterborne disease outbreaks in the United States, 1948–1994. *Am. J. Public Health* **91**, 1194–1199.

Furtado, C., Adak, G. K., Stuart, J. M., Wall, P. G., Evans, H. S. & Casemore, D. P. 1998 Outbreaks of waterborne infectious intestinal disease in England and Wales, 1992–5. *Epidemiol. Infect.* **121**, 109–119.

Galbraith, N. S., Barrett, N. J. & Stanwell-Smith, R. 1987 Water and disease after Croydon: a review of water-borne and water-associated disease in the UK 1937–86. *J. Inst. Water Environ. Manage.* **1**, 7–21.

Hunter, P. R. 2003 Climate change and waterborne and vector-borne disease. *J. Appl. Microbiol.* **94**(Suppl), 37S–46S.

Miettinen, I. T., Zacheus, O., von Bonsdorff, C. H. & Vartiainen, T. 2001 Waterborne epidemics in Finland in 1998–1999. *Water Sci. Technol.* **43**, 67–71.

Nichols, G. L., Chalmers, R. M., Lake, I., Sopwith, W., Regan, M., Hunter, C. A., Grenfell, P. Harrison, F. & Lane, C. 2006 *Cryptosporidiosis: A Report on the Surveillance and Epidemiology of Cryptosporidium Infection in England and Wales*. Drinking Water Directorate Contract Number DWI 70/2/201. Drinking Water Inspectorate. http://www.fwr.org/waterq/dwi0849.htm

Rose, J. B., Epstein, P. R., Lipp, E. K., Sherman, B. H., Bernard, S. M. & Patz, J. A. 2001 Climate variability and change in the United States: potential impacts on water- and foodborne diseases caused by microbiologic agents. *Environ. Health Perspect.* **109**(Suppl 2), 211–221.

Said, B., Wright, F., Nichols, G. L., Reacher, M. & Rutter, M. 2003 Outbreaks of infectious disease associated with private drinking water supplies in England and Wales 1970–2000. *Epidemiol. Infect.* **130**, 469–479.

Schuster, C. J., Ellis, A. G., Robertson, W. J., Charron, D. F., Aramini, J. J., Marshall, B. J. & Medeiros, D. T. 2005 Infectious disease outbreaks related to drinking water in Canada, 1974–2001. *Can. J. Public Health.* **96**, 254–258.

Smith, A., Reacher, M., Smerdon, W., Adak, G. K., Nichols, G. & Chalmers, R. M. 2006 Outbreaks of waterborne infectious intestinal disease in England and Wales, 1992–2003. *Epidemiol. Infect.* **134**, 1141–1149.

Tillett, H. E., de Louvois, J. & Wall, P. G. 1998 Surveillance of outbreaks of waterborne infectious disease: categorizing levels of evidence. *Epidemiol. Infect.* **120**, 37–42.

Thomas, K. M., Charron, D. F., Waltner-Toews, D., Schuster, C., Maarouf, A. R. & Holt, J. D. 2006 A role of high impact weather events in waterborne disease outbreaks in Canada, 1. *Int. J. Environ. Health Res.* **16**, 167–180.

Willocks, L., Crampin, A., Milne, L., Seng, C., Susman, M., Gair, R., Moulsdale, M., Shafi, S., Wall, R., Wiggins, R. & Lightfoot, N., for the Outbreak Investigation Team, 1998 A large outbreak of cryptosporidiosis associated with a public water supply from a deep chalk borehole. Outbreak investigation team. *Commun. Dis. Public Health* **1**, 239–243.