Epidemiol. Infect. (1998), **120**, 43–54. Printed in the United Kingdom

Surveillance data for waterborne illness detection: an assessment following a massive waterborne outbreak of *Cryptosporidium* infection

M. E. PROCTOR[1],* K. A. BLAIR[2] AND J. P. DAVIS[1]

[1] *Bureau of Public Health, Wisconsin Division of Health, Madison, WI*
[2] *City of Milwaukee Department of Health, Milwaukee, WI*

(*Accepted 29 September 1997*)

SUMMARY

Following the 1993 Milwaukee cryptosporidiosis outbreak, we examined data from eight sources available during the time of the outbreak. Although there was a remarkable temporal correspondence of surveillance peaks, the most timely data involved use of systems in which personnel with existing close ties to public health programmes perceived the importance of providing information despite workload constraints associated with an outbreak. During the investigation, surveillance systems which could be easily linked with laboratory data, were flexible in adding new variables, and which demonstrated low baseline variability were most useful. Geographically fixed nursing home residents served as an ideal population with non-confounded exposures. Use of surrogate measurements of morbidity can trigger worthwhile public health responses in advance of laboratory-confirmed diagnosis and help reduce total morbidity associated with an outbreak. This report describes the relative strengths and weaknesses of these surveillance methods for community-wide waterborne illness detection and their application in outbreak decision making.

INTRODUCTION

Since 1984 at least 13 community-wide outbreaks of *Cryptosporidium* infection acquired through public drinking water systems have been reported in the United States [1–7], England [8–13], Scotland [14], and Canada [15]. In response to public concerns about the safety of community drinking water systems, partnerships have been formed between water utility companies, local government and public health officials to discuss methods to ensure the quality of the local water supplies. Following the 1993 Milwaukee *Cryptosporidium* outbreak which occurred during late March and early April, we examined surveillance data from eight sources which were available during the time of the outbreak. Data from five of the sources were used during the outbreak investigation (water treatment plant effluent turbidity logs, clinical laboratory diagnoses, nursing home diarrhoeal rates, hospital emergency room logs, and random digit dialing telephone surveys) and data from three other sources (water utility complaint logs, school absentee logs, and pharmacy sales of over-the-counter anti-diarrhoeal drugs) were examined following the outbreak.

This report summarizes relative strengths and weaknesses of these surveillance data and related methods for routine community-wide waterborne illness detection. The feasibility and timeliness in obtaining these types of data during a waterborne outbreak and our experience in applying these data to outbreak decision making are discussed.

* Author for correspondence: Mary E. Proctor, Communicable Diseases Section, Wisconsin Bureau of Public Health, 1414 E. Washington Avenue, Room 241, Madison WI 53703-3044, USA.

METHODS OF INVESTIGATION

The methods for obtaining data from water treatment plant effluent turbidity logs, clinical laboratory testing, and random digit dialing telephone surveys during the 1993 Milwaukee cryptosporidiosis outbreak are described in detail elsewhere [5] and summarized briefly here. We include detailed methods for the other five surveillance mechanisms used during or evaluated following the outbreak investigation but not previously described. While data using each of the surveillance techniques exist for different intervals, the study interval for each method in this report is 1 March, 1993 through 30 April, 1993 unless otherwise indicated. Outbreak awareness occurred on 5 April, 1993 when the City of Milwaukee Health Department (MHD) staff were informed of diarrhoeal illness associated with widespread school and vocational absenteeism, shortages of antidiarrhoeal medications at some pharmacies, and shortages of bacterial enteric culture media at some hospital laboratories. Early notification of the MHD occurred during calls from municipal health departments in Milwaukee County and print media. The MHD notified the Wisconsin Division of Health (DOH) of the outbreak on 5 April. On 7 April when the DOH investigation team arrived on-site, (1) retrospective routinely collected data were requested from the water utilities and clinical laboratories, (2) new nursing home and emergency room surveillances were initiated with requests for retrospective data, and (3) prospective data were requested from all four data sources as follows:

Water treatment plant effluent turbidity logs

The Milwaukee Water Works (MWW) staff routinely monitor plant effluent (treated water leaving the plant) turbidity expressed in nephelometric turbidity units (NTU) from the Linnwood Water Treatment (North) and the Howard Avenue Water Treatment (South) Plants which provide Milwaukee municipal drinking water (see Fig. 1). Logs of plant effluent turbidity were received from the MWW and examined during the outbreak investigation [5]. Daily maximum plant effluent turbidity is plotted for the study period.

Clinical laboratory diagnosis

Staff of 14 clinical microbiology laboratories which conduct parasitic examination of stool specimens in the Milwaukee vicinity provided the number of *Cryptosporidium* tests performed in their laboratory to the MHD retrospectively for the period 1 March through 7 April and prospectively for the period 8 April through 30 April and the number of positive tests during these intervals. An acid fast staining technique was used in 13 of these laboratories and an immunofluorescent stain was used in one laboratory [5]. Laboratory confirmed cases reported during the study period were plotted by laboratory test date.

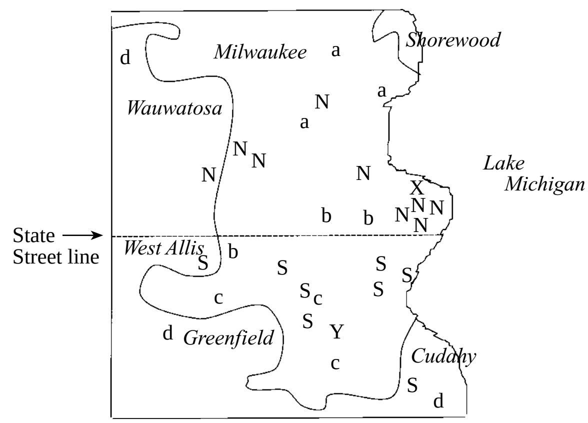

Fig. 1. Relative location of the Milwaukee Water Works North (X) and South (Y) water treatment plants, the 9 north (N) and 8 south (S) nursing homes participating in the outbreak surveillance, and the 12 emergency rooms providing data during the outbreak which were located in the midzone (b), north (a) or south (c) of State Street in Milwaukee county, or were located outside (d) Milwaukee county in the 4 surrounding communities.

Nursing home diarrhoeal rates

On 7 April, 17 nursing homes (NHs) listed in the Milwaukee area telephone directory Yellow Pages were selected as representative of geographically distinct locations within Milwaukee county and four surrounding communities (see Fig. 1). When both water treatment plants are operating, the nine northern NHs received water predominately from the North Plant and the eight southern NHs received water predominately from the South Plant. An infection control practitioner at each NH performed retrospective chart review to determine the number of residents with diarrhoea for each day during the period 1 March through 7 April and provided prospective twice weekly updates on diarrhoea episodes in residents at their facility from 8 April through 30 April. For purposes of the surveillance, diarrhoea was defined as three or more loose stools in a 24 h period. A table providing date and number of residents with new onset of diarrhoea was facsimile transmitted twice weekly to the MHD for entry into a spreadsheet used to calculate diarrhoea rates based on nursing home census that week. Prevalence of diarrhoea was compared between residents of north and south NHs for the study period.

https://doi.org/10.1017/S0950268897008327 Published online by Cambridge University Press

Hospital emergency room logs

On 7 April, 12 emergency rooms in Milwaukee county and four surrounding communities were selected for surveillance to include three hospitals from each of four geographically distinct areas (Fig. 1) delineated by municipal water source: southern region (South Plant), northern region (North Plant), midzone (both plants), and areas not receiving MWW water (outside the service area). An infection control practitioner at each selected hospital provided retrospective data on emergency room (ER) visits for the period 1 March through 7 April and prospective data for the period 8 April through 30 April. Data included total number of ER visits and GI-related visits (chief complaint of diarrhoea, gastroenteritis, vomiting, dehydration or 'stomach flu') per day and were facsimile transmitted twice weekly to the MHD for entry into a spreadsheet. Data were used to calculate the daily proportion of total ER visits which were GI-related in different regions of the Milwaukee vicinity (north or south of State Street or outside the MWW service area) during the study period.

Random digit dialing telephone survey

To determine the magnitude and monitor progress of the outbreak, four random digit dialing (RDD) telephone surveys were conducted by the DOH and the Wisconsin Survey Research Laboratory among residents in the Milwaukee vicinity. We include data from a RDD survey conducted between 28 April and 2 May [5]. Reported date of onset of watery diarrhoea during the period from 1 March through 28 April was plotted for 436 individuals with illness meeting the clinical case definition during this telephone survey.

Milwaukee utility consumer complaint logs

Consumer complaint logs maintained at both MWW water treatment plants include data on date of call, name of caller, address, telephone number, nature of the complaint regarding water quality, explanation provided to the caller by MWW staff, and plants for further action or follow-up. Data from handwritten complaint logs maintained at both water treatment plants relevant to the study period were evaluated after the outbreak for trends [16].

School absentee logs

Student absentee data for the period 1 March through 16 April were solicited from public and private schools in the Milwaukee vicinity. A standardized form requested name, address, zip code, total enrollment of the school, and the number of students absent during each day of the study period. Data were evaluated after the outbreak for usefulness as a surveillance tool.

Pharmacy surveillance

Shortly after outbreak recognition, the MHD was contacted by a local pharmacist with a computerized inventory system that could track monthly sales of over-the-counter drugs. A computer log of anti-diarrhoeal drug sales for March and April 1993 received from this pharmacist appeared to closely match the epidemic curve of human illness observed during the outbreak. In January 1994 a questionnaire sent to Milwaukee vicinity member pharmacies of the Wisconsin Independent Pharmacy Association asked pharmacists if they had the ability to monitor over-the-counter antidiarrhoea drug sales and would be willing to provide this information to the MHD on a weekly basis as a potential disease surveillance tool; the response was poor. These data were available only as monthly totals and are presented as total number of over-the-counter sales of three different antidiarrhoeal medications (Imodium®, Pepto Bismol®, and Kaopectate®) by month of sale during 1993 by a single pharmacy located in the area served by the South Plant. Data from this pharmacy included three different liquid and tablet sizes of Imodium, 5 regular and 3 maximum strength liquid Pepto-Bismol sizes, and 8 different regular and 2 maximum strength liquid Kaopectate sizes. Comparable sales data from this same pharmacy by month for 1994 and through July 1995 are included for comparison.

Statistical methods

Baseline mean and standard error for outcome variables for each surveillance system were calculated from pre-outbreak data recorded during 1 March and 15 March using Epi-Info, version 6.02, CDC. Signal ratio was defined as the ratio between the peak signal value and the mean pre-outbreak outcome variable baseline value measured between 1 March and 15 March 1993 for each surveillance system.

RESULTS

Surveillance data trends during the outbreak investigation period

Serial plots of data from seven of the surveillance

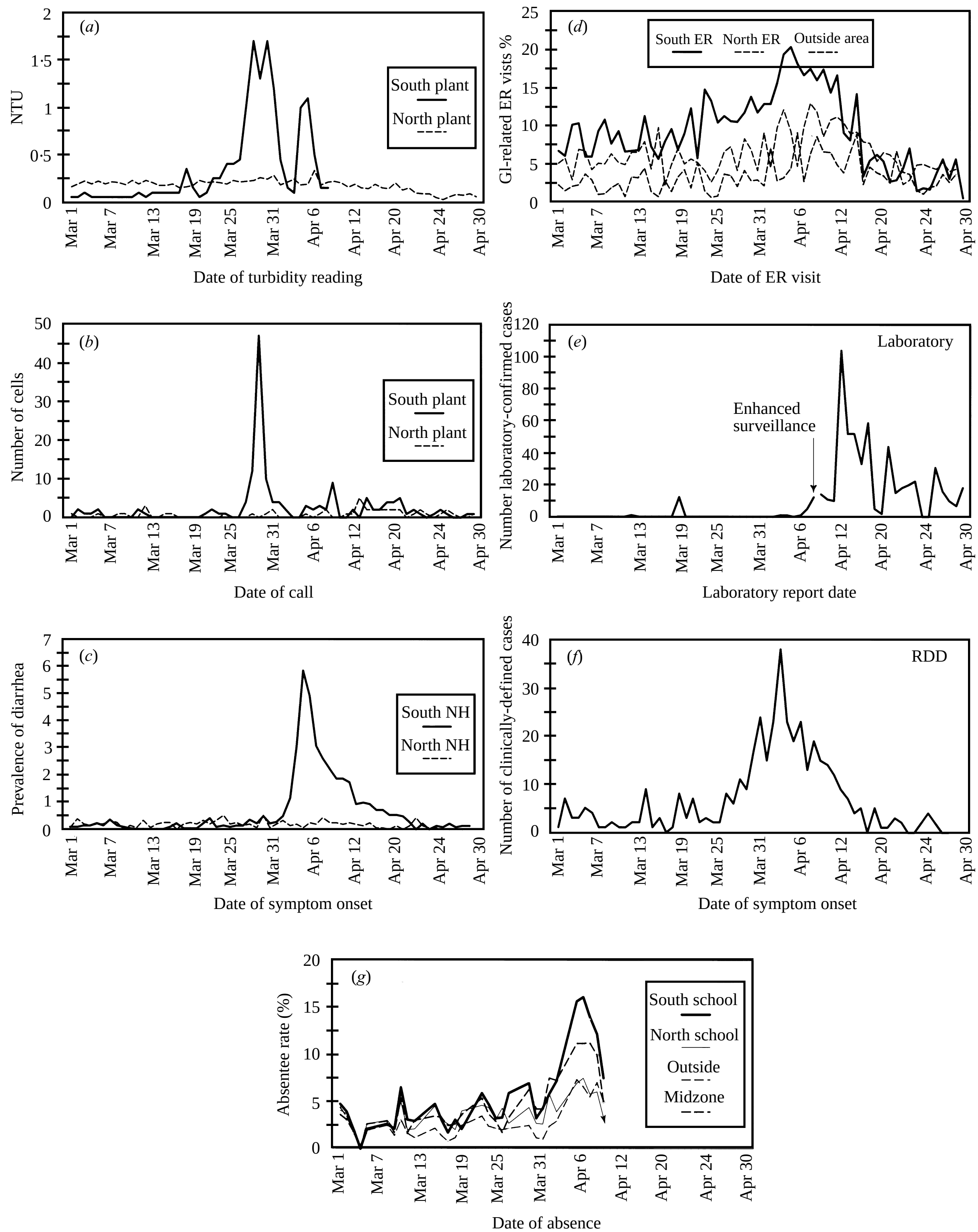

Fig. 2. Comparison of outcome variable trends between 1 March and 30 April among seven surveillance systems available at the time of the 1993 Milwaukee *Cryptosporidium* outbreak investigation. (a) Daily maximum water treatment plant effluent turbidity by treatment plant; (*b*) daily number of water utility customer complaints by treatment plant; (*c*) daily nursing home (NH) diarrhoea prevalence rates per 100 residents by geographic location of nursing home in Milwaukee Water Works (MWW) service area; (*d*) percentage of GI-related visits among total visits to hospital emergency rooms (ERs) by geographic location of the ERs in the MWW service area; (*e*) number of clinical laboratory diagnoses of *Cryptosporidium* infection by laboratory report date; (*f*) daily number of cases of watery diarrhoea clinically defined during a random digit dialing survey; (*g*) daily school absentee rates by location of school in the MWW service area.

https://doi.org/10.1017/S0950268897008327 Published online by Cambridge University Press

sources (Fig. 2) demonstrate the difference between south and north-specific data during the outbreak period and allow comparison of surveillance data peaks relative to each other. The first South Plant finished water turbidity peak followed a gradual daily increase in turbidity and occurred on 28 March (Fig. 2*a*) followed by peak southern Milwaukee MWW customer complaints on 29 March (Fig. 2*b*), peak date of onset of watery diarrhoea on 3 April among Milwaukee area residents contacted by RDD (Fig. 2*f*), peak southern ER GI-related visits on 4 April (Fig. 2*d*), peak southern NH diarrhoeal rates on 5 April (Fig. 2*c*), peak absentee rates among south schools on 6 April (Fig. 2*g*), and peak number of laboratory diagnosed cases of *Cryptosporidium* infection on 12 April (Fig. 2*e*).

Surrogate measures of morbidity in early detection of waterborne outbreaks

A quantitative comparison of statistical parameters and timeliness in identifying peaks associated with the seven surveillance systems is summarized in Table 1. While the date of peak signal heights occurred between 28 March and 12 April (column one), these peak dates were less important than the dates we physically received data from different surveillance systems and could rapidly identify trends in disease occurrence (column two). For purposes of this report, timeliness measurements (column two) are provided relative to the initial peak in South Plant treated water turbidity on 28 March (reference). Knowledge of the customer complaint and nursing home diarrhoea rate peaks were most timely with a lag of 2 and 11 days beyond the reference date, respectively. Knowledge of ER GI-related visit and laboratory diagnosis peaks followed 15 days beyond the reference date. Knowledge of surrogate morbidity peaks using RDD survey (35 days) and school absence logs (64 days) were the least timely in providing information.

South Plant complaint logs were very timely in detecting the peak signal height and had low mean baseline (pre-outbreak) signal, low mean baseline variability (standard error mean), and a high signal ratio (peak signal height to mean pre-outbreak outcome variable measurement). Although clinical laboratory data provided the highest signal ratio and had low mean baseline signal and variability, the 15 day lag in learning about the peak limited the utility of this surveillance method for early outbreak detection. Among the remaining four surveillance methods, nursing home diarrhoea prevalence rates provided the largest signal ratio as well as the lowest mean baseline signal and mean baseline variation and were very timely. High mean baseline variation values for ER GI-related visits, school absentee rates and RDD clinically-defined cases made earlier identification of unusual disease occurrence more difficult to recognize with certainty using these data.

Strengths and weaknesses of surveillance data sources during the Milwaukee outbreak

Acceptability of providing and timeliness in receiving data along with other attributes (simplicity, flexibility, sensitivity, predictive value positive, representative-ness, and usefulness) for evaluating surveillance systems [17] were the basis for comparing strengths and weaknesses of the data sources in relation to our outbreak investigation (Table 2).

Simplicity, timeliness and acceptability

During the initial phases of the investigation, we sought data for their simplicity because they were either routinely monitored (treatment plant effluent turbidity, clinical laboratory diagnosis) or we thought they could be rapidly extracted from routinely collected information into outbreak-specific data (NH diarrhoeal illness, ER GI-related visits, school absentee data). With the exception of the water treatment plant effluent turbidity readings and clinical laboratory diagnosis which were computer generated, all data sought was provided in hand written tabular form during the outbreak but this did not limit use of the data for assimilation into epidemiologic decision making. The RDD survey was least simple to perform and was both labour intensive and costly.

The most timely data involved the use of systems in which personnel with existing close ties to public health programmes perceived the importance of providing the data despite workload constraints associated with the outbreak itself (acceptability). For example, NH, ER, and laboratory data were readily provided twice weekly by facsimile transmission on outbreak-specific one page summary sheets for easy entry into a spreadsheet by the outbreak investigators. Providing school absentee data was not perceived by school administrators as good use of staff resources and the data, much of it incomplete, reached us a month after the outbreak was over.

Table 1. *Comparison of the timeliness and variability of surveillance data available at the time of the 1993 Milwaukee cryptosporidiosis waterborne outbreak*

Surveillance system	Date of peak signal height	Timeliness* in learning about peak in days (date)	Peak† signal height (a)	Mean baseline signal‡ (b)	Mean baseline variation§ (std. error mean)	Signal ratio (a/b)‖
South water treatment	28 Mar	Ref.	1·7	0·07	0·01	24·3
South water utility	29 Mar	2 (30 Mar)	47·0	0·60	0·21	78·3
RDD detected cases	3 April	35 (2 May)	38·0	2·87	0·63	13·2
GI-related south ER	4 Apr	15 (12 Apr)	20·4	7·90	0·53	2·6
South nursing home	5 Apr	11 (8 Apr)	5·9	0·09	0·02	65·6
South school absentee	6 Apr	64 (2 June)	16·1	3·15	0·49	5·1
Clinical laboratory	12 Apr	15 (12 Apr)	104·0	0·07	0·07	1485·7

* Number of days between south water treatment plant effluent turbidity peak on 28 March, 1993 (referent) and date we learned about peak signal for other surveillance methods.

† Maximum daily outcome variable signal height detected by given surveillance system between 1 March and 30 April, 1993.

‡ Mean daily outcome variable measured between 1 March and 15 March, 1993 (pre-outbreak) for given surveillance system.

§ Standard error of the mean daily outcome variable measured between 1 March and 15 March, 1993 (pre-outbreak).

‖ Ratio of peak signal height (a) and mean pre-outbreak outcome variable baseline measurement (b).

Flexibility, sensitivity, and representativeness

Some surveillance methods (NH diarrhoeal illness and RDD surveys) were more flexible during the investigation because additional questions could be added to those routinely collected as the need arose. Water utility complaint logs would have also been flexible if they had been employed at the time of the outbreak investigation. Some data sources collected fixed information (water treatment plant effluent turbidity, school absentee data, hospital ER log data, diagnostic laboratory data, and pharmacy sales) and could not be altered in a rapid enough manner to provide needed epidemiologic information.

Surveillance data sources differed significantly in their representativeness or the sensitivity in monitoring the population of interest. For example, in comparison with RDD surveys where a specific case definition could be developed, other surveillance data were less sensitive because relatively few people make utility company customer complaints, individual filter monitoring would be more sensitive than monitoring total water treatment plant effluent turbidity, and generally only those who are most ill will visit a hospital emergency department and will have laboratory tests collected by their physician for testing. We found that effluent turbidity data, customer complaint logs and NH residents were representative of subpopulations at risk by specific water source, that ER patients and NH residents could be evaluated by specific zip code, and that the NH population was representative of an immobile population. Finally, while RDD and ER surveillance captured a wide range of age groups, school absentee data and NH illness provided information on age-specific subpopulations; this could be a strength or a weakness as a surveillance tool depending on the aetiologic agent and the immune status of that subpopulation to that agent.

Usefulness and predictive value positive

Individual surveillance methods had varying perceived utility during different phases of the outbreak investigation and in meeting a broad range of surveillance objectives. Telephone calls to the MHD on 5 April about widespread school and workplace absenteeism and newspaper reports of shortages of antidiarrhoeal medications were the first indication that something unusual was occurring. These calls prompted inquiries by the MHD to emergency departments and clinical laboratories regarding specific symptoms and diagnoses possibly associated with these anecdotal reports. Knowledge on 5 April of high South compared to North Plant effluent water turbidity levels provided the first clue that the outbreak was waterborne and prompted the boil water advisory. The sixfold differential in diarrhoea prevalence rates between south and north NH residents (data received and calculated on 8 April) combined with differences in finished water turbidity between South and North Plants strengthened the hypothesis regarding a South Plant point source waterborne outbreak and was instrumental in the decision to temporarily close the

South Plant on 9 April. One RDD survey [5] provided the best estimate of the magnitude of the outbreak and the other RDD surveys provided information regarding the end of the outbreak.

Finally, although we did not perform a mathematical evaluation, the source of data with the lowest frequency of false positive case reports (high predictive value positive [PVP]) would be diagnostic laboratory testing. Elevated effluent water turbidity and increased customer complaints may not correlate with actual morbidity in the community, and increased absentee rates in schools, elevated diarrhoeal rates in nursing homes, or increased emergency department visits may be due to non-illness reasons (school field trips, vacations), increased GI problems among NH residents, or other GI agents circulating in a community, respectively, and would presumably have lower PVP. During the outbreak we instituted a courier system between participating NHs and the MHD diagnostic laboratory. Whenever a NH resident experienced new onset of diarrhoea, the MHD was contacted and the courier was dispatched to pick up the specimen for delivery to the laboratory. This link of surrogate NH surveillance with rapid laboratory diagnosis increased the PVP of NH surveillance during the outbreak. We found it difficult during the outbreak investigation to link names of those seen in the ER or absent from school (because the data was not electronically available) with laboratory confirmed case data, thus reducing the PVP of these surrogate surveillance systems.

Surveillance for community-wide waterborne outbreak detection in Milwaukee since the outbreak

Since the 1993 outbreak, Milwaukee has actively pursued collection of ongoing surveillance data from a variety of data sources including clinical laboratory diagnosis, water treatment plant effluent turbidity and particle counts, individual filter bed turbidity (not used during the outbreak), prevalence of diarrhoea among residents in sentinel NHs (with courier assisted laboratory linkage), and sale of antidiarrhoeal agents at sentinel pharmacies. Since September 1993, the continuous provision of nursing home, clinical laboratory, and water quality data has been readily sustained. The initial interest by 15 pharmacies to provide data has waned; a single pharmacy has provided 29 consecutive months of data (Fig. 3). In addition to the poor response rate, the main weakness of pharmacy surveillance in the Milwaukee area has been the inability to transmit the data electronically, difficulty collapsing data from different size bottles and flavours into usable information, and lack of timeliness because data is available on a monthly rather than weekly basis (Table 2).

In addition to prospective surveillance, since the outbreak Milwaukee has established two task force groups: The Interagency Clean Water Advisory Council, an executive level group which reports quarterly to elected officials, and a more technically focused group, the Water Quality Work Group. These two groups meet separately on a monthly basis and jointly on a quarterly basis. Both are staffed by representatives from local and state health communicable disease, laboratory and environmental health specialties, water utility and water regulatory representatives, waste water treatment plant staff, and public information personnel. The groups review current water treatment methodologies, assess the vulnerability of the local drinking water, determine which water testing results constitute a 'trigger event' which will necessitate follow-up and response, and identify a chain of notification if a water-related emergency occurs. This multilevel and multidisciplinary structure assures involvement and acceptance at all levels of management and provides a process for advancing ideas and change with a common voice.

The end result of the formation of these two committees has been the development of a comprehensive community action and response plan in the event of a waterborne emergency. Group members have provided leadership for the recent CDC publication [18] that includes specific 'trigger events' for various surveillance data (total coliform rule violation, surface water rule violation, water filtration breakdown, an unusual number of customer complaints about water quality, pathogens found in the finished water, increased reports of diarrhoeal illness or laboratory confirmed cases reported to local health departments), levels of response to these trigger points (no health risk suspected, health risk indeterminate, health risk suspected, boil water advisory), and notification chains regarding who will be notified (immunocompromised populations, general public, food and beverage manufacturers). Another positive result of these two task forces has been the preparation of template materials for media releases and development of protocols for use by special audiences including hospitals and clinics, renal dialysis units, nursing homes, daycare facilities, dental offices,

Table 2. Attributes [17] of surveillance data available during the 1993 Milwaukee cryptosporidiosis outbreak: applications for routine community-wide waterborne illness surveillance and outbreak decision making

Surveillance data source	Strengths	Weaknesses
Water treatment plant effluent turbidity	Simplicity (routinely monitored/summarized daily) Acceptability (readily provided by MWW) Timeliness (may be computer generated/readily available) Useful (elevated readings can stimulate enhanced effluent testing and human diagnostic laboratory testing)	Flexibility (data collected is fixed) Sensitivity (monitoring individual filter effluent turbidity is more sensitive) Predictive value positive and Representative (elevated turbidity may not correlate with presence of disease-causing organisms in treated water)
Water utility customer complaint log	Simplicity (routinely monitored) Flexibility (can ask caller additional questions) Acceptability (readily provided by MWW) Representative (monitors subpopulations at risk by specific water source) Timeliness (handwritten logs are easy to facsimile transmit) Useful (elevated complaints can stimulate additional effluent testing and human diagnostic laboratory testing)	Sensitivity (relatively few people make complaints when a problem exists) Predictive value positive (poor water quality may not correlate with morbidity)
School absentee data	Simplicity (routinely collected) Sensitivity (broad view of all absences by grade level) Timeliness (may be computer generated) Useful (stimulates looking at other surveillance patterns)	Flexibility (information collected not standard between schools; reason for absence and address not usually noted) Acceptability (a burden for staff to collate) Predictive value positive (some absences due to other etiologies or for non-illness reasons, e.g., field trips, vacation) Representative (completeness of data poor) Useful (poor linkage of student names with laboratory data)
Nursing home diarrhoeal disease data	Simplicity and Sensitivity (routinely collected) Flexibility (can select information to abstract from chart) Representative (immobile population) Timeliness (data rapidly collapsed/transmitted by ICN) Useful (elevated prevalence can stimulate laboratory testing of residents)	Acceptability (data generally must be abstracted from hand written resident records) Predictive value positive (reports of increased diarrhoeal rates above background must be coupled with laboratory confirmation because of more GI problems in nursing home residents than in the general population, including the elderly) Representative (age-specific subpopulation)
Hospital ER log data	Simplicity (routinely collected, sometimes computerized) Representative (if patient zip code collected may help identify problem geographically) Timeliness (data rapidly transmitted by ICN) Useful (elevated levels of GI-related visits can stimulate laboratory testing of clients)	Flexibility (information collect is generally fixed) Acceptability (tabulation of data considered a hardship by staff already burdened during an outbreak) Sensitivity (only most ill visit ER) Predictive value positive (other GI etiologies may confound the data) Representative (ER catchment data not necessarily representative of geographic location of hospital; represents mobile population)

Table 2. (*cont*.)

Surveillance data source	Strengths	Weaknesses
Diagnostic laboratory test data	Simplicity (routinely performed, may be computer generated/electronically transmitted) Acceptability (readily provided) Predictive value positive (identifies etiologic agent if testing available and requested) Useful (baseline data readily available if diagnostic test is widely available, clinically recognized and requested by physician)	Flexibility (address, zip code, and onset date not always noted) Sensitivity (only those most ill likely to be tested) Representative (most laboratories do not routinely test for *Cryptosporidium* as part of parasitic enterics panel unless specifically requested) Timeliness (available surveillance data lags 7 days beyond nursing home and ER surveillance data)
Households with telephones: random digit dialing surveys	Flexibility (can design specific questionnaire) Acceptability (usually good response during a community outbreak) Sensitivity (can develop a case definition) Useful (provides good estimate of the magnitude of the outbreak and clinical description during a large outbreak)	Simplicity (labor intensive, costly) Representative (captures only households with telephones) Timeliness (availability of surveillance data was delayed relative to other surveillance systems)
Pharmacy sales of antidiarrhoeal medications	Simplicity (may be computerized) Timeliness (sales above baseline surveillance levels may provide early indication of an outbreak/prompt early public health responses/inquiries) Sensitivity (may represent those with mild illness and more severe forms of illness)	Flexibility (inventory differences such as different sized bottles, flavors, and formulations may be difficult to collapse large volumes of sales data into usable information) Acceptability (poor response rate among initial participants) Predictive value positive (illness may be due to a variety of etiologies) Representative (represents mobile populations; purchase site not necessarily reflective of place of residence) Timeliness (availability of surveillance data lags unless transmitted twice weekly; data may only be available as monthly totals)

commercial establishments (restaurants, hotels, convenience stores) and public consumers of water supplies. These protocols are also incorporated in the *Cryptosporidium and Water* handbook [18].

DISCUSSION

The staggered sequence of surveillance data peaks illustrated during the outbreak period represent points of detection and intervention along the continuum of cause-effect-response which follow exposure to a potential pathogen in the community water supply. Source water which is significantly unacceptable (turbid, off colour, bad odour or taste) may elicit the response of only a small number of customers within a day of the distribution of potentially contaminated water but should trigger further investigation by designated individuals. Depending on the infectious agent, this event typically occurs one to several days before signs and symptoms can be passively monitored (e.g., logging of diarrhoeal episodes in NH notes), and much before individuals will decide (active response) they are symptomatic enough to purchase anti-diarrhoeal medication, miss work or school, or visit their physician or an emergency department. Ultimate recognition of morbidity involves laboratory confirmation which occurs last in this sequence and only if individuals visit a medical care provider, the provider orders an appropriate test and the test result is available at an accessible facility. A sensitive waterborne illness surveillance and response system should be responsive to multiple early surrogate indicators of potential morbidity.

Early identification of a potential community-wide

https://doi.org/10.1017/S0950268897008327 Published online by Cambridge University Press

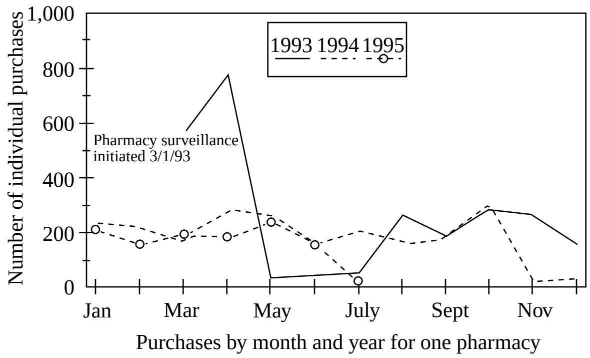

Fig. 3. Number of individual purchases of antidiarrhoeal agents at one Milwaukee pharmacy south of State Street by month, 1993–5.

waterborne problem can initiate appropriate public health interventions such as temporarily closing a water treatment plant, recommending boiling water or purchasing bottled water, and ultimately will have an impact on reducing the morbidity and mortality associated with an outbreak. Surrogate indicators of a potential waterborne outbreak, such as knowledge of increases in NH or other monitored diarrhoea prevalence, should prompt health department review of water utility water quality data and requests of local medical care providers for enhanced specimen collection and testing. Based on analysis of RDD data [5], further delays in waiting for more definitive data instead of acting on available plant turbidity data, absentee data, and early clinical laboratory results available on 7 April would have resulted in substantially more *Cryptosporidium* related morbidity. Only those surveillance methods which were collecting baseline data at the time the outbreak first occurred (treatment plant effluent turbidity, nursing home diarrhoeal disease, and to some extent emergency department visits) were useful in the first week of the outbreak investigation in establishing that an outbreak was occurring, in confirming that the outbreak was waterborne in nature, and in establishing the point source. Familiarity with expected occurrence of illness and variability of these data is essential for sensitive outbreak detection. Surrogate measures of morbidity in early detection of WBOs can provide timely, recognizable signals provided baseline measurements are available, are routinely evaluated and that appropriate individuals are notified when set trigger points are exceeded.

There are recognized weaknesses for all proposed surrogate waterborne surveillance systems as early signals of real disease morbidity. For example, in Washington, D.C. in 1993, elevations in effluent turbidity were not associated with the detection of disease-causing organisms in treated water or increased morbidity [19]. Conversely, as was noted during a recent cryptosporidiosis outbreak in Las Vegas, treated water can meet all water quality standards and still contain sufficient levels of *Cryptosporidium* oocysts to cause a community-wide outbreak [6, 7]. Increased prevalence of diarrhoea illness among NH residents or ER clients needs to be interpreted with caution because of a variety of GI problems in the elderly and because multiple human pathogens circulating in a community may complicate interpretation of the data. Enhanced laboratory testing of NH and ER clients should be initiated following identification of gastrointestinal illness beyond expected background occurrence.

Because of the uniqueness of the water distribution system in Milwaukee using two large treatment plants, the ability of individual surveillance systems to monitor subpopulations at risk by usual water source when both treatment plants are operating, especially immobile populations, was a strength during the investigation. Other waterborne disease outbreak investigations have examined illness among nursing home populations [3] and immobile populations with geographically distinct water supply sources [20] for disease outbreak detection and for point source identification. In contrast to individuals who reside in geographic locations which differ from their place of employment, schooling, or source of medical care, geographically fixed populations such as nursing home, boarding school, or correctional facility residents may serve as ideal populations with a singular water source throughout the day. Since most communities rely on a single water supply, some surveillance methods described here for outbreak investigation and prospective surveillance after an outbreak may not be directly applicable.

We note the value of alternate data sources as early warning systems which can complement laboratory diagnosis. Collin and colleagues reported the use of gastrointestinal medication sales to evaluate water supplies in Meurthe and Moselle, France during 1981 [21]. Recently, Ashendorf and colleagues reported the use of drug sales as a means of surveillance for diarrhoeal illness in New York City [22]. The Italian National Health Service records of drug prescriptions were reported by Maggini and colleagues to estimate that the prevalence of turberculosis in Italy is seven times higher than official notifications indicate [23]. Finally, several of the surrogate systems described in this report were also used in Washington, D.C. following recognition in December 1993 of inadequately filtered public drinking water with an increase in finished water turbidity to 9.0 NTU; data from a random digit dialing telephone survey and hospital emergency room, nursing home and laboratory surveillance were evaluated to determine there was no associated detectable morbidity [19].

Based on our examination of surveillance data from various sources during the outbreak period and during a sustained post-outbreak period, we recommend that communities with populations greater than 100000 whose water supply is derived from surface water, should consider developing one or more surveillance systems in addition to laboratory diagnosis to establish baseline data for those systems. Since no single set of recommended surveillances will be applicable to all communities, a combination of surveillance options should be developed locally drawing on existing disease surveillance methods and expanding with new partnerships. Those surveillance systems which can be easily linked with laboratory data, are flexible in adding new variables, and which demonstrate low baseline variability may be most helpful in detecting waterborne illness.

While the existence of baseline data is essential for recognition of an unusual occurrence, having a community-wide plan for critically and systematically evaluating these data is the second component of effective waterborne illness surveillance. We recommend that communities establish an interagency task force whose charge should be to meet regularly, develop new or adapt other community policies, set protocols for community-wide notification and response, develop and distribute educational material for the general public and high risk groups, designate specific individuals to review ongoing surveillance data, and communicate periodically in the revision of these protocols and policies. The best time to prepare for an outbreak is before it occurs. Predetermined set point values to trigger public health notification and response should be developed as part of a community-wide outbreak control protocol which should involve representatives from water utilities and water regulatory agencies, local and state health departments, and other local governmental agencies.

REFERENCES

1. D'Antonio RG, Winn RE, Taylor JP, et al. A waterborne outbreak of cryptosporidiosis in normal hosts. Ann Intern Med 1985; **103**: 886–8.
2. Gallaher MM, Herndon JL, Nims LJ, et al. Cryptosporidiosis and surface water. Am J Publ Health 1989; **79**: 39–42.
3. Hayes EB, Matte TD, O'Brien TR, et al. Large community outbreak of cryptosporidiosis due to contamination of a filtered public water supply. N Engl J Med 1989; **320**: 1372–6.
4. Leland D, McAnulty J, Keene W, Stevens G. A cryptosporidiosis outbreak in a filtered water supply. J Am Water Works Assoc 1993; **85**: 34–42.
5. 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**: 171–7.
6. Goldsmith ST, Juranek DD, Ravenholt O, et al. Cryptosporidiosis: an outbreak associated with drinking water despite state-of-the-art water treatment. Ann Intern Med 1996; **124**: 459–68.
7. Roefer PA, Monscvitz JT, Rexing DJ. The Las Vegas cryptosporidiosis outbreak. J Am Water Works Assoc 1996; **88**: 95–106.
8. Rush BA, Chapman PA, Ineson RW. A probable waterborne outbreak of cryptosporidiosis in the Sheffield area. J Med Microbiol 1990; **32**: 239–42.
9. 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.
10. Richardson AJ, Frankenberg RA, Buck AC, et al. An outbreak of waterborne cryptosporidiosis in Swindon and Oxfordshire. Epidemiol Infect 1991; **107**: 485–95.
11. Bridgeman SA, Robertson RMP, Syed Q, et al. Outbreak of cryptosporidiosis associated with a disinfected groundwater supply. Epidemiol Infect 1995; **115**: 555–66.
12. Outbreak Control Team. An outbreak of cryptosporidiosis in South and West Devon, August–September, 1995. The Public Health Medicine Department, South and West Devon Health Commission, Plymouth, England.
13. Maguire HC, Holmes E, Hollyer J, et al. An outbreak of cryptosporidiosis in South London; what value the *p* value? Epidemiol Infect 1995; **115**: 279–87.
14. Smith HV, Patterson WJ, Hardie R, et al. An outbreak of waterborne cryptosporidiosis caused by post-treatment contamination. Epidemiol Infect 1989; **103**: 703–15.
15. Pett B, Smith F, Stendahl D, Welker R. 1993. *Cryptosporidium* outbreak from an operations point of view, Kitchener-Waterloo, Ontario, spring 1993. In: Proceedings of the American Water Works Association Water Quality Technology Conference, Miami, FL, 7–11 November 1993; 1739–66.
16. Report of the Common Council Water Crises Fact Finding Committee. Milwaukee, June 1993: 1–14, 25.
17. Centers for Disease Control. Guidelines for evaluating surveillance systems. MMWR 1988; **37**(S-5): 1–17.
18. Cryptosporidium and Water: A Public Health Handbook. Atlanta, Georgia: Working Group on Waterborne Cryptosporidiosis, 1997.

19. Centers for Disease Control and Prevention. Assessment of inadequately filtered public drinking water – Washington, D.C., December 1993. MMWR 1994; **43**: 661–3.
20. Lopez CE, Dykes AC, Juranek DD, et al. Waterborne giardiasis: a community wide outbreak of disease and a high rate of asymptomatic infection. Am J Epidemiol 1980; **112**: 495–507.
21. Collin JF, Melet JJ, Morlot M, Foliguet JM. Eau d'adduction publique et gastroenterites en meurthe-et-moselle. J Frangals d'hydrobiologie 1981; **35**: 155–74.
22. Ashendorf A, Principe MA, Seeley A, et al. Watershed protection for New York City's supply. J Am Water Works Assoc 1997; **89**: 75–88.
23. Maggini M, Salmaso S, Alegiani SS, et al. Epidemiological use of drug prescriptions as markers of disease frequency: an Italian experience. J Clin Epidemiol 1991; **44**: 1299–307.