United States
Environmental Protection
Agency

Office of Water
4607

EPA 815-R-98-005
August 1998

EPA

U.S. EPA and CDC Workshop on Waterborne Disease Occurrence Studies - Summary Report

March 12-13, 1997
Atlanta, Georgia

Summary Report

U.S. Environmental Protection Agency and the Centers for Disease Control and Prevention

Workshop on Design of Waterborne Disease Occurrence Studies

Workshop Held in Atlanta, Georgia, March 12-13, 1997

Workshop Organizers: Sue Binder, Fred Hauchman , and Ron Hoffer

Report prepared by Gunther Craun

Draft Report Reviewed by Rebecca Calderon, William MacKenzie, Donald Reasoner

Table of Contents

Introduction 3

Workshop Agenda 3

Epidemic and Endemic Waterborne Disease and a National Estimate of its Occurrence 4

Epidemiological Study Designs 6

Table 1. Types of Epidemiological Studies* 6

Systematic Bias in Epidemiological Studies 8

Random Error 9

Measures of the Magnitude of Risk 10

Selection of a Study Design 10

Table 2. Advantages and Disadvantages of Epidemiological Study Designs to Estimate the National Occurrence of Waterborne Disease 13

Table 3. Possible Sources of Systematic Basis and Methods for Control in Studies to Estimate the National Occurrence of Waterborne Disease 13

Selection of a Health Outcome for Study 14

Populations to be Studied 15

Selection of a Study Site 16

Table 4 Matrix of Water Quality and Treatment/Compliance Considerations for Selecting a study Site 17

Water Exposures and Home Treatment Devices To Consider for Intervention Studies 17

Table 5 Evaluation of Bottled Water and Home Water Treatment Units for Use in Intervention Studies 19

Appendix A. Participant List 20

Appendix B. Abstracts of Presentations 25

Appendix C. Draft Request for Proposal 36

Appendix D. Causality of an Epidemiological Association 38

Introduction

The Safe Drinking Water Act (SDWA) Amendments of 1996, Section 1458(d), require the Director of the Centers for Disease Control and Prevention (CDC), and the administrator of the U.S. Environmental Protection Agency (EPA) to jointly:

> *within 2 years after the date of enactment of this section, conduct pilot waterborne disease occurrence studies for at least 5 major United States communities or public water systems; and within 5 years after the date of enactment of this section, prepare a report on the findings of the pilot studies, and a national estimate of waterborne disease occurrence.*

EPA has set aside research funds to begin this collaborative effort with CDC, and several interagency discussions have taken place. The next step in the planning of this research was to convene a work group to discuss the types of epidemiological studies that may be appropriate to estimate the prevalence of waterborne disease and identify the various issues related to their design and feasibility. Invited epidemiologists, scientists, physicians, and engineers from EPA, CDC, state public health agencies, and universities met in Atlanta on March 12 and 13, 1997 (list of participants in Appendix A). Participants were asked to address the strengths and weaknesses of various study designs and other important issues, such as the selection of geographic areas, population groups, health outcomes, and water exposures which should be studied.

Considering EPA's overall strategy for implementing the SDWA and the legislative history of the SDWA amendments, the following assumptions were made regarding the proposed pilot studies:

- Studies should be designed primarily to provide a national estimate of waterborne disease occurrence rather than address other issues surrounding waterborne disease or environmental epidemiological methods.
- Studies should focus on diseases caused by both well-known and "emerging" waterborne pathogens (bacteria, viruses, and protozoa).
- Because of recent concerns about microbial contaminants and limited funding, studies will not address effects of acute or chronic exposure to chemical contaminants in water.

Workshop Agenda

Technical presentations at the beginning of the workshop provided participants with current information to assist in the design of the proposed epidemiological studies:

- Water sources and treatment used in the United States.
- Occurrence of waterborne pathogens in water sources and tap water.
- Etiologic agents causing waterborne disease.
- Water system deficiencies associated with reported waterborne outbreaks.
- Analytical methods for detecting pathogens in water samples and clinical specimens.
- Human dose response issues for Norwalk virus and *Cryptosporidium*.

Participants were also informed about the objectives, study designs, and preliminary results of recently completed and ongoing waterborne disease epidemiological studies:
Study of risk factors including water for cryptosporidiosis in an HIV-infected and a general population cohort at Cornell Medical Center, New York City.

- An incidence case-control study of waterborne and other risks of cryptosporidiosis in HIV-infected persons in New Orleans and Los Angeles.
- Study of the sero-prevalence of *Cryptosporidium*-specific antibodies in populations using surface and groundwater systems and an NHANES national population.
- Longitudinal enteric disease studies in communities using surface water sources where water filtration facilities are being installed.
- Community intervention studies of endemic waterborne disease in Canada.

Abstracts of these presentations are included in Appendix B.

A general discussion of the fundamental questions surrounding the proposed epidemiological studies followed the technical presentations. Participants then met separately in three break-out groups for further discussions. Each group summarized their deliberations at the end of the first day. During the second day, workshop participants identified and discussed the important issues that should be considered in designing and conducting the proposed epidemiological studies. This report summarizes these issues.

Epidemic and Endemic Waterborne Disease and a National Estimate of its Occurrence

Workshop participants were reminded that not all waterborne disease results in diarrheal illness and that to fully understand waterborne disease risks requires the consideration of other symptoms. For example, nausea and vomiting are the primary symptoms of illness caused by Norwalk virus, an important waterborne pathogen, and hepatitis A; however, symptoms do not usually include gastroenteritis. In addition, many waterborne infections result in mild illness or may be asymptomatic, and ingestion is not the only route of exposure. The important route of exposure for several potential waterborne pathogens (e.g., *Mycobacteria, Legionella*) can be from the inhalation of aerosolized organisms.

Waterborne illness can be both epidemic (outbreaks) and endemic. Several workshop participants cautioned that in the United States unrecognized endemic waterborne illness may be the greater problem and this should be the focus of the proposed epidemiological studies. Little information is currently available about the endemic waterborne disease risks. It is not known whether the endemic incidence of waterborne illness is a relatively low or high, the proportion of gastroenteritis cases is relatively constant over time, or whether illness occurs as sporadic cases due to very small outbreaks. These small outbreaks may occur relatively frequently but are not likely detected by current disease surveillance systems, and epidemiological studies must be

specifically designed to determine the endemic waterborne risk. Although several studies have assessed endemic waterborne disease risks, the information is not sufficient to estimate the occurrence of waterborne disease in the United States with any degree of certainty. Waterborne outbreaks are voluntarily reported, and information about their occurrence and causes has been systematically collected by CDC and EPA since 1971. Analysis of waterborne outbreak data are primarily used to help identify important etiologic agents and water system deficiencies.

National estimates of waterborne disease occurrence have been reported by several investigators. Hauschild and Bryan (1980) estimated the annual incidence of both food- and waterborne disease in the U.S. for 1974 and 1975 to be 1,400,000 to 3,400,000 cases. Morris and Levin (1994) estimated that 1.8 million cases of waterborne disease and 1,800 deaths occur annually, and Bennett *et al.* (1987) estimated the annual incidence of waterborne disease to be more than 900,000 cases and almost 900 deaths. These estimates were not based on epidemiological studies specifically designed to obtain such an estimate, and the estimates are highly uncertain. For example the estimate prepared by Bennett *et al.* (1987), the most often cited reference, was based solely on expert opinions about the underreporting of specific enteric diseases and the proportion of cases likely to be waterborne. Workshop participants agreed that a better estimate is needed for the occurrence of waterborne illness and this estimate should be obtained from appropriately designed epidemiological studies.

To estimate the national occurrence of waterborne disease requires two components -- the population attributable risk (PAR) of illness due to drinking water exposures and the incidence of acute gastroenteritis (AGI) or specific enteric diseases from all causes[1]. Well known standard procedures are available to design and conduct a national survey of AGI incidence or other selected microbial enteric disease. However, more careful thought must be given to choosing and designing the kinds of epidemiological studies needed to estimate the PAR.

The PAR measures the relative proportion of disease attributed to water compared to other exposures, and this measure will likely differ for each etiologic agent and each geographic area (major city, region, and water system type or source). These possible differences should be considered in the proposed study designs and must be taken into account in computing the national estimate. The selection of health outcome measures, and other issues such as the consideration of symptomatic and asymptomatic cases and herd immunity must be also be considered. For example, the entire U.S. population using public water systems may not be at risk of waterborne infection because of recent infection and immunity to certain etiologic agents. This may be an especially important when evaluating the risk for waterborne cryptosporidiosis. For example, if it is found that only 50% of the population is at potential risk from illness due to *Cryptosporidium* water exposures, the overall PAR for all waterborne disease will be greatly reduced. The workshop examined the designs of epidemiological studies that should be considered to estimate the PAR, disease or health related endpoints that should be selected for study, water exposures that should be evaluated, and how sites should be selected. These issues are further discussed in a later section of this report.

---

[1] Waterborne Disease = (PAR) x (Disease from all Causes)

Workshop participants felt that a single study design for all geographic areas may not be the best approach to provide information for a national estimate of waterborne disease. Rather, more convincing evidence of a waterborne disease risk and its magnitude would be provided by results from several different study designs, especially if results were consistent among the different studies.

Epidemiological Study Designs

Observational and experimental studies can be conducted to assess associations between waterborne infectious disease risks and drinking water sources, treatment, and contaminants. This section describes the kinds of studies that were considered. Because of slight differences in the terminology used to describe epidemiological study designs, the scheme proposed in Table 1 is proposed to avoid confusion about the types of studies.

Table 1. Types of Epidemiological Studies*

I. Experimental
- A. Clinical
- B. Population-Based Intervention

II. Observational
- A. Descriptive (Includes Community-Based Intervention)
  - 1. Disease Surveillance and Prevalence Surveys
  - 2. Ecological
- B. Analytical (Includes Community-Based Intervention)
  - 1. Longitudinal
    - a. Cohort or Follow-up
    - a. Case-Control
  - 2. Cross-sectional

*Adapted from Monson, 1980

Estimates of the magnitude of the waterborne disease risk can be obtained from moderately large, well-designed, and well-conducted experimental and observational epidemiological studies (see Appendix C). Guidelines are available to assess the causality of an epidemiological association (see Appendix D), and each study must be carefully evaluated to determine whether its design is appropriate, the study size and power are adequate to detect a meaningful difference in risk of illness and the systematic bias.

Experimental Studies. Experimental epidemiological studies include population-based intervention studies and clinical trials. The experimental epidemiological study design is the only design approved by the Food and Drug Administration to determine the effectiveness of new pharmaceuticals. If this study design is found to be scientifically and economically feasible to evaluate water exposures, it can provide solid epidemiological evidence for an estimate of waterborne disease in the United States.

This type study considers the effect of varying some characteristic or exposure which is under the investigator's control, much like a toxicological study. Comparable individuals are randomly assigned to a treatment or intervention group (e.g., bottled, tap, and home filtered water groups) and observed for a specific health-related outcome (e.g., signs and symptoms of

illness, positive stool specimens to identify cases of giardiasis, antibodies in serum or saliva). The study is longitudinal in nature, much as a cohort study, as persons in each group are followed prospectively. Ethical concerns must be fully addressed, and periodic analyses may be necessary to assess risks throughout the study. Once sufficient information is available to suggest an increased or decreased risk in any group, participants must be informed. Once informed, participants may choose to drop out of the study.

Observational Studies. Two basic kinds of observational epidemiological studies can be conducted, descriptive and analytical. These approaches differ primarily in the supportive evidence they can provide about a possible causal association between drinking water and disease. An ecological study does not link individual health outcome events to individual exposure or confounding characteristics; an analytical study does. In an ecological study, information about exposure and disease is available only for groups of people but not for individuals that comprise the groups. Thus, information regarding possible confounders (alternative explanations) of associations between exposure and disease are not available. Investigators have examined limitations of these studies and when and under what assumptions this type study may be appropriate (Piantadosi, 1994; Greenland and Robins, 1994a, 1994b; Susser, 1994a, 1994b; Poole, 1994). Ecological studies can help develop hypotheses and may be useful for studying infectious disease where the overall community effect is the important consideration. For example, they can provide useful information about herd immunity in the community and possible changes in community disease rates associated with changes in water sources or treatment methods. These studies will not, however, provide reliable quantitative information about the magnitude of risk for a national estimate.

Analytical epidemiological studies can provide information about the causality of an association and estimate its magnitude. Information is obtained about disease and exposures for each person included in the study. Analytical studies are either longitudinal or cross-sectional. In a longitudinal study, the time sequence can be inferred between exposure and disease, i.e., exposure precedes disease. In a cross-sectional study, exposure and disease information relate to the same time period, and it is not always easy to determine if exposure precedes disease or vice-versa. Cross-sectional studies of infectious disease are less of a problem in this regard, but results must be evaluated to ensure that exposure caused disease.

Longitudinal studies are of two opposite approaches: the cohort study and the case-control study. The cohort study begins with the identification of individuals having an exposure of interest (e.g., persons using tap water, untreated groundwater, or unfiltered surface water) and an unexposed population (e.g., persons using bottled, disinfected groundwater, or filtered surface water) for comparison. Disease (e.g., AGI, cryptosporidiosis) or other health-related outcomes (e.g., serum or saliva antibody levels) are then determined for each group. A number of different diseases or symptoms can be studied. A cohort study can be either retrospective or prospective and sometimes a combination retrospective-prospective approach is used. Two or more groups of people are assembled for study strictly according to their exposure status at the present time (prospective) or in the past (retrospective). Disease incidence rates are compared between exposed and unexposed groups.

In a case-control study, the investigator identifies persons having a disease (e.g., AGI, cryptosporidiosis) or health outcome of interest (antibody level, stool positivity) and a control or

comparison group of individuals without the disease of interest. Past water exposures and other risk factors are evaluated in these persons. A variety of types of exposures (e.g., bottled water, tap water, filtered water, volume of water consumed) and potential risk factors (e.g., foreign travel, consuming untreated water while camping) can be studied. Controls (the comparison group) can be randomly selected from the either the general population or a special population within a specified geographic area (e.g., hospitals, HMOs, or nested in a selected cohort), but procedures must be adequate to ensure controls are free of disease. Different types of water exposures must also be available in the selected geographical area. There are several design variations and options for obtaining information (e.g., how cases and controls are selected; the geographical area studied; how information on exposures, risk factors, and confounding factors is obtained; who is interviewed; etc.). Both cohort and case-control studies can take advantage of a 'natural experiment'; that is, a community that changes water treatment or water sources or uses water sources of different quality for distribution to several distinct areas. In these 'natural experiments' exposure is assessed for each person included in the study rather than on an ecological level.

Systematic Bias in Epidemiological Studies

Bias can occur because of the way a study is designed and conducted, leading to a false or spurious association or a poor measure of risk (one that departs systematically from the true value). Procedures in the study's design and conduct are used to prevent or reduce possible systematic bias. When information for exposure and disease is collected by methods that are not comparable for each participant (e.g., selective recall of disease by exposed persons), an incorrect association will be due to observation bias. When the criteria used to enroll individuals in the study are not comparable, the observed association between exposure and disease will be due to selection bias. A wrong diagnosis of disease or assessment of exposure can result in misclassification bias. Random or non-differential misclassification bias almost always biases a study toward not observing an effect or observing a smaller change in risk than may actually be present, but non-random or differential misclassification can result in either higher or lower estimates of risk, depending on the distribution of misclassification.

Confounding bias may convey the appearance of an association; that is, a confounding characteristic rather than the putative cause or exposure may be responsible for all or much of the observed association. To cause confounding, a characteristic must be associated with the exposure being evaluated, a risk factor for the disease, and not part of the exposure-disease pathway. Information on known or suspected confounding characteristics is collected to evaluate and control this bias during data analysis. Matching is a technique used to prevent confounding bias. For example, if a child in day care is thought to be a possible confounding characteristic, an equal number or proportion of families with children in day care can be selected for study in order to avoid confounding bias from this exposure.

In an experimental epidemiological study, randomization is possible; that is, each individual in the study has an equal or random chance of being assigned to an exposed or unexposed group. Because of this random assignment of exposure, all characteristics, confounding or not, tend to be distributed equally between the selected study groups of different

exposure. This is the major reason experimental studies are held in such high esteem.

Random Error

Random error is governed by chance and influenced primarily by the size of the study (number of study participants). The likelihood that a positive association is due to random error is assessed by calculating the level of statistical significance ("p" value) or confidence interval (C.I.). A small "p" value or a C.I. that does not include unity (1.0) suggests chance may be an unlikely explanation for an observed association but does not completely rule it out. A statistically significant association may still be spurious, however, because of observation, selection, misclassification, and confounding biases discussed previously.

A study should be to conducted within a reasonable time period and without overwhelming logistical problems. The number of study participants should be reasonably large based on an expectation of detecting a meaningful difference in risk of illness between the exposed and unexposed groups.

When computing sample size requirements, either two-tailed or one-tailed statistical tests can be considered. One-tailed tests require fewer study participants for the same level of significance and can be justified because there is no anticipated beneficial effect of drinking contaminated water. But in an intervention study, the possibility of introducing harm by the intervention, although this is considered highly unlikely, should be taken into account in designing study size. Sample size requirements primarily introduce logistical problems - how to collect sufficient numbers of participants in a reasonable time period. If a particular study design is desired and it is not possible to obtain the required number of participants within a reasonable time, either the study period can be extended or the study area expanded in order to collect the necessary cases. Another less desirable option is to lower expectations about the magnitude of risk that can be detected. Study design changes can also be considered. For example, use of more sensitive measures of illness might allow detection of the same level of risk with fewer participants. Sample size requirements can be greatly reduced by use of more sensitive indicators, but supporting biological data may be needed to show the relationship between serum or saliva antibodies and illness or infection for selected diseases. These considerations are all part of assessing the feasibility of a particular study design.

Investigators must consider what magnitude of risk is desired to be detected and the number of participants required to detect that risk can be computed for each type of study. Determining the magnitude of risk that may be meaningful is often difficult, but the anticipated risk differences should be at levels sufficient for establishing public health policy. For drinking water associated risks, a relatively small difference in illness between the exposed (tap water) and unexposed groups should be considered. In the recently completed intervention study by Payment *et al.*, the study size was computed based on the detection of a 5% increased risk of AGI due to tap water consumption; a 14 to 40% PAR (depending on the age group) was observed in the study.

Designing a study to detect very small risks is especially important in the event no association is observed in one or more of the proposed studies. If systematic error is not suspected in a reasonably large study, the study's power to observe an association should be

considered. *A priori* expectations of the magnitude of risk and assumptions about its variability are used to determine an appropriate number of study participants. Upon completion of the study, information from that study is available to evaluate the study's power to detect an association of a certain magnitude. For example, studies in which an association is not observed often include a statement similar to the following, "no association was observed but the power of this study could only detect a RR of 1.9 or greater." This suggests the possibility that there may be an association, however, the risk of that association is less than could be detected in this study.

Measures of the Magnitude of Risk

The two basic measures of an association between exposure and disease in analytical studies are: the rate ratio or relative risk (RR) and exposure-odds ratio (OR). A RR or OR of 1.0 indicates no association; any other ratio signifies either a positive or negative association. For example, a RR or OR of 1.8 indicates an 80 percent increased risk among the exposed. The size of the RR and OR is also used to help assess an observed association. Based on Monson's (1980) experience, a RR or OR of 0.9 to 1.2 is considered too weak to be detected by epidemiological methods. It is difficult to interpret a RR or OR of 1.2 to 1.5 because one or more confounding characteristics can easily lead to a weak association between exposure and disease, and it is difficult to identify, measure, or control weak confounding bias. A larger RR or OR, however, is unlikely to be completely explained by some uncontrolled or unidentified confounding characteristic. Because the RR or OR for endemic waterborne disease is expected be less than 1.5, unidentified weak confounding bias will be an important consideration. The size of a RR has no relevance for assessing the occurrence of observation, selection, or misclassification bias.

Selection of a Study Design

Discussion of appropriate study designs to provide a national estimate of waterborne disease focused on five designs:

1) Population-based intervention study where persons in a community are randomly assigned to drink tap water or an alternative source of water that is known to be free of pathogens (bottled, boiled, or further filtered and disinfected at the home).
2) Cohort study where persons in a specified community are enrolled into several groups based upon their current water consumption patterns (always drink bottled water; frequently drink bottled water; always drink home filtered water which is effective in removing pathogens; always drink tap water without further treatment; drink only coffee, tea, soft drinks, etc.).
3) Case-control study where cases (AGI, severe AGI, persons with certain antibody levels, or a specifically-diagnosed disease) are selected from a broad geographic area with different types of community water systems or in an area or community where the only difference in water exposures will be individual choices about use of the community supply or alternatives (bottled, filtered, etc. as noted previously).

4) Cross-sectional study where the prevalence of illness or a health-related outcome is determined for populations residing in communities or areas with various water sources and types of treatment. Community-based intervention study (analytical rather than ecological design). In this study, the selection of communities depends upon planned improvements to water quality, increased water treatment, changes in water sources, or communities where separate water sources serve distinct areas. The study can be cross-sectional, cohort or case-control. Each of these study designs can provide evidence for a national estimate, and each offers certain advantages in terms of scientific and economic feasibility. However, each also has certain disadvantages and limitations. Participants generally agreed that the population-based intervention study would provide the strongest epidemiological evidence of waterborne disease risk and best estimate of the PAR for drinking water exposures, but it was also acknowledged that for both economic and scientific reasons several different designs could be employed. In the intervention study, it may be difficult to completely control systematic bias by not effectively blinding study participants about water exposures. Results of studies of a different design (and different sources of possible bias) can be used for comparison with results from intervention studies.
5) The population-based intervention study is the most expensive to conduct, and funds may not be available to conduct this type study in all five cities. Similar evidence of an association can be obtained from the cohort or case-control design at less cost, but the feasibility of conducting other types of studies depends largely on the number of persons in that area who do not use tap water without further treatment or use bottled water. In the intervention study, a equal number of persons will be assigned to the various water exposure groups, whereas, in cross-sectional, cohort, and case-control studies, 30% or fewer in the population studied may by choice consume bottled or water other than tap water. The cost of a cohort study will be much less than an intervention study because the cost of alternative drinking water sources is not provided by investigators, and because of this lower prevalence of exposure, a larger number of participants may be needed to detect the same magnitude of risk that can be detected by the intervention study. The study size of a cohort study may be reasonable to study certain health outcomes (AGI or persons with increased antibody levels), but this design may not be feasible to study specific waterborne diseases such as cryptosporidiosis, giardiasis, or shigellosis, as a very large number of study participants will be required. Sample size requirements, however, may be reasonable for a case-control study of specific waterborne diseases or severe cases AGI, and these studies can be conducted within a cohort study. A possible disadvantage of the non-intervention study designs is the self selection of persons who regularly consume bottled water. Are they different in ways that may cause confounding bias? This question needs to be answered before these studies are initiated, and an appropriate study of bottled water users should be conducted to obtain this information.

Workshop participants recommended that each of these types of study be further evaluated for possible systematic biases that may need to be controlled, methods available for their control, number of participants needed for a statistically stable estimate of risk, and other advantages and disadvantages. Each of the study designs should also be evaluated for their ability to consider each of several possible health outcomes (see subsequent section on selecting a health outcome). For example, if specific diseases such as giardiasis, cryptosporidiosis, or

Norwalk virus caused illness are to be studied, a case-control design (population-based, hospital-based, or nested within a cohort) may be more feasible than a cohort design and may be the most informative approach, whereas, if symptoms of AGI are to be studied one of the other study designs may be more appropriate. Discussions about the case-control design suggest that this design is probably useful primarily as a complement to the cohort study, but its use as a way to obtain information about AGI (mild or severe) should not be completely dismissed without further consideration. A possible problem with the case-control approach is the selection of appropriate controls who do not have the illness or health condition being studied.

An advantage of the cross-sectional study design is that a random sample of the U.S. population can be selected for study so that results will be representative of the general population. The cost of this study is also less than the cohort because persons are not followed over time, but a disadvantage is that the health and exposure of a person are determined at a single point in time. This problem can be remedied by conducting a series of cross-sectional studies in the area over a selected time period, but the costs would then approach a cohort study. Another advantage is that the study allows investigators to determine base-line levels of health, monitor water quality or other measures of effective treatment for adverse changes, and then conduct health monitoring studies as water quality changes. Serological measures can be incorporated in the cross-sectional design, but because paired sera are necessary to assess illness caused by Norwalk-like viruses, this study design can not be used to assess this illness; the cohort and case-control study could. It was suggested that a cross-sectional study could be conducted as part of NHANES by including questions about drinking water use in the NHANES questionnaire. This possibility should be explored, but the difficulty of collaborating with NHANES was also noted.

The following tables should be completed and used as a guide for selecting one or more study designs. Table 3 presents a systematic approach to identifying the advantages and disadvantages of each of the study designs, and Table 2 allows for an evaluation of possible systematic bias and methods for its control. The desired features of an epidemiological study of waterborne disease are presented as candidates for possible consideration.

After a study design(s) is selected, its scientific and economic feasibility must be assessed. Certain information may be required before further development of a particular study design. For example, information may be needed about the proportion of the population that regularly uses bottled water and their characteristics to help design the study and assess its feasibility. Procedures for reducing systematic bias, collecting information about the disease or health-outcome measure, exposure, and other risk factors, and selecting and enrolling participants should also be tested before a series of large epidemiological studies are initiated. Good management procedures and controls should also be tested and developed. Overall management of the studies through a multi-center approach should be investigated.

Table 2. Advantages and Disadvantages of Epidemiological Study Designs to Estimate the National Occurrence of Waterborne Disease

Important Features of a Study Design	Individual-based Intervention	Community-based Intervention	Cohort	Case-Control	Cross-Sectional
Control of Observation and Selection Bias	* Discuss advantages and disadvantages of study design in this regard				
Control Confounding Bias					
Minimal Exposure Misclassification					
Incorporate Dose-Response Assessment					
Minimal Disease Misclassification					
Sample Size Required to Detect a Reasonable Risk					
Time Required to Collect Sufficient Number of Study Participants					
Likely High Rate of Participation in Study					
Incorporate a Measure of Disease Severity					
Study AGI					
Study Specific Diseases of Interest					
Incorporate a Quantitative Measure of Exposure to Waterborne Pathogens or Overall Microbial Contamination					
Ability to Generalize Results					
Measure Worst Case and Average Exposures					
Include Subgroups of the Population					

Table 3. Possible Sources of Systematic Basis and Methods for Control in Epidemiological Studies to Estimate the National Occurrence of Waterborne Disease

Possible Sources of Bias and Comments	Individual-based Intervention	Community-based Intervention	Cohort	Case-Control	Cross-Sectional
Misclassification of Exposure	* Discuss sources and how likely control efforts might be successful				
Misclassification of Disease					
Selection bias					
Observation bias					
Confounding bias					

Selection of a Health Outcome for Study

In selecting the health outcome(s) that should be considered, it is important to remember that waterborne pathogens can be transmitted by both the ingestion and inhalation routes of exposure and that nausea and vomiting are important symptoms in addition to diarrhea. These are important considerations for the design of epidemiological studies. For example, if the study compares bottled and tap water exposures and a high proportion of persons are found to have symptoms associated with exposure to *Mycobacteria* or *Legionella* are noted in the study there will be no way to evaluate whether these symptoms are associated with the selected water exposures unless appropriate methods are included in the design of the study. Questions can be included to evaluate inhalation risks, and questionnaire development for assessing risk factors and identifying exposure will be an important part of all study designs.

There are really two important questions of concern in estimating the national occurrence of waterborne disease -- what is the magnitude of risk (PAR) and what are the important etiologic agents. The implications for water treatment are quite different if most waterborne illness is viral compared to bacterial or protozoan. Thus, it is important to consider the inclusion of procedures to identify the etiologic agents. Stool, serum, or saliva specimens can be collected and analyzed. One method is to bank sera from selected persons in the study so they can be tested at a later date as new procedures become available for identifying the agents. This should be considered at one or two sites. More information is needed to better assess the relationship between an antibody response to *Cryptosporidium* exposure, infection, and illness.

Work group participants felt the following health outcomes should be considered for inclusion in one or more epidemiological studies:

- Signs and symptoms of acute gastrointestinal illness (AGI), probably limited to highly credible gastroenteritis illness (HCGI) as defined by Payment *et al.*
- Respiratory symptoms.
- Antibody response (serum or saliva specimens). Collection of a single specimen to look for antibodies will not be appropriate for waterborne viruses except hepatitis. At the current time, Norwalk-like viruses require acute and convalescent sera.

Related issues include internal controls, herd immunity or the immune status of the population studied, secondary attack rates, severity of illness, and access to health care. Each must be evaluated for possible inclusion in the proposed study. Internal controls (e.g., if questions are asked about symptoms include symptoms that are unrelated to water exposures) may be needed to help assess possible observation bias. The immune status of the population will affect the size of population considered to be at risk and if immunity levels are high, a larger number of study participants may be needed to detect the desired levels of risk. The question of secondary attack rates relates to whether secondary transmission of illness should be considered waterborne. If so, assessment of secondary cases should be included in the study design. Participants generally felt that if the initial case within a family is waterborne and the illness is then transmitted person to person, the secondary cases should be considered waterborne.

Symptoms of waterborne disease are generally mild and many persons do not seek medical care. Persons can also become infected but show no symptoms. Thus, should the national estimate focus on waterborne infection, symptomatic illness, or severe illness? Evaluation of the severity of waterborne illness requires additional discussion because it may be useful to identify the economic impact of severe waterborne illness. Severe consequences of waterborne disease in the United States may be limited to special; high risk groups (young, elderly, immunocompromised), and these groups can be studied separately from the general population. There is concern that if the risk of only severe cases is considered (e.g., hospitalized cases of AGI) in an epidemiological study of the general population, little will be learned about the much larger number of mild cases of waterborne illness in this population. Mild illness in a small percentage of the general population can also result in large economic losses if persons must miss work or school due to the illness. Another concern about the study of only severe cases is that the waterborne transmission of hospitalized cases may be much different than the waterborne transmission of mild AGI. Severe cases may not be representative of all cases nor do they necessarily cause the greatest economic impact (e.g., in the Milwaukee outbreak the highest cost was associated with cases not hospitalized). The study population's access to health care should be evaluated in studies of hospitalized cases or cases from HMO's. There may be additional severe cases in the community that are not hospitalized because the persons do not have health care coverage. Participants felt that the primary source of study participants should be population based and at the household level; selecting individuals or families from a HMO may introduce bias due to health care access.

Populations to be Studied

Different populations may be at high risk for different pathogens, and the primary questions to be considered about selection of populations to study are:

- Should the proposed study be representative of the general population or representative of the types of water systems used by the U.S. population?
- Should studies focus on possible high-risk populations groups, such as young children, the elderly, or immunocompromised persons?
- Should studies be conducted to determine waterborne disease in populations exposed to water systems with highly contaminated water or average water quality?

To obtain results that are generalizable to the U.S. population, a random sample of the population is needed, and the cross-section study would be the study of choice. To obtain results that are representative of the water system types or water quality received by the general population requires a stratified random sample. Since these type studies consider water systems with good as well as poor water quality, large numbers of study participants are required to detect the likely very small risk. To optimize detection of an effect, populations with the poorest water quality should be selected (see next section). This is the worst case scenario and will likely define the upper bounds of the risk estimate.

The question of studying special populations that may be of high risk should be examined further, but participants agreed that the argument is in favor of first studying the general

population. If you find a risk in the general population, the risk in the immunocompromised population can be assumed as least as high. If you find no risk in the general population, you cannot conclude there is no risk in the immunocompromised population. If you study only the immunocompromised population, no conclusions can be made about the general population regardless of the findings. The study of children may be difficult, as approval from Review Boards may not be given (no invasive samples). However, children can likely be included in non-intervention studies where questions are asked about symptoms of illness.

Selection of a Study Site

The health outcome(s) selected for study, the study design, and study area all affect the feasibility of conducting a study. Criteria have been suggested for selecting a study design(s) and health outcomes, and similar criteria should be developed for selecting suitable study locations. Much of the success of the proposed study will depend on the area chosen, and its location should be selected with care. The most optimal study sites may not necessarily be in those areas where enhanced surveillance activities are being conducted for emerging pathogens, enteric disease, or food- or waterborne outbreak surveillance. The primary site selection criterion is water source/treatment/quality, and a major concern is to conduct the study in an area where water exposures are optimum for detecting an association (high levels of contamination or non-compliance and large numbers of people exposed). Since this represents the worst case situation and not the average or a representative exposure, the national occurrence of waterborne disease would be over-estimated. This must be balanced with the likely results from studies in areas with little contamination. Observing no associations in these study areas provides little or no information to estimate waterborne disease with any certainty. Results may be uninterpretable. To ensure that a difference in risk is detected requires that the most polluted water source be studied. Selecting unpolluted water sources reduces the likelihood of observing a risk. Since sites are pathogen specific and risk-specific it is important to have a diversity of water sources and quality. The alternative is to select a site where the water source is continuously impacted by sewage. It may be important to include a control community, that is a community with a high quality water source and good treatment. Additional discussion of this is needed, since it may be difficult to select this control community.

Participants generally agreed that a study site should:

- Provide a representative sample of the general population's exposure to water sources and treatment.
- Be contaminated to the extent that a health effect can be observed with an epidemiological study.

However, it may not be possible to meet both conditions.

Most people are served by large community water systems but small communities may be at greater risk because most water quality violations occur in small water systems.. It was suggested that at least one study area contain a variety of community water systems sizes. Possible sources of information about water sources, treatment, and quality include the EPA regulatory data bases, EPA special water quality surveys, EPA inventory, USGS data bases on source water quality, water quality information from state agencies, the AWWA and AWWA

Research Foundation, and the ICR. The methods for pathogen detection in the ICR are likely to be too poor to be of use in assessing exposures in epidemiological studies. However, the mini ICR will collect more frequent samples and use improved laboratory techniques and thus, may be useful. It may be possible to coordinate the selection of epidemiological study sites with selection of sites for the mini ICR. A number of water quality data bases are available and these should be further examined for useful information to select study sites. Individual water treatment facilities and state regulatory agencies or their regional district offices are the best sources of information about water quality. These sources should be used to obtain routine water quality monitoring data and other information about water exposures for use in the epidemiological studies.

Communities are also planning to change water sources and treatment technologies and these locations can be considered for possible study. Information was presented about water systems changing disinfection practices, installing filtration (e.g., granular, membrane filtration). The question was also raised about possible health effects associated with water distribution systems, and this should be considered in the design if possible.

Highly contaminated water systems may define high exposure but this does not take into account the susceptibility of the population at risk. It may be necessary to chose sites based on a matrix of considerations. Table 4 considers a matrix of water contamination and water treatment/regulatory compliance. A similar table should be developed for water quality and population characteristics.

Table 4 Matrix of Water Quality and Treatment/Compliance Considerations for Selecting a study Site

	No Water Treatment	Limited Water Treatment or Not in Compliance with Regulations	Good treatment and in Compliance.
No Contamination	*	*	*
Low Contamination	*	*	*
High Contamination	Few Systems	****	*****

*Study to be costly and unlikely to observe an increased risk

****Good candidate sites for epidemiological studies

Water Exposures and Home Treatment Devices To Consider for Intervention Studies

In intervention studies there is concern about systematic bias in reporting symptoms of illness because of knowledge about water exposures. Persons assigned to the tap water group may know their source of water and provide a biased response about the occurrence of illness. Likewise, persons assigned to the group using bottled or further treated tap water may know their source of water and also provide a biased response about the occurrence of illness. It is important to blind study participants about their source of drinking water. Internal controls, as discussed earlier, can also be included to help evaluate possible observation bias. Persons in the intervention study should also not change their normal behavior patterns related to the amount of water consumed and other behavior habits. These changes can be documented during the study for evaluation in the analysis.

Thus, any home water treatment unit considered for an intervention study should remove bacteria, protozoa, and viruses, but also should be able to be modified so that a sham water

treatment unit can also be installed. The sham unit must not modify the tap water quality in any way. Household members should also be unable to detect a difference in water quality from the sham or treatment units. The exterior of the units should be similar in appearance and not change the taste or appearance of the water. Routine service and maintenance must also be performed on both units and at the same frequency. A question was raised about whether water for the entire house or only the water at the kitchen sink should be treated. Since some people may drink at other than the kitchen sink, the study should specify that water be treated water for the entire house. Since whole house treatment is much more expensive, studies should be conducted to determine the compliance of study participants who agree to drink water only from the kitchen tap. It may be feasible to consider treating water only at the kitchen sink.

A discussion of the home water treatment units that should be considered for an intervention study identified the most likely candidates (Table 5). Use of these units to further treat tap water and use of bottled water sources should be evaluated. Ultraviolet light disinfection will destroy bacteria and viruses but is not very effective for protozoa. An exception is a newly developed pulsed unit, but this may not be readily available for purchase. An undetectable sham unit can be installed to operate at an ineffective light-wave length. Cartridge filters can remove particles, bacteria, and protozoa depending on their effective size but must be followed by ultraviolet light for destruction of viruses. Disinfection will be more effective when combined with cartridge filtration because of it ability to remove particles and turbidity. A sham cartridge filter unit can be installed. Flow must be restricted to simulate the pressure drop associated with the filtration process. A cartridge with a larger effective size should not be employed as a sham filter because it will filter some particles, and the water leaving the unit will not be the same quality as delivered to the tap. Reverse osmosis units are effective in removing chlorine so there may be problems of bacterial regrowth. The water storage reservoir may also become contaminated with bacteria which will represent a different type water exposure. If these units are installed, health department may have concerns because of HPC bacteria levels in treated water. Reverse osmosis treated water has a different taste so a sham unit without the reverse osmosis membrane may be recognized by the homeowner. Reverse osmosis also results in lots of wasted water.

An important point was made regarding the funding of the installation of the water treatment units. Manufacturers may assist in providing the units at or below cost. Workshop participants agreed that this could save money, but the study will be more credible if the industry that may profit from the results is not involved in its funding or conduct.

Table 5 Evaluation of Bottled Water and Home Water Treatment Units for Use in Intervention Studies

Home Treatment Unit or Alternative Water Source	Removal	Remarks about Sham Unit
Bottled Water	Should Test for Quality if Purchased	
Ultraviolet Light	Not Effective for Protozoa	Wavelength can be Changed to be Ineffective
Cartridge Filter	Particle, Bacterial, Protozoan Removal Depends on Effective Size; Cannot remove all viruses	Filter housing can be installed with a flow restriction device to simulate pressure drop through filter.
Cartridge Filter + UV	Good for All Pathogens if <2 Micron Pore Size	See Above
Reverse Osmosis	Excellent for All Pathogens; Lots of waste water; improves Taste	May be Difficult to Disguise Sham Unit as An operating Unit

References

Beaglehole R., Bonita R., and Kjellström T., "Basic Epidemiology" World Health Organization, 1993

Bennett *et al.* Infectious and parasitic diseases. In Closing the Gap: The Burden of Unnecessary Illness, Amler, R.W. and Dull, H.B., Eds., Oxford Univ. Press, 102, 1987.

Hauschild A.H.W. and Bryan F.L. "Estimate of cases of food and waterborne illness in Canada and the United States. J. Food. Prot. **43: 435-440**, 1980.

Hill A.B., "The environment and disease: association or causation?" Proceedings of the Royal Society of Medicine **58:295-300**, 1965.

Morris R. and Levin R. " Estimating the incidence of waterborne infectious disease related to drinking water in the United States." , *Proceedings of the International Symposium, International Assoc. Hydrological Sciences,* Rome, September 13-17, 1994.

Rothman K.J., "Causal Inference in Epidemiology" **Modern Epidemiology** pp 7-21, 1986.

Appendix A. Participant List

EPA/CDC Waterborne Disease Workshop Participants

Name/Address	Telephone #	FAX #	E-mail address
Ben Beard DPD, NCID, CDC Mail Stop F-12 4770 Buford Hwy N.E. Chamble, Ga 30341	770 488-4939	770 488-4454	cbb0@CDC.GOV
Sue Binder, MD DPD/NCID/CDC, MS F22 4770 Buford Highway, NE Atlanta, GA 30341	770 488-7750	770 488-7794	SCB1@CDC.GOV
Paul Blake, MD, MPH Epidemiology and Prevention Branch Division of Public Health Georgia Dept. Oh Human Resources 2 Peachtree St., NW, Room 325 Atlanta, GA 30303	404 657-2588	404 657-2608	PAB1@PH.DHR.STATE.G A.US
Joe Bresee, MD DVRD/NCID/CDC, MS G17 600 Clifton Road, NE Atlanta, GA 30333	404- 639-4651	404- 639-3866	JSB6@CDC.GOV
Dr. Rebecca Calderon U.S. EPA ORD/NHEERL (MD-58A) Research Triangle Park, NC 27711	919- 966-0617	919- 966-7584	CALDERON.REBECCA@ EPAMAIL.EPA.GOV
Dr. Cynthia Chappell University of Texas at Houston School of Public Health 1200 Herman Pressler P.O. Box 20186 Houston, TX 77225	713- 500-9372	713- 500-9364	CCHAPPELL@UTSPH. SPH.UTH.TMC.EDU
Paul Cieslak, MD Oregon Health Division 800 NE Oregon St., #21, Suite 772 Portland, OR 97232	503- 731-4024	503- 731-4798	CIESLAK@STATE.OR.US
Dr. Bob Clark U.S. EPA ORD/NRMRL (MD-689) 26 W. Martin Luther King Dr. Cincinnati, OH 45268	513 569-7201	513 569-7685	CLARK.ROBERT@EPAM AIL.EPA.GOV

Jack Colford, MD PhD Assistant Professor School of Public Health UC Berkeley 140 Warren Hall Berkeley, CA 94720	510 643-1076	510 643-5163	JCOLFORD@UCLINK2.BERKELEY.EDU
Mike Cox U.S. EPA OW/OGWDW (4603) 401 M St., S. W. Washington, DC 20460	202 260-1445	202 260-3762	COX.MICHAEL@EPAMAIL.EPA.GOV
Gunther Craun Gunther Craun and Associates 101 West Frederick St., Suite 104 Staunton, VA 24401	540 8861939	540 886-1939	
Vance Dietz MD DPD, CDC Mail Stop F22 4770 Buford Hwy N.E. Atlanta, GA 30341	770 488-7771	770 488-7761	vdx0@CIDDPD2.EM.CDC.GOV
Dr. Al Dufour U.S. EPA ORD/NERL (592) 26 W. Martin Luther King Dr. Cincinnati, OH 45268	513 569-7303	513 569-7464	
Kim Fox U.S. EPA ORD/NRMRL (B-24) 26 W. Martin Luther King Dr. Cincinnati, OH 45268	513 569-7820	513 569-7172	FOX.KIM@EPAMAIL.EPA.GOV
Dr. Floyd Frost Lovelace Medical Foundation Center for Health and Population Research 2425 Ridgecrest Dr., S.E. Albuquerque, NM 87108	505 262-3471	505 262-7598	FLOYD@AUDREY.TLI.ORG
Roger Glass, MD, PhD DVRD/NCID/CDC Mail Stop G-04 1600 Clifton Road N.E. Atlanta, GA 30333	404 639-3577	404 639-3645	RIG2@CIDDVD1.EM.CDC.GOV
Lee H. Harrison, MD Associate Professor Epidemiology and Medicine University of Pittsburgh 521 Parran Hall 130 De Soto Street Pittsburgh, PA 15261	412 624-3332	412 624-2256	LHARRISO@EDC1.GSPH.PITT.EDU

Dr. Fred Hauchman U.S. EPA ORD/NHEERL (MD-51A) Research Triangle Park, NC 27711	919 541-3893	919 541-0642	HAUCHMAN.FRED@ EPAMAIL.EPA.GOV
Craig Hedberg, PhD Surveillance and Disease Investigations Unit Minnesota Department of Health 717 Delaware Street, SE P.O. Box 9441 Minneapolis, MN 55440-9441	612- 623-5414	612 623-5743	CRAIGH@MDH-DPC. HEALTH.STATE.MN.US
Ron Hoffer U.S. EPA OW/OGWDW (4607) 401 M St., S. W. Washington, DC 20460	202 260-7096	202 260-3762	HOFFER.RON@EPA MAIL.EPA.GOV
Dr. Christon Hurst U.S. EPA ORD/NRMRL (MD-387) 26 W. Martin Luther King Dr. Cincinnati, OH 45268	513 569-7461		HURST.CHRISTON@ EPAMAIL.EPA.GOV
Dr. Walt Jakubowski U.S. EPA ORD/NERL (592) 26 W. Martin Luther King Dr. Cincinnati, OH 45268	513 569-7385	**513 569-7411**	**JAKUBOWSKI.WALTE R @EPAMAIL.EPA.GOV**
Dennis Juranek, DVM, MPH DPD/NCID/CDC, MS F22 4770 Buford Highway, NE Atlanta, GA 30341	770 488-7783	770 488-7761	DDJ1@CDC.GOV
Jan Keithly, Ph.D. Wadsworth Center, NYS DOH David Axelrod Inst. Pub. Health P.O. Box 22002 120 New Scotland Avenue Albany, NY 12201-2002	518 473-2692	518 473-6150	KEITHLY@ WADSWORTH.ORG
Laura A Klug, MPH OD/NCID/CDC Mail Stop C-12 1600 Clifton Road N.E. Atlanta, GA 30333	404 639-4733	404 639-4698	LBK1@CDC.GOV
Pat Lammie, PhD DPD/NCID/CDC, MS F13 4770 Buford Highway, NE Atlanta, GA 30341	770 488-4054	770 488-4108	PJL1@CIDDPD2. EM.CDC.GOV
Bill Mac Kenzie, MD DPD/NCID/CDC, MS F22 4770 Buford Highway, NE Atlanta, GA 30341	770 488-7784	770 488-7761	wrm0@CDC.GOV

Altaf Lal DPD, NCID, CDC Mail Stop F-12 4770 Buford Hwy N.E. Chamble, Ga 30341	770 488-4047	770 488-4454	AAL1@CDC.GOV
Dr. Christine Moe, PhD Dept. of Epidemiology School of Public Health Univ. of North Carolina at Chapel Hill CB-7400 Chapel Hill, NC 27599-7400	919 966-1420	919 966-2089	CHRISTINE_MOE@ UNC.EDU
Steve Monroe, PhD. DVRD/NCID/CDC, MS G04 1600 Clifton Road, NE Atlanta, GA 30333	404 639-2391	404 639-3645	STM2@CDC.GOV
Anne Moore, MD, PhDDPD/NCID/CDC, MS F22 4770 Buford Highway, NE Atlanta, GA 30341	770 488-7776	770 488-7761	AYM2@CDC.GOV
Randall Nelson, DVM Dept. of Public Health 410 Capital Avenue, Mail Stop 11 EPI Hartford, CT 06134	860 509-7994	860 509-8286	NELS127W@WONDER.E M. CDC.GOV
Dr. Patricia Murphy U.S. EPA NCEA (MD-190) 26 W. Martin Luther King Dr. Cincinnati, OH 45268	513 569-7226	513 569-7916	MURPHY.PATRICIA@ EPAMAIL.EPA.GOV
Pierre Payment, PhD Institut Armand-Frappier 531 boulevard des Prairies Laval, Quebec H7N 4Z3 Canada	514 687-5010	514 686-5626	PIERRE.PAYMENT@ IAF.UQUEBEC.CA
Bob Pinner, MD OD/NCID/CDC, MS C12 1600 Clifton Road, NE Atlanta, GA 30333	404 639-2859	404 639-4698	RWP1@CDC.GOV
Dr. Don Reasoner U.S. EPA ORD/NRMRL (MD-387) 26 W. Martin Luther King Dr. Cincinnati, OH 45268	513 569-7234	513 569-7328	REASONER.DONALD@ EPAMAIL.EPA.GOV
Sudha Reddy DBMD/FDDB Mail Stop A-38 1600 Clifton Road N.E. Atlanta, GA 30333	404 639-4399	404 639-4080	VBR8@CDC.GOV

Robin Ryder, MD Yale University School of Medicine 60 College Street P.O. Box 208034 New Haven, CT 06520-8034	203 785-2919	203 785-7552	RYDERRW@MASPO3. MAS.YALE.EDU
Dr. Steve Schaub U.S. EPA OW/OST (4304) 401 M St., S. W. Washington, DC 20460	202 260-7591	202 260-1036	SCHAUB.STEPHEN@ EPAMAIL.EPA.GOV
Drew Voetsch, MPH DBMD/FDDB Mail Stop A38 Atlanta, GA 30333	404 639-4637	404 639-4080	AAV6@CDC.GOV
Duc Vugia, MD, MPH Division of Comm. Dis. Control California Dept. Of Health Services 2151 Berkeley Way Berkeley, CA 94704	510 540-2566	510 540-2570	DVUGIA@HW1. CAHWNET.GOV
Debra Walsh U.S. EPA HSD (MD-58C) Research Triangle Park, NC 27711	919 966-0636	919 966-0655	WALSH.DEBRA@ EPAMAIL.EPA.GOV
Julie Ziegler SRA Technologies 8110 Gatehouse Rd. Suite 600 West Falls Church, VA 22042	703 205-8664	703 205-6260	JZIEGLER@SRATECH. COM

Appendix B. Abstracts of Presentations

DRINKING WATER TREATMENT AND DISTRIBUTION: ITS ROLE IN PREVENTING WATERBORNE OUTBREAKS

Robert M. Clark[a], Donald A. Reasoner[b], Kim R. Fox[c] and Christon J. Hurst[d]

For the last 100 years, drinking water utilities in the United States (U.S.) have played a major role in protecting public health through the reduction of waterborne disease. For example, in the 1880's for one year, the typhoid death rate was 158 deaths per 100,000 in Pittsburgh, Pennsylvania but by 1935 the typhoid death rate had declined to 5 per 100,000. These reductions in waterborne disease outbreaks were brought about by the use of sand filtration, disinfection and the application of drinking water standards. Despite this excellent record, occasional drinking water quality problems are a reminder of the need for constant vigilance. For example, in 1993, Milwaukee, Wisconsin suffered a cryptosporidiosis outbreak in which it was estimated that more than 400,000 people were ill. In July of 1993, Manhattan, New York was placed on a boil water order, as was Washington, D.C. in December of 1993.

Nearly all of the utilities in the U.S. that use surface water practice some type of treatment and many practice what might be called conventional treatment. A standard water treatment train that makes up conventional treatment usually consists of: coagulant chemical feed, rapid mix, flocculation, sedimentation, filtration, and disinfection.

In addition to being concerned over the role of water treatment in preventing waterborne disease, we should also be concerned over the role of distribution systems in causing or preventing waterborne disease. An illustration of the water quality problems associated with failures in distribution comes from two recent studies. During the period December 15, 1989 to January 20, 1990, residents and visitors to Cabool MO, population 2090, experienced 240 cases of diarrhea and 7 deaths. Escherichia coli serotype 0157:H7, a bacterium associated with the feces of healthy, dairy cattle, was isolated in many of the stool samples of ill people. Following an investigation by the Centers for Disease Control (CDC), with assistance by U.S. EPA's Water Supply and Water Resources Division, it was concluded that the illness was caused by waterborne contaminants that entered the distribution system through a series of pipe breaks and meter replacements that occurred during unusually cold weather. Another example of infrastructure failure occurred in Gideon, Missouri in November 1993, when 400-500 people of a population of 1000, contracted *Salmonella typhimurium*. The *Salmonella* outbreak contributed to the death of 6 elderly individuals. It is presumed that bird droppings contaminated water in storage tanks. As with Cabool, the city used a nondisinfected ground water. One consequence of these investigations is the finding that immunocompromised people (elderly and young) are often less capable of surviving waterborne illness than younger, healthier people. The issue of water borne disease from municipal water supplies is of growing concern in the U.S.

---

a. Director, Water Supply and Water Resources Division, NRMRL, U.S. Environmental Protection Agency, 26 W. Martin L. King Dr., Cincinnati, OH 45268.

b. Microbial Contaminants Control Branch, NRMRL, U.S. Environmental Protection Agency, 26 W. Martin L. King Dr., Cincinnati, OH 45268.

c. Treatment Technology Evaluation Branch, NRMRL, U.S. Environmental Protection Agency, 26 W. Martin L. King Dr., Cincinnati, OH 45268.

d. Microbial Contaminants Control Branch, NRMRL, U.S. Environmental Protection Agency, 26 W. Martin L. King Dr., Cincinnati, OH 45268.

Pathogen Occurrence
Walter Jakubowski, NERL, USEPA, Cincinnati, OH

EPA and CDC maintain a database of waterborne disease occurrence. This database was examined for the time period 1971-94 to determine what microbiological agents have been associated with waterborne disease outbreaks in the United States. For this time period, 737 waterborne disease outbreaks were reported. Eighty-nine of these were in recreational water and 69 were chemical etiologies. Removing these resulted in 579 drinking water outbreaks having a microbiological or unknown etiology producing acute gastrointestinal illness (AGI).

Drinking water outbreaks were reported from non-community (50%), community (40%) and individual (10%) supplies. About 60% of the outbreaks were associated with well or spring waters, 25% with rivers or lakes and 15% with a mixture of surface or ground waters, or the supply was unknown. Most of the outbreaks in non-community supplies were associated with well/spring waters. In the community water supplies, about 42% were from surface water sources, 39% from ground waters and 20% from mixtures or unknown sources.

AGI with no specific agent identified accounted for 56% of the outbreaks; protozoa were responsible for 22%, bacteria for 14%, and viruses for 8%. The bacterial agents identified were: *Shigella* (51% of outbreaks in the bacterial category), *Salmonella* (23%), *Campylobacter* (20%), *Yersinia* (3%), *Vibrio cholerae* (3%), and *Escherichia coli* (1%). The protozoan agents identified were: *Giardia* (90% of the protozoan outbreaks), *Cryptosporidium* (8%), *Entamoeba histolytica* (1%), and *Cyclospora* (1%). Viruses identified were: hepatitis A (57% of viral outbreaks), Norwalk-like (41%), and rotavirus (2%). Information on the densities of these organisms in feces, sewage/wastewater, surface water, ground water and drinking water will be summarized.

DETERMINATION OF THE INFECTIOUS DOSE OF NORWALK VIRUS IN HUMAN VOLUNTEERS

Christine Moe, Department of Epidemiology, University of North Carolina

The objectives of this study are to 1) determine the dose-response range of Norwalk Virus (NV) infectivity in adult human volunteers with various levels of preexisting antibodies; and 2) determine the risk of Norwalk virus at low levels of virus that are representative of levels in contaminated water and food. Norwalk and related viruses are predominant waterborne and foodborne viral pathogens, as evidenced by their role in numerous outbreaks and the high seroprevalence rates in the population. The first NV challenge study was in 1971. Since then, numerous human volunteer studies have been conducted to examine histopathology, patterns of illness and immune response associated with NV infection. However, none of these studies systematically addressed the issue of infectious dose because existing technology was inadequate to accurately titer the inoculum. The limited dose-response data from these studies is inconsistent and contradictory.

The study will test 45 human volunteers by having them ingest safety-tested, PCR-titered NV inoculum at different dose levels. Conducted in three rounds of 15 volunteers, each round tests three doses (five volunteers per dose) which approximate the ID_{10}, ID_{50} and ID_{90}. Samples collected include sera, stool, vomitus and saliva. Outcome measures are clinical illness (nausea, vomiting, diarrhea, abdominal paid), detection of NV in stool and/or vomitus by PCR and seroconversion.

Based on the results of this study a second phase is proposed which will be a closer examination of the response in the critical low dose region of the dose-response curve. The second phase is necessary to estimate the risk of NV infection from the low levels of virus that may occur in water and food. The lower two or three dose levels where volunteer response in observed in Phase 1 will be evaluated with a larger number of volunteers and more closely space dilutions, such as half-log or three-fold dilutions. The results of these studies will be invaluable for estimating the risk of NV gastroenteritis associated with exposure to contaminated water, food and other vehicles.

Timeline: July 1994 - June 1997

FIELD SURVEY OF CRYPTOSPORIDIOSIS: RISK FROM CONTAMINATED DRINKING WATER

Floyd Frost, Lovelace Medical Institute, Albuquerque, NM

Failure to detect cryptosporidiosis outbreaks among persons consuming oocysts contaminated drinking water may result from the insensitivity of disease surveillance programs, which are likely to only detect large outbreaks. Alternatively, it is possible that the presence of detected oocysts in the drinking water may be unrelated to the risk of disease outbreaks. A larger issue is whether elevated endemic levels of disease or infection, which do not result from disease outbreaks, are found among residents of cities consuming surface drinking water. The focus of this series of studies is to develop information to evaluate the endemic risk of infection among consumers of surface drinking water and to compare this risk with that for persons consuming municipals drinking water derived from deep wells.

The serological method used is based on a serological response to *Cryptosporidium* infection. Following infection, a large fraction of persons will develop antibodies to proteins contained on the surface of the *Cryptosporidium* merozoite. Serum IgG antibodies appear several weeks to several months following infection and remain, in many individuals, elevated for months to several years. Subsequent *Cryptosporidium* infections appear to increase the levels of these antibodies. Two different antibodies appear most commonly among previously infected people and are used in this study as markers for prior infection. These antibodies are characterized by molecular weight of the antigen or protein to which they attach The two antigens have wights of 17 kDa and 27 kDa.

Several studies to estimate the prevalence of *Cryptosporidium* antibodies have begun using blood bank volunteer donors as the source of sera. Blood bank donors have several advantages for the study. Access to these people is obtained through local blood banks. Donors are screened to exclude persons with behaviors that may increase certain risks of *Cryptosporidium* infection. Donors also tend to be in an age group that is less likely to have small children in diapers. By using phoresis donors, who regularly donate blood, it is possible to follow a large fraction of these individuals over an extended period of time. The sera from these donors also have a low risk of HIV and are, therefore, relatively safe to work within the laboratory.

In this ongoing study, sera are drawn from similar sub-populations in each of several communities. These communities represent different drinking water sources, such as cities which use unfiltered surface water, filtered surface water and deep well water. The sub-population in each city will be phoresis blood donors. Although these donors are not representative of the general community population, they should be similar between communities in all characteristics other than their drinking water source. Compared with the general population, they are likely to be easier to follow over time because of their strong community ties and they will likely have fewer other exposures that increase their risk of infection.

Each participant completes a two-page questionnaire on possible exposures (e.g., livestock, children, foreign travel, drinking water source and duration of residence in the community) and demographic information (age, sex). The questionnaire information is linked to the serology results and analyzed to determine whether residents consuming surface water have higher serological evidence of prior infection, after adjusting for other risk factors. In some communities, a second blood draw is performed and a second, much shorter, questionnaire is completed.

COMMUNITY ENTERIC DISEASE STUDY

Rebecca L. Calderon, Epidemiology & Biomarkers Branch, NHEERL, USEPA

Microbial organisms that cause enteric disease and their sources are a major concern for EPA. To conduct risk assessments or determine environmental health policy, information is needed on the level of disease, factors that influence that level, specific microbial organisms that cause illness, and possible sources of those organisms. Approximately 50% of food and waterborne disease outbreaks are of unknown etiology. Current surveillance programs do not provide adequate information on background rates of enteric illness and the relative source contribution of environmental sources of organisms that cause disease. In addition, current surveillance does not provide information on the effectiveness of environmental policy or management decisions in lowering exposure or reducing disease.

The goal of this study is to obtain information on enteric disease rates in the U.S. Enteric disease rates are needed to determine environmental health policy and management strategies for environmental sources of microorganisms. The objectives of this study are to: 1) determine the enteric disease rates in various communities across the country; 2) determine the relative source contribution of environmental factors associated with enteric disease; 3) determine etiologic agents associated with enteric disease; and 4) evaluate methods of surveillance. These studies will examine alternative surveillance methods versus longitudinal studies as means to obtain information for trend analysis.

Site selection. To vary environmental ranges of environmental factors, communities of different geographic location, size, drinking water sources and drinking water treatment have been identified. Ideal communities are those served by utilities that are about to change either source or treatment. Studies in these communities would provide additional information on the health benefit and cost-effectiveness of changes.

Timeline: June 1996 to December 1997 for completion of first community.

Site criteria: Community utility is either changing source water or changing treatment (e.g., adding filtration, changing disinfectants). The population served by this utility should be 15,000. Communities with one source water type are preferred.

Measures of enteric disease:

Longitudinal health study. Approximately 300 families with children between the ages of two and 10 will be asked to record gastrointestinal symptomatology on a daily basis. At the beginning of the study, those individuals will be asked to report baseline information on current use of medicines, occupation, underlying health conditions and childcare arrangements. Those individuals reporting symptomatology will be ask extensive environmental factors questions, i.e., eat out, recreational activities, contact with animals, travel, use of antibiotics, contact with diapered children and contact with ill individuals. These families will be monitored from June to December of 1996 and again June to December of 1997.

Hospital admissions. Admission rates for gastrointestinal disease (including emergency room visits) will be collected starting in June 1996 and ending in December 1997.

Nursing home surveillance. Select nursing homes will be asked to report incidences of gastrointestinal illness based on defined symptomatology. This surveillance will be conducted from June 1996 to December 1997. An addition a pilot study of a community retirement center will be conducted to correlate their reported symptomatology with nursing home and longitudinal measurements.

Antidiarrheal sales. Pharmaceutical distributors will be asked to report sales information from June 1996 to November 1997.

Clinical laboratory reporting. Clinical diagnostic laboratories will be asked to report number of stool specimens reported and results of those analyses.

Analysis of etiologic agents. Approximately 30 individuals from the health diary study will be asked to provide baseline sera and stools for etiologic analysis. During the course of the study, those individuals will be asked to provide stools when ill and acute and convalescent sera. At the end of the health diary, all individuals will be asked to provide sera and stools for analysis.

Cross sectional serologic survey. Sera from approximately 260 college students will be collected the first year to determine seroprevalence to known viral and protozoan agents. These students will be asked to donate twice more in six month intervals.

Serologic Assays for Cryptosporidium

Patrick Lammie, Ph.D.

The use of parasitologic techniques for surveillance of cryptosporidiosis underestimates the true prevalence of infection; oocytst shedding is intermittent and tests to detect *Cryptosporidium* oocysts in stool are insensitive. Serologic assays may provide more sensitive methods for detection of *Cryptosporidium* infection.

CDC investigators have developed an immunoblot to detect antibody responses directed against two specific *Cryptosporidium* antigens. Monoclonal antibodies directed against the 27- and 17- kDa antigens localize these proteins on the surface of sporozoites, the infective stage of the parasite. Antibody responses directed against these antigens are associated with protection from symptomatic infection. IgA and IgG responses to the 27- and 17-kDa antigens develop within 2 to 3 weeks after infection and persist for weeks and months, respectively. Bases on these preliminary studies, the immunoblot assay shows promise as a tool for epidemiologic studies designed to define the risk of waterborne cryptosporidiosis. For example, monitoring IgA responses to specific *Cryptosporidium* antigens may provide a measure of the incidence of infection. Unfortunately, the laborious nature of the immunoblot assay and the requirement for trained laboratory personnel to perform it may restrict the size of such studies and their ability to provide definitive estimates of risk. For these reasons, the current consensus is that it would be advantageous to convert the immunoblot to an ELISA format. Considerable progress has been make in the development of techniques to purify the native 27- and 17-kDa antigens and in the characterization of these antigens. In addition, one of these antigens has now been cloned and expressed in a bacterial expression system; thus, it is now feasible to determine whether simple assays based on this recombinant antigen can be developed. Additional work is now required to determine whether the sensitivity and specificity of the immunoblot is retained when purified or recombinant *Cryptosporidium* antigens are used in an ELISA format. It also will be necessary to determine the time course of *Cryptosporidium* antibody responses using these assays and the influence of repeated exposures to *Cryptosporidium* on antibody responses.

Molecular Approaches for Environment Monitoring of *Cryptosporidium* spp.
Altaf Lal, Ph.D.

DNA and RNA-based diagnostic procedures have the potential to detect *Cryptosporidium* spp in a species and isolate specific manner. The advantages of nucleic acid-based procedures include greater sensitivity, rapid analysis of many samples, relatively low cost, simultaneous detection of many pathogens, and the ability to differentiate between viable and non-viable parasites. The availability of species/isolate specific and sensitive diagnostic tests would address the question of : 1) what is the source of contamination in environmental samples? 2) to what extent are surface and finish water contaminated with parasites incapable of infecting humans? and 3) what is the zoonotic potential of most animal isolates?

Accordingly, CDC's *Cryptosporidium* molecular diagnostics program focuses on the development of diagnostic that would allow: 1) detection of *C. parvum* in environmental samples; 2) differentiation of *C. parvum* strain in both outbreak and non-outbreak settings; 3) differentiation of *C. parvum* strains from *C. muris*, *C. wrairi*, *C. meleagridis*, *C. baileyi*, *C. serpentis*, and *C. nasorum*; and 4) differentiation of viable from non-viable parasites.

We have developed PCR diagnosis procedures that use small subunit rRNA and antigen genes genus specific detection of *Cryptosporidium*. In addition to agarose gel-based diagnostic read out, PCR procedure has been adapted to bioluminescence and ELISA format. We are also investigating the use of ribosomal rRNA gene-based primers, developed by Australian investigators, to differentiate between human and non-human *C. parvum* strains. We have also invasion-PCR-based tests that allow identification of viable parasites.

DPD investigators are also focusing on sequence characterization of multiple genes of *C. parvum* (from human and non-human), *C. muris, C. baileyi, C. serpentis* toward the development of a robust, sensitive species and isolate specific diagnostic test. Preliminary analysis of the TRAP-C2 gene suggests that several point mutations separate humans *C. parvum* from non-humans parasites.

The existing resources at DPD allow us to support epidemiologic investigation by molecular diagnosis or pathogen. We are interested in working collaboratively with the water treatment facilities to test the presence of *Cryptosporidium* in source and filtered water. The test available would allow us to identify the presence of *Cryptosporidium* in water samples, and the tests in the process of development are designed to reveal the source of water contamination. The present staff and resources would allow us to handle between 10-15 samples a week. In some of these samples we will also be able to determine the viability to oocytsts. Analysis of larger sample size, as may be required in systematic monitoring of water samples, would require additional funding.

In about a four-six months time frame, we will test whether procedures developed in Australia would allow differentiation of *C. parvum* strains. In about a years time, using probes and primers developed through our multi-locus gene analysis, we may be able to distinguish *C. parvum* from non-*C. parvum* parasites. These procedures are being developed with a turn around time such that water utilities can use information developed by laboratory investigations as part of a decision making algorithm.

Prevalence of Risk Factor for Cryptosporidiosis:
Drinking Water Usage Patterns and Other Exposure Risks for NYC Patients
Dennis Juranek, Rosemary Soave, and Larry Davis

Objective: To assess the prevalence of known and theoretical exposure risks for Cryptosporidium parvum in HIV-infected individuals and a general population cohort at Cornell Medical Center.

Methods: A 17 page questionnaire was administered to 160 HIV infected patients and 153 ambulatory outpatients (representing the general population) at The New York Hospital-Cornell Medical Center. Persons representing the general population were recruited from general medicine, nephrology, and otorhinolaryngolgy medicine departments/services.

Incident Cases of Cryptosporidiosis in HIV-positive Persons: Risk Factors Associated with Infection and Development of Laboratory Methods for Detection

Anne Moore, John Ward, Frank Sorvillo, Susan Troxler, Pat Lammie, Ben Beard, Norman Pieniazek, Gilden Beall, Andrea Kovacs, Elisabeth Clark

Objectives: The study will characterize the epidemiology and risk factors for acute cryptosporidiosis in HIV-infected persons, characterize the serologic response to infection, and determine the prevalence of latent colonization with *Cryptosporidium* after acute illness and its relationship to CD4 and viral load.

Project Description: From a cohort of ~2600 active HIV clinic patients currently followed in New Orleans (60%) and Los Angeles (40%), this case-control study will enroll persons with acute laboratory-confirmed cryptosporidiosis along with two controls (one without diarrhea and one with acute, non-cryptosporidial diarrhea). Demographic and risk factor data will be obtained. Serum and stool specimens will be collected to facilitate the evaluation of PCR, molecular typing, and serologic assays for *Cryptosporidium* now in development.

Study implementation: A study questionnaire has been collaboratively developed with the participating sites and has been piloted in New Orleans and Los Angeles. Procedures for identification, contact, and interview of study subjects have been outlined. Methods of specimen acquisition, handling, storage, and shipping have been specified for each site. Study coordinators have been hired in each location. IRB clearance has been obtained at two participating hospital clinics (LAC/USC, Beer Medical Group). Clearance by CDC, and the remaining two sites (Harbor/UCLA, LSUMC) is pending and expected shortly. Study implementation with collection of specimens is targeted to commence February 15, 1997.

Progress in laboratory method development: PCR methods in development are specific for *Cryptosporidium*; other parasites tested (e.g. *Cyclospora*) are nonreactive. Molecular typing has been assessed at 3 genetic loci in specimens from 7 sources. All 3 loci show nucleotide variation among isolates from different sources, one position may show strain-specific variation for the Milwaukee outbreak, and intra-isolate variation is detected among multiple samples from the same outbreak. The data are promising, but are too preliminary to assign patterns.

Appendix C. Draft Study Proposals

Evaluating the Burden of Waterborne Diseases

EIP sites may apply for one or more of the following studies

A. Recipient activities for cross-sectional studies of the population to estimate the incidence of gastrointestinal illness and water consumption patterns

1. Supplementation of random digit dialing surveys already being conducted by the EIP site in the FOODNet activities to include information regarding gastrointestinal illness and water consumption (e.g., quantity of unboiled tap water consumed, quantity of bottled water consumed, public or private water supply, approximate fraction of water consumed at various sites - home vs. work, etc.)
2. Willingness to work collaboratively with EPA, CDC, and other EIP sites in the formulation of questions for cross-sectional surveys
3. Use of clinical signs as outcomes for gastrointestinal illness [e.g., a) clinically defined diarrhea, b) vomiting].

B. Recipient Activities for Feasibility Evaluation of Household Intervention Studies

1. Work collaboratively with CDC and EPA to identify intervention devices that can eliminate viable pathogens from water, can be altered to be ineffective in a blinded manner, do not change participant drinking habits, and that can be used in a limited feasibility study. Evaluate the cost of obtaining the intervention devices and installing and maintaining the devices.
2. Identify a cohort of randomly or systematically selected households of immunocompetent persons with at least one child in the household to participate in a feasibility study
3. Conduct a limited, blinded household intervention trial (feasibility study) to evaluate the effectiveness of blinding participants to the effectiveness of the intervention device, logistical considerations regarding maintenance of the intervention device, frequency of follow-up for data collection and collection of specimens (if desired). At the end of the study provide cost estimates for conducting specific portions of the intervention trial (e.g. data and specimen collection, specimen testing, etc.)
4. Measurement of outcomes during the feasibility study - Examples of such outcomes could include: a) clinically defined diarrhea, b) vomiting, c) laboratory studies of stool from cooperative, ill participants which would be tested broadly for bacterial, parasitic, and viral pathogens, and/or d) antibody response to specific pathogens such as Cryptosporidium and Calici viruses in study participants willing to give specimens (serum, saliva).
5. The recipient(s) will be required by EPA to document their Quality Assurance Project Plan (QAPP) for assuring that the environmental data collected are of the expected quality for their intended use and are properly assessed.
6. Work with CDC and EPA in identifying specific characteristics (e.g. water supply) would constitute a "good site" for doing future household intervention studies.

er Studies contributing to determining the burden of waterborne disease may also be ed. Examples of such studies include case-control studies using cases of severe outcomes gastroenteritis or community intervention studies to evaluate changes in the incidence of come (e.g. gastrointestinal illness, antibody response to a specific pathogen) before and change in the community's water treatment.

Appendix D. Causality of an Epidemiological Association

For each epidemiological study, the validity of the association observed between exposure and disease must be assured before its causality can be evaluated. Because of the observational nature of epidemiology, no single study can provide a definite answer about causality even if systematic bias is minimal. A body of evidence must be obtained from studies conducted in different geographic areas and populations. Guidelines are available to help epidemiologists assess the possible causality of associations observed in well-designed and conducted studies. The interpretation of epidemiological data should be made with caution and in the context of all available scientific information. Epidemiologists apply the following criteria to assess evidence about causality (Beagle *et al.*, 1993; Rothman, 1986; Hill, 1965):

- Biological Plausibility. When the association is supported by evidence from clinical research or toxicology about biological behavior or mechanisms, an inference of causality is strengthened.
- Temporal Association. Exposure must precede the disease, and in most epidemiological studies this can be inferred. When exposure and disease are measured simultaneously, it is possible that exposure has been modified by the presence of disease.
- Study Precision and Validity. Individual studies which provide evidence of an association are well designed with an adequate number of study participants (good precision) and well conducted with valid results (i.e., the association is not likely due to systematic bias).
- Strength of Association. The larger the RR or OR the less likely the association is to be spurious or due to confounding bias. However, a causal association cannot be ruled out simply because a weak association is observed.
- Consistency. Repeated observation of an association under different study conditions supports an inference of causality, but the absence of consistency does not rule out causality.
- Specificity. A putative cause or exposure leads to a specific effect. The presence of specificity argues for causality, but its absence does not rule it out.
- Dose-Response Relationship. A causal interpretation is more plausible when a risk gradient is found (e.g., higher risk is associated with larger exposures).
- Reversibility. An observed association leads to some preventive action, and removal of the possible cause leads to a reduction of disease or risk of disease.