A Critical Assessment Of Epidemiological Studies For The Investigation Of
The Health Risk Of Drinking Untreated Rainwater

S. Rodrigo¹, M. Sinclair², D. Cunliffe³, K. Leder²

¹Monash University and the CRC for Water Quality and Treatment,
srod5@student.monash.edu.au
²Monash University and the CRC for Water Quality and Treatment
³Department of Health, South Australia

Abstract
Rainwater has been used for many purposes globally including drinking, however whether untreated rainwater is a
risk factor for health problems remains inconclusive. Rainwater tanks generally serve households and in many
instances illness may not be reported. There are some published studies which have investigated the role of rainwater
in the occurrence of illness. Additionally, several outbreaks associated with rainwater have been reported in the
literature. The many epidemiological study designs which have been used in these investigations will be discussed.
This paper will critically assess the advantages and limitations of the various epidemiological designs and present a
design which may provide a more conclusive answer to the role of rainwater on the health of persons who use it for
drinking. A current randomized controlled trial assessing the risk associated with drinking untreated rainwater
amongst people who choose rainwater as their primary source of drinking water will be discussed.

Introduction
In some parts of the world today, rainwater is still the only source of water available for domestic
purposes. In other instances the poor aesthetic quality (colour, taste and hardness) of mains water
supplies results in people using of rainwater for household purposes (1). In Australia
dissatisfaction with mains water quality and use of rainwater tanks is highest in South Australia
(1).

Rainwater is generally regarded by consumers as a clean and pure source of water compared to
other water sources. However, consumers fail to consider that rainwater may be contaminated
with the faecal deposits of birds and other small animals on the catchment surface. In addition,
leaching of heavy metals from the roof and gutters into the rainwater run-off, and the deposition
of industrial pollutants on the roof surfaces in urban areas are other safety concerns for rainwater
harvesting. While many studies have shown that harvested rainwater is contaminated (2), this is
not reflected in the number of reported cases of illness due to rainwater consumption. This may
be due in part to the fact that rainwater tanks generally serve households and as a result, illness is
often underreported. Even if entire family becomes ill, the outbreak is unlikely to be noticed by
the health authorities and the source of illness will often not be traced to the rainwater tank.
Hence the majority of the outbreaks implicating rainwater which have been reported in the
literature have involved communal tanks serving larger numbers of people (3-5).

There are a few epidemiological studies published which have attempted to establish a
relationship between rainwater consumption and health risk. The epidemiological designs which
have been used are case reports, case control and cohort studies. Thus far no study has used the
randomized control design (the gold standard) to investigate the health risk involved with
rainwater consumption. This paper aims to assess the published study designs of the health
effects associated with rainwater consumption and to present an alternative design, namely a
randomized control study, for the investigation of the health risks of drinking untreated rainwater.
Epidemiological Studies

Epidemiological studies can be classified into two broad groups, observational or experimental. Observational studies can be descriptive or analytical while experimental studies are analytical. Studies can also be grouped depending on the way data is collected: retrospective (past exposure information) or prospective (forwards in time). Observational studies may be retrospective or prospective while experimental studies are always prospective.

Descriptive studies show the distribution of diseases and risk factors in populations, characterizing the distribution of disease in terms of person, place and time. The advantages of these studies are that they are relatively cheap and quick to complete and they give a useful initial overview of a problem that may point to the next appropriate step in its investigation. These studies are therefore used to generate causal hypotheses which can be tested in more rigorous study designs. Observational analytical studies analyse the relationship between disease and exposure but do not control the exposure. Experimental studies measure the effectiveness and safety of particular types of health care interventions and can be applied to the study of environmental exposures. Ethical considerations are particularly important when considering the design and execution of any kind of intervention study.

In all epidemiological studies errors can be introduced and these may be random (due to chance) or systematic (leading to bias). Chance errors cannot be completely eliminated however they can be minimized by increasing the sample size and/or the number of repeated measurements. Bias occurs when there is a definite trend in the data which shows a consistent deviation from the truth. The types of bias encountered include selection, observer (measurement) and information. Selection bias occurs when study participants have different characteristics from those not selected and are therefore not representative of the population. Differences in the measurement of the risk or health status by the study personnel results in observer or measurement bias. This can be reduced by standardization in the classification of health outcomes and the questions asked, and blinding of observers to the health status of the participant and/or treatment groups. Information (recall) bias arises as a result of the difference in reporting of exposure between those with and those without the outcome of interest. Cases may be motivated to recall past exposure hence there will be an increase in the estimation of the risk. Retrospective studies are prone to recall bias due to the time lag between exposure and manifestation of the outcome.

The epidemiological studies which have been used for studying associations between rainwater consumption and illness are case reports, cross-sectional, case control and cohort. However, to definitively link rainwater consumption and illness and show causation, prospective cohort or randomized control studies are optimal. Retrospective cohort and case control studies may be able to suggest an association between exposure to rainwater and adverse health outcomes but cannot prove causation as they cannot confirm the exposure preceded the effect. These study types will now be discussed in relation to published studies involving rainwater.

Case report

Case reports use retrospective data and present anecdotal evidence to provide information on possible risk factors for disease and to serve as a foundation for stimulating further research. Brodribb (6) reported on a case involving an elderly woman admitted to hospital with gastrointestinal symptoms. The woman was treated and discharged, however illness recurred. On investigation, it was noted that the only source of drinking water was a rainwater tank from
which *Campylobacter fetus*, which caused the bacteraemia, was isolated. Advice was given to boil the water and since no further illness was reported the rainwater tank was linked with her illness. Case reports are limited since there is generally not enough evidence to reach a conclusion regarding the cause of the illness and are usually biased toward abnormal outcomes. Another limitation is that the number of people exposed, for example those using rainwater tanks, is unknown therefore the event rate cannot be calculated.

**Cross-sectional study**

Cross-sectional studies are fact-finding surveys used to generate hypotheses and are useful when assessing the health care needs of the population. Often the potential risk factors identified in cross-sectional studies are investigated further using a more rigorous design.

The cross-sectional study by Heyworth et al. (7) investigated the role of drinking water sources (including rainwater) on the prevalence of gastroenteritis in young children. Parents were required to complete a questionnaire for two weeks prior to the health check date. The preliminary analysis presented in the paper was not able to show a statistically significant relationship between drinking water source and gastroenteritis.

In cross-sectional studies, recall bias can affect the quality of data collected since participants are required to give past information. Also since exposure and outcome are determined at the same time, it difficult to distinguish cause from effect. In the case of rainwater, whether illness led to rainwater consumption or rainwater caused the illness cannot be clearly distinguished in such a study; that is water consumption pattern may be changed as a result of illness.

Another problem encountered is confounding, which occurs when a risk factor is associated with both the factor under investigation and the outcome. The confounding factor changes the effect of the exposure on the outcome which can result in an apparent association between exposure and outcome when in reality there is none. Techniques such as matching, stratification and restriction can be used to reduce confounding at the design stage. Adjusting for multiple variables by using statistical models can control for confounding at the analysis phase. However, in cross-sectional studies, it is difficult to control for confounding. There was no information in the South Australian study as to the effect of potential confounders such as socioeconomic status.

**Case control study**

The case control design investigates the exposure of persons with and without diseases to several risk factors. In a case control study, cases of a particular disease (e.g. gastroenteritis caused by *Salmonella*) occurring in a specific geographic area in a particular time interval are recruited and their exposures during the likely time of infection are investigated. Controls are recruited amongst people from the same area who did not experience the disease. Controls are sometimes matched to the cases with respect to age, gender and other factors which may confound the relationship between the disease and exposure.

The main advantage of the case-control study is that it enables the study of rare health outcomes without having to follow thousands of people. This type of study is cost and time effective, can investigate a variety of exposures, and requires a relatively small sample size. However, this design has the disadvantage of selecting cases and controls after both the outcome and the exposure have occurred. Additionally case control studies cannot directly estimate the risk of
Case control studies can provide evidence of an association between illness and exposure to water sources. The results of these studies highlight that the importance of widespread waterborne risks depends on the system type and vulnerability to contamination, and for some illnesses, the immune status or age of the population (8). However selection and recall bias can occur with case control studies. For example, selection bias could occur if people drinking particular water sources are more likely to be included in the study (9, 10).

Several case control studies have implicated rainwater as the risk factor for illness (5, 11, 12). In these studies, cases were selected based on being infected with a specific pathogen and the focus of these papers was on the illness experienced. There are no published case control studies which have primarily aimed to investigate the risk of rainwater consumption (8) since these studies start with the disease as the focus of investigation and not the exposure.

In the case control study by Weinstein et al (11) in South Australia, it was suggested that rainwater had a protective effect in the transmission of cryptosporidiosis compared with other water types. Cases were selected from laboratory confirmed cases while controls of similar age and same sex were nominated by cases which may have resulted in selection bias. Merritt et al (4) used a case control design to investigate an outbreak of campylobacteriosis in Queensland. Illness was found to be significantly associated with a water dispenser and rainwater tank (OR = 6.65, 95% CI = 1.45 – 33.1), with the water dispenser being filled from the rainwater tank. Eberhart-Phillips et al. (12) also used a case control design to investigate campylobacteriosis and found a greater risk to be associated (OR = 2.50 95% CI = 1.04 – 4.62) with rainwater as the source of drinking water in New Zealand. The cases were mainly persons of European descent, higher income, educated and living in urban areas. However, in New Zealand, persons using roof-collected rainwater live mostly in areas not served by public mains water supply which are often rural areas. Persons living in rural areas may be less prone to visit the doctor in cases of GI illness compared with those in urban areas, thereby contributing to selection bias. It is also known that the number of notified cases is a small fraction of the actual cases since many cases are not reported, hence, there may be considerable bias in the selection of cases for the study.

More recently, Ashbolt et al. (13) reported the association of untreated water collected in rainwater tanks with cases of salmonellosis due to Salmonella Mississippi in Tasmania (adjusted OR = 5.08, 95% CI = 2.00 – 12.90). In Queensland, an outbreak of salmonellosis at a construction site led the investigators to conclude that there was strong evidence suggesting that tanks which collected rainwater led to the outbreak of illness (5). The tanks were proven to be contaminated with Salmonella Saintpaul.

Cohort study
A cohort study which may be retrospective or prospective is a longitudinal study where information on exposures and development of illness is recorded for a group of people. While the forward directionality of prospective cohort studies make them less prone to recall bias and less likely for selection bias to occur since exposure is evaluated before the health status is known, it is expensive and time-consuming. Since exposure precedes outcome, prospective cohort studies establish a temporal relationship and thus the causal link between exposure and disease. The disadvantages result from loss of subjects during the lengthy follow-up period and the possibility
of differential loss which leads to bias. Another bias in these studies may be due to the exposed being followed more closely than the unexposed which can negatively affect the study validity.

Cohort studies do not normally involve blinding of researchers and participants nor is randomization part of the design. There may be bias in reporting by the participants and/or researcher bias in eliciting responses from participants. In addition, there may be differences between the groups – in socioeconomic and demographic factors which affect the health outcome independently of the exposure of interest.

There is one published cohort study which investigated the risk of gastroenteritis in children according to their drinking water sources with its main focus on rainwater (14). The study concluded that there was a protective effect from rainwater consumption compared to tap water consumption. The prospective nature of the study enabled a determination of the temporal sequence between exposure and the outcome under investigation, with recall bias being kept at a minimum. Since participants were selected before illness, bias due to selection was also minimized. However, since participants were not blinded there may be some bias associated with responses. In addition, the analysis of the data allowed for the control of potential confounders. However, the follow up period for the study was 6 weeks which was a relatively short period of time. A longer period may have given more extensive data however the cost and loss to follow up would be increased with longer follow up times.

**Outbreak reports**

Outbreak reports, while not a unique epidemiological study design, provide information on the association between risk factors and illness. There are investigations of disease outbreaks which have resulted in the implication of rainwater as a risk factor for gastrointestinal illness. In most instances these reported outbreaks have occurred where rainwater tanks supplied water for large groups of people at resorts, camps or work sites. In outbreaks there is a time delay between exposure and illness which can make it difficult to prove whether or not rainwater was the source. Another issue in the case of outbreaks is that the surveillance systems do not recognise or investigate all outbreaks especially if small (9, 10). Rainwater tanks generally serve individual households and in such cases, the surveillance systems may not detect the outbreak.

The outbreaks reported in the literature which have implicated rainwater as a risk factor for illness include an outbreak of gastrointestinal illness attributed to *Salmonella arechevalata* at a camp in rural Trinidad (3). *Salmonella* was found in water samples from the kitchen taps connected to the rainwater tank; however a delay in collection of samples from the rainwater tank resulted in tank water samples being negative for the organism. An outbreak of gastroenteritis also occurred in a family in New Zealand (15). *Salmonella* Typhimurium phage type I was isolated from family members and the tap water supply which was fed from a rainwater tank. Other outbreaks include dengue type 2 in Queensland which was linked to the breeding of *Aedes aegypti* in vessels collecting rainwater such as tanks and discarded containers (16); and a mixed outbreak of cryptosporidiosis and giardiasis at a school camp in Victoria was found to be due to contamination of an in-ground rainwater tank supply by sewage effluent (17).

**Randomized control trial (RCT)**

A randomized controlled trial (RCT) is a prospective experimental study or intervention in which participants are randomly assigned into two groups to receive or not receive the treatment under
study. The results are then analysed by comparing health outcomes in the two groups. The advantages of this design lies in the randomization which should result in an even distribution of unmeasured variables among the groups. This study type also controls for bias and confounding.

The most ideal randomized controlled trial design would involve randomizing persons to receive either rainwater or mains water. In this design persons who are currently using mains water could potentially be asked to drink untreated rainwater and vice versa. This study would present ethical issues since the current knowledge on the health risks of drinking rainwater is incomplete.

Another approach to investigating the health risks of drinking rainwater while avoiding potential ethical issues is to include people only if they already currently drink rainwater as their primary drinking source. All enrolled households would receive a water treatment unit, but the households would be randomly allocated to receive either an active filter treatment unit which removes microorganisms or to receive a sham unit. Health outcomes in both groups would be monitored. Blinding of the study participants and researchers as to the nature of the water treatment device allocated to participants during the study would reduce potential recall and observer biases. However, the cost involved with RCT studies may be higher than other study designs and may involve prolonged study periods.

An RCT is currently being conducted in South Australia which has the highest prevalence of rainwater tanks compared to the national average. Three hundred households, each having at least 4 persons, will be enrolled and will be randomized to receive a real or sham water treatment unit. Real units will remove microorganisms while the sham units will not. The sample size has been calculated to detect a 25% reduction in the overall rate of highly credible gastroenteritis episodes in the filter group with 80% power using a two-sided 5% significance level. This reduction in the rate of gastrointestinal diseases is considered clinically relevant and realistic when considering the population size and available funding.

Participants will be recruited from the Adelaide metropolitan area using a combination of methods including use of the electoral roll extract, and advertising and distribution of material in schools and health check centres. Recruitment material will be provided advising the interested and potentially eligible families to telephone the Study Centre to find out more about the study. At that time, a screening questionnaire will be administered to determine the eligibility of the household. Eligible households will then be sent an Information Booklet containing details on the conduct and requirements of the study. Households will be contacted after receipt of the booklet to determine interest in participation, and a second screening questionnaire will be administered. If the household remains eligible and willing to take part in the study, an enrolment visit will be scheduled during which demographic information for all eligible household members will be obtained. Additionally a questionnaire on tank characteristics (e.g. size, materials) and maintenance (use of first flush diverters etc.) will be administered at this visit.

Recruited households will be required to complete a health diary for every household member over a one year period. The health diary will focus mainly on symptoms of gastroenteritis, but will also include questions regarding other symptoms. In addition to the diaries, questionnaires will be administered when cases of highly credible gastroenteritis are recorded. The rate of gastroenteritis in the real and sham groups will then be compared to determine whether removal of microbes from rainwater among the group with the real filter results in a decrease in illness
rates compared to those with a sham filter. This will provide evidence regarding whether or not there are pathogens in untreated rainwater that are causing gastroenteritis among consumers. A nested case control study will be performed, with both cases and controls being interviewed to determine the activities in the week prior to the gastrointestinal event. An unmatched design will be used for the case control study since finding the required matched controls may not always be possible. In addition, matching is not necessary since adjustments for the relevant factors can be performed during the analysis. Unmatched cases and controls will therefore be analysed with adjustment for sex, an adult/child indicator and age.

The prolonged time over which randomised controlled trials are conducted can lead to attrition bias, so dropout must be minimized. The current study will establish ongoing communication with study participants via monthly newsletters. Several incentives during the data collection period will be important to achieve a low dropout rate (≤10%) – shopping vouchers for prompt return of diaries, contests and prizes for children. Similar strategies were used in the Melbourne Water Quality Study which had a 6.8% dropout rate (18).

**Conclusions**

Variations in water quality may affect the differences reported in the risk of illness from rainwater consumption. Rainwater quality is a reflection of the catchment surface, the surrounding environment, the tank characteristics, and its proximity to animals. Population age and the susceptibility to microorganisms may also affect the reporting of illness.

While a few published studies have suggested rainwater is a risk factor for gastrointestinal illness, these – with the exception of the cohort studies conducted in South Australia – have not primarily had rainwater as their focus. As such the association between rainwater needs to be explored by a more rigorous study design. The proposed randomized controlled trial aims to address the shortcomings of previous designs.

The underlying hypothesis of the Adelaide rainwater RCT is that waterborne pathogens in rainwater supplies may contribute to gastroenteritis in households drinking untreated rainwater. If the rate of gastroenteritis is greater in the unfiltered group, then this will imply there is a beneficial effect to filtration of rainwater. If, however, there is no difference in illness rates between the groups, then this will suggest that consumption of untreated rainwater is safe. This study will not provide information regarding the comparative safety of drinking rainwater versus mains water, which could only be addressed by comparing health outcomes among consumers of rainwater versus tap water. However, our current lack of knowledge regarding the safety of rainwater consumption means that ethical concerns would preclude performing a blinded RCT to address this. The Adelaide rainwater RCT will however facilitate endorsement of rainwater usage for a number of household purposes, including activities that may be associated with inadvertent consumption such as bathing and showering, should drinking unfiltered rainwater be found to be safe. Therefore the findings will be invaluable in helping to advise government and shape policy response outlining acceptable uses of rainwater in an urban setting serviced by an existing reticulated supply.

**References**