

**Rolling Revision of the  
WHO Guidelines for Drinking-Water Quality**

**Draft for review and comments  
(Not for citation)**

**How to interpret epidemiological associations**

**By G.F. Craun and R.L. Calderon**



**World Health Organization  
August 2004**

# How to interpret epidemiological associations

---

*G.F. Craun and R.L. Calderon*

## 1 INTRODUCTION

Cardiovascular disease is among the leading causes of morbidity and mortality in industrialized countries. Hypertensive disease, smoking, excessive alcohol consumption, diet, and physical inactivity have been identified as risk factors, but they do not entirely explain the worldwide variability of cardiovascular disease (Sauvant and Pepin 2002; Hornstra *et al.* 1998). It has long been suspected that cardiovascular disease mortality may be associated with geochemical constituents, especially water hardness.

### 1.2 Development of the water hardness hypothesis

Our knowledge of the association of cardiovascular disease with water hardness is based largely on observational epidemiological studies that have been conducted over the past half century. The hypothesis is that hard water or minerals, such as calcium and magnesium found in hard water, may help decrease cardiovascular risks or constituents associated with soft waters (e.g., corrosion products from distribution pipes, sodium from home softening units) may increase the risk.

Concern about a water factor can be traced to research reported in the 1950s. Enterline and Stewart (1956) called attention to the marked geographical variation in death rates from heart disease for the United States during the period 1949–1951 and suggested that place of residence might be an important risk factor. Kobayashi (1957) reported an association of mortality from apoplexy in Japan and the acidity of river water (i.e., the ratio of sulfates to carbonates). After meeting with Kobayashi, Schroeder sought to confirm the findings in the United States using the 1949–1951 mortality statistics (Comstock 1979a, b). Since drinking water hardness was the most widely available water quality measure in the United States, Schroeder (1960a, 1960b, 1969) evaluated its relationship to death rates on a state-by-state basis (Comstock 1979). Schroeder (1960a, 1960b) found statistically significant inverse (protective) correlations between water hardness and deaths from all causes ( $r=-0.36$ ), all cardiovascular diseases ( $r=-0.56$ ), coronary heart disease ( $r=-0.31$ ), other cardiovascular diseases ( $r=-0.36$ ), and stroke ( $r=-0.33$ ). In a second study of 163 metropolitan areas of the United States, Schroeder (1969) also found an inverse association between coronary heart disease mortality rates for white males 45–64 years of age and water hardness ( $r=-0.29$ ), magnesium ( $r=-0.30$ ), and calcium ( $r=-0.27$ ). In the years since these landmark studies, epidemiologists around the world have studied the association between water hardness and cardiovascular disease. Results of these epidemiological studies are reviewed in subsequent chapters (see Calderon and Craun, Monarca

and Donato, Kozisek, and Frost). This chapter discusses epidemiological study designs and other important aspects of the studies to help readers better interpret the observed associations.

## 2 TYPES OF EPIDEMIOLOGICAL STUDIES

Observational epidemiological studies can be descriptive or analytical (Table 1). Descriptive epidemiology is primarily used to summarize disease information, assess geographical or temporal patterns of disease, and develop hypotheses about disease etiologies. Ecological studies, sometimes called geographical, correlation, group, or aggregate studies, are used to explore possible relationships between available health statistics (e.g., cardiovascular disease mortality) and geographical location, population characteristics, or environmental and water quality measures (e.g., water hardness). Investigators often conduct these studies to help develop hypotheses for further evaluation with analytical studies. Because ecological studies consider exposures and outcomes determined in the aggregate for groups of people, they rely on available statistics. Thus, they are relatively inexpensive and easy to conduct. However, the observed associations should be viewed with caution (Greenland and Robins 1994a, 1994b; Piantadosi 1994; Poole 1994). Since the health, exposure, and demographic measures characterize population groups, inferences from associations observed in an ecological study may not necessarily pertain to the individuals within the group, especially when outcomes from long-term exposures are studied. Population migration and other demographic changes may have occurred, and group exposures (e.g., water quality) and health statistics may not be consistent over the relevant study period. For example, migration into the study area by older people may increase the mortality rates, but the increase may reflect exposures at previous residences. Neither theoretical nor empirical analyses have offered consistent guidelines for the interpretation of ecological analyses. However, ecological analyses have a distinct advantage because of their statistical power to detect small risks. If information is available to adequately assess population characteristics, relevant exposures, and the health outcome of interest, these studies can help identify potential problems. This design is especially useful for environmental health studies.

Table 2.1. Types of observational epidemiological studies\*

Descriptive Studies	Analytical Studies
Disease Surveillance/Surveys	Cross-sectional
Ecological	Longitudinal <ul style="list-style-type: none"> <li>• Cohort or Follow-up</li> <li>• Case-Control</li> </ul>

\*Adapted from Monson (1990)

Many ecological studies have assessed the association between water hardness and cardiovascular disease mortality. In these studies, investigators considered the available population information about mortality, water exposures, and demographic characteristics. Although many studies considered water hardness, some considered calcium and magnesium, the principal constituents that contribute to the water hardness. Several community-intervention studies evaluated mortality changes associated with water hardness changes.

Analytical studies are used to test specific hypotheses, and they can provide a quantitative estimate of the relative risk and other information to help scientists assess causality (Monson 1990). Information about disease, exposures, and important behaviours or characteristics is obtained from each study participant. Analytical studies can be either longitudinal or cross-sectional (Table 2.1). In the cross-sectional study, information about exposure and disease relate to the same time period. These studies are most useful for studying diseases with a short latency period. The longitudinal study allows a time sequence to be inferred between an exposure and disease. Two distinct, opposite approaches are used. The cohort or follow-up study begins with an exposure or characteristic of interest and evaluates disease consequences of the exposure or characteristic. The case-control study begins with a disease or health condition of interest and evaluates previous exposures of interest and risk factors associated with the disease. Because past exposures are studied, a case-control study may be labelled a retrospective study. It is sometimes called a case-referent or case-comparison study. The case-control study is usually less costly than the cohort study, since fewer study participants are required for adequate statistical power.

To avoid selection bias in a case-control study, participants should enter the study solely on the basis of their

disease status, and the investigator should have no knowledge of their exposure status. A single disease or health outcome (e.g., cardiovascular mortality) is studied. Persons with the disease or outcome are selected from a defined geographical area, hospital(s), clinic(s), or even a cohort. A comparison group of persons in which the condition or disease is absent is also selected, preferably at random from the same population from which the cases were selected. Existing or past attributes and exposures thought to be relevant in the development of the disease are determined for all participants, and the frequency of exposure is compared among persons with and without the disease. Information about the relevant exposure or behavior is obtained by interview or other means. For example, participants may be asked about their smoking habits (number of cigarettes smoked per day) or a blood specimen may be collected and analyzed for DNA adducts. A residence history is required to determine long-term drinking water exposures. Water quality information (e.g., water hardness, magnesium, or calcium levels) can then be obtained from water quality records at the appropriate water system. If information about current tap water quality is important, a sample can be collected and analyzed. The participant may also be questioned about water consumption patterns and other exposures (e.g., dietary calcium and magnesium intake). Sometimes, information must be obtained by questioning a spouse or care giver. When collecting information from participants or others, it should be recognized that it may be difficult to accurately recall exposures or events that occurred many years ago. Investigators should ensure that the quality of the information is similar for both cases and controls.

An advantage of the cohort study is that more than one health-related outcome or disease can be studied. To avoid selection bias, the cohort should be assembled solely on the presence or absence of certain characteristics, a specific event, or their exposure status (e.g., water hardness; high, moderate, or low levels of calcium or magnesium in drinking water). Morbidity or mortality incidence is determined for the diseases of interest, and rates are compared for exposed and unexposed groups in the cohort. A cohort can be based on currently-defined exposures or historical exposures and followed forward from that point in time. Thus, the approach can be prospective, retrospective, or both. For example, if a cohort were established based on known drinking water exposures to magnesium from 1980 to the present, 23 years of exposure would have already occurred. The analysis could evaluate risks associated with this period of exposure or continue to follow the cohort for two or five years and consider 25 or 30 years of exposure.

A cohort study can be conducted when a community changes its water treatment practice or water source. Either individual- or group- information can be evaluated. Community-intervention studies have helped demonstrate the effectiveness of water fluoridation in preventing dental caries, and as noted earlier, several studies have evaluated cardiovascular disease associated with changes in hardness of a community's drinking water. An advantage of this type study is that water quality is changed at all places where persons may consume water (e.g., home, school, work, restaurants) minimizing exposure misclassification bias. However, an important limitation to consider is that many years of exposure may be required to effect a change in the disease risk. Population demographics, behaviours, and other risk factors may also change during the study period and should be taken into account when analyzing the data. Also, the areas may not be optimal in terms of desired water quality or population characteristics.

### **3 THE EXPOSURE-DISEASE ASSOCIATION**

The correlation coefficient ( $r$ ) frequently reported in ecological studies does not provide reliable, quantitative information about the risk or benefit of an observed association. In contrast, analytical studies can provide an estimate of the magnitude of the risk or benefit. The basic measures are the rate difference (RD) and rate ratio (RR). The RD is a measure of the absolute difference between two rates (e.g., incidence rate of cardiovascular disease for the exposed minus the incidence rate for the unexposed in a cohort study). The RR is a relative measure of two rates (e.g., incidence rate for the exposed divided by the incidence rate for the unexposed in a cohort study). The RR is also called the relative risk. A reported RR of unity (1.0) indicates no association and no increased risk; any other value signifies either increased or decreased risk. For example, a RR of 1.8 indicates an 80 percent increased relative risk of disease; a RR of 0.8 indicates a decreased risk of 20 percent. Because participants in a case-control study are selected according to their disease status, the exposure odds ratio (OR) is determined. The OR is the odds or chance of disease among the exposed divided by the odds of disease among the unexposed. The OR is essentially equivalent to the RR (Monson 1990). Another important, but often misunderstood measure, is the population attributable risk (PAR). Often computed as a percentage, the PAR provides an estimate of the incidence of a disease in a population that is associated with or attributed to the exposure or risk factor in question, provided the association is causal (Last, 1995). A PAR computed for

water hardness should be cautiously interpreted, since cardiovascular disease has many known and suspected risk factors. The individual PARs computed for several potential risk factors for a disease may add up to over 100% (Monson, 1990).

### 3.1 Random and systematic error

The epidemiological association should be evaluated for possible systematic and random error (Table 2.2). The likelihood that an observed association is due to random error is assessed by the level of statistical significance ("p" value) or the confidence interval (C.I.). The C.I. is the preferred measure because it provides a range of possible values consistent with the risk estimate. For example, a reported RR of 1.8 with a 95% C.I. of 1.6-2.0 indicates both a precise and statistically significant estimate because the C.I. is narrow and does not include 1.0. A reported RR of 2.7 with a C.I. of 0.8-14.5 indicates that the risk estimate is not statistically significant (the C.I. includes 1.0) and imprecise (a wide range of values). It should always be remembered that random error or chance can never be completely ruled out as the explanation for an observed association and that statistical significance does not imply causality, biological significance, or the lack of systematic error.

Table 2.2. Assessing bias for reported associations

Lack of Random Error (Precision)	Lack of Systematic Error (Validity)
Study Size and Statistical Power	Misclassification Bias Selection Bias Observation Bias Confounding

Systematic error affects the validity of an observed association and can occur in the design or conduct of the study. In either case, it leads to a false or spurious association and a measure of risk that departs systematically from the true value. Selection bias occurs when criteria are not comparable for enrolling cases and controls or exposed and unexposed persons into the study. Observation bias occurs when disease or exposure information is collected differently from participants. For example, cases may selectively or differentially recall the exposure (Monson 1990).

The erroneous classification of a study participant's exposure or disease will result in misclassification bias. Differential misclassification bias can result in associations that either under or over estimate the magnitude of risk. Non-differential misclassification will almost always result in not observing an association when one may actually be present, thereby under estimating the risk. In environmental epidemiological studies where the magnitude of the association is often small, accurate assessment of exposure is critical, as the impact of exposure misclassification can be severe. The imprecise estimate of water hardness in many of the cardiovascular disease studies presents a potential for misclassification bias.

A confounding characteristic rather than the suspected cause or exposure may be responsible for all or much of the observed association. Confounding does not necessarily result from an error of the investigator. It is potentially present in all epidemiological studies and should be considered as a possible explanation for any observed association. Cigarette smoking may confound many associations.

Investigators may control confounding in the study design and assess it during the analysis (e.g., regression techniques and stratification). A confounder is a characteristic that can cause or prevent the disease and is associated with the exposure being evaluated. Thus, if the suspected characteristic (e.g., smoking) can be shown to have no association with exposure (e.g., water hardness), the characteristic cannot confound the association that may be observed between the exposure and the disease. Confounding can be controlled by using a technique known as matching in the design of the study. For example, controls may be selected to have similar characteristics (e.g., smoking status) as cases. In this example, all study participants will either be smokers or non-smokers. Effect modification refers to a change in the magnitude of the effect of a putative cause (Monson 1990) and should not be confused with confounding. For example, if cardiovascular disease risks associated with water hardness differed among smokers and non-smokers, smoking would be considered an effect modifier.

Table 2.3. Assessing the strength of an epidemiological association \*

Rate Ratio (Increased risk)	Rate Ratio (Decreased risk)	Strength of Association
1.0 – 1.2	0.9 – 1.0	None
1.2 – 1.5	0.7 – 0.9	Weak
>1.5	< 0.9	Moderate to Strong

\*Adapted from Monson (1990)

### 3.2 Strength of an association

An increased risk of less than 50% (RR=1.0–1.5) or a decreased risk of less than 30% (RR=0.7–1.0) is considered by many epidemiologists to be either a weak association or no association (Table 2.3). Based on Monson’s experience (1990), it may be difficult to interpret associations with these RRs. Confounding can lead to a weak association between exposure and disease, and it is usually not possible to identify and adequately measure or control weak confounding characteristics. For weak associations, investigators should thoroughly evaluate the possibility that the association is affected by uncontrolled confounding. On the other hand, a very large increased or decreased RR is unlikely to be completely explained by an unidentified or uncontrolled confounding factor. The magnitude of a RR, however, has no bearing on the possibility that an association is due to observation, selection, or misclassification bias. Systematic error, even in studies with a large RR, may lead to a spurious association.

### 3.3 Causality of an association

Epidemiologists have debated how to make causal inferences from observed associations. Even though an association is repeatedly observed, investigators may question whether the association constitutes an “empirical demonstration that serves as a valid platform for (causal) inference” or whether “the process is still steeped in uncertainty” (Rothman 1986). Interpretation of epidemiological results should always be made with caution and in the context of all relevant biological information about the disease. Epidemiologists generally agree that no single epidemiological study, even one with little systematic error, can provide a definitive answer about the exposure and its effect. Results from a relatively large number of studies in various geographical areas allow for a more definitive assessment of the causality of an association. However, the design, precision, and validity of the individual studies must be evaluated before the evidence for causality is considered.

Epidemiologists judge the causality of an epidemiological association (Hill 1965; Rothman 1986; Beaglehole *et al.* 1993) based on the following guidelines:

- Temporal Association. Exposure must precede the disease, and in most epidemiological studies this can be inferred. In studies where exposure and disease are measured simultaneously or exposure is measured after the occurrence of disease, the temporal association should be evaluated.
- Specificity. A supposed cause or exposure leads to a specific disease rather than to a general effect or multiple diseases. The presence of specificity argues for causality, but its absence should not rule it out.
- Consistency. Repeated observation of an association under different study conditions supports an inference of causality, but its absence should not rule it out.
- Biological Plausibility. When the association is supported by evidence from clinical research or basic sciences (e.g., toxicology, microbiology) about biological mechanisms, an inference of causality is strengthened.
- Strength of Association. As noted earlier, the greater the magnitude of risk or benefit, the less likely the association is to be spurious or due to confounding bias. However, a causal association should not be ruled out simply because a weak association is observed.

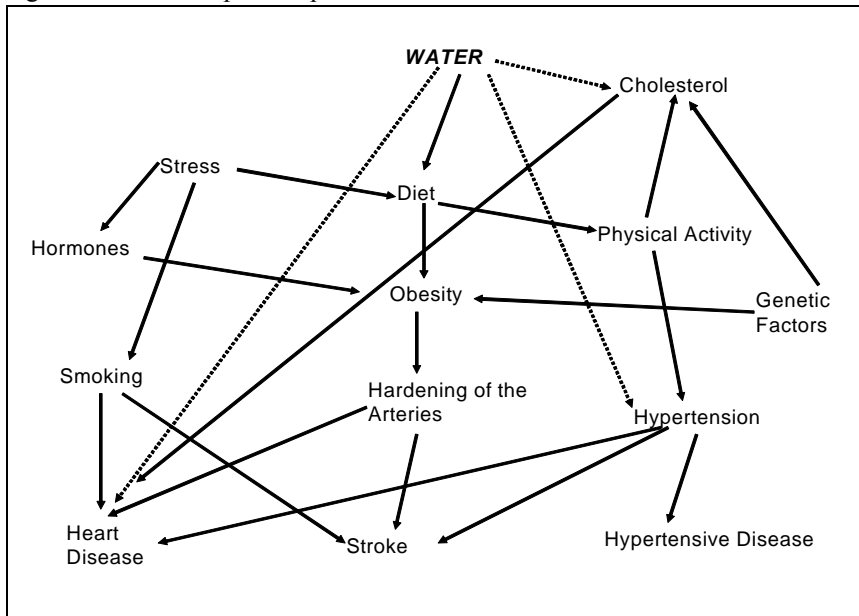
- Dose-Response Relationship. A causal interpretation is more plausible when an exposure- or dose-response gradient is found (e.g., a higher risk is associated with larger exposures).
- Reversibility. An observed association leads to some preventive action, and removal or reduction of the exposure should lead to a reduction of disease or risk of disease.

Scientific evidence for causality may be sparse and inconsistent, and scientists may offer conflicting opinions about causality. When environmental policy makers and regulators are confronted with epidemiological associations that suggest the need for action, they should consider the uncertainties about causality. Most regulatory actions will require a high level of certainty about the causal nature of the association. However, it should also be recognized that effective public health actions may be taken even with incomplete knowledge about causality. For example, Dr. John Snow in mapping cholera deaths in London more than one hundred and fifty years ago was able to associate increased deaths with use of the Broad Street well. This information was considered sufficient to remove the pump handle and preventing use of the well. This action was taken well before the etiological agent *Vibrio cholerae* was identified and scientists understood the importance of the transmission of cholera through water contaminated by human feces. More recently, the mounting epidemiological evidence of increased lung cancer risks among smokers prompted public health warnings about cigarette smoking in the 1960s long before specific carcinogens were identified.

### 3.4 Web of causation

Many diseases have multiple exposures or risk factors that cause the disease or increase the disease risk, and the disease process is often complex. This complexity is evident in an example of a conceptual model that might be used to describe the relationship between various exposures and risk factors for cardiovascular disease (Figure 2.1). This model, often referred to as the web of causation, places less emphasis on the role of a single agent (e.g., a water constituent) in favour of other factors that may be important in the onset of disease (Rockett 1994). When evaluating the role of water hardness in the etiology of cardiovascular disease, it is important to consider how various risk factors might affect not only disease but also exposure. In the example provided, a dotted line is used to suggest that additional evidence may be warranted for waterborne exposures and how these exposures may affect the disease process.

Figure 2.1 An example of a possible disease model for cardiovascular disease\*



\*Adapted from Rockett (1994)

## 4 CONCLUSIONS

## 5. REFERENCES

Numerous ecological and analytical studies during the past forty-five years have provided epidemiological information about health benefits associated with hard water. Results from these studies can be evaluated to assess the causality of the association and potential magnitude of the benefit. This information can help public health officials make decisions about recommendations for adding certain constituents to demineralised and desalinated waters.

- Beaglehole, R., Bonita, R., Kjellstrom, T. (1993) *Basic Epidemiology*, World Health Organization, Geneva, 1993, pp. 71-81.
- Comstock, G. (1979a) Water hardness and cardiovascular diseases. *Amer. J. Epidemiol.* **110**(4):375-400.
- Comstock, G. (1979b) The association of water hardness and cardiovascular diseases: An epidemiological review and critique. *Geochemistry of Water in Relation to Cardiovascular Disease*, National Academy of Sciences, Washington, DC, p. 48-68.
- Enterline, P. and Stewart, W. (1956) Geographic patterns in deaths from coronary heart disease. *Public Health Rpts.* **71**:849-855.
- Greenland, S., and Robins, J. (1994a) Invited commentary: ecologic studies--biases, misconceptions, and counter examples. *Amer. J. Epidemiol.* **139**:747-760.
- Greenland, S., and Robins, J. (1994b) Accepting the limits of ecologic studies. *Amer. J. Epidemiol.* **139**:769-771.
- Hill, A. (1965) Environment and disease: association or causation? *Proc. Royal Soc. Med.* **58**:295-300.
- Hornstra, G., Barth, C.A., Galli, C., Mensink, R.P., Mutanen, M., Riemersma, R.A., Robertfroid, M., Salminen, K., Vansant, G., Verschuren, P.M. (1998) Functional food science and the cardiovascular system. *Brit. J. Nutr.* **80**(suppl. 1): S113-S146.
- Kobayashi, J. (1957) On geographical relationship between the chemical nature of river water and death-rate from apoplexy. *Berichte des ohara institutes fur landwirtschaftliche biologie*, **11**:12-21.
- Piantadosi, S. (1994) Invited commentary: ecologic biases. *Amer. J. Epidemiol.* **139**: 71-64.
- Poole, C. (1994) Editorial: ecologic analysis as outlook and method. *Amer. J. Public Health* **84**: 715-716.
- Last, J.M. 1995 *A Dictionary of Epidemiology*, 3<sup>rd</sup> edition, Oxford University Press, New York, Oxford, Toronto, 180 pp.
- Monson, R. (1990) *Occupational Epidemiology*, 2<sup>nd</sup> edition, CRC Press Inc., Boca Raton, Florida, 291 pp.
- Rothman, K. (1986) *Modern Epidemiology*, Little, Brown and Company, Boston, pp. 7-21.
- Rockett, I. (1994) Population and health: An introduction to epidemiology, *Population Bull.* **49**(3):11.
- Sauvant, M-P. and Pepin, D. (2002) Drinking water and cardiovascular disease. *Food Chem. Toxicol.* **40**:1311-1325.
- Schroeder, H. (1960a) Relations between hardness of water and death rates from certain chronic and degenerative diseases in the United States. *J. Chronic Dis.* **12**:586-591.
- Schroeder, H. (1960b) Relationship between mortality from cardiovascular disease and treated water supplies. Variations in states and 163 largest municipalities of the United States. *J. Amer. Med. Assoc.* **172**:1902-1908.
- Schroeder, H. (1969) The water factor. *New Engl. J. Med.*, **280**:836-838.