AJIC current methodological concepts

Causal inference: The case of hygiene and health

Allison E. Aiello, MS[a]
Elaine L. Larson, PhD, RN, FAAN, CIC[b]
New York, New York

A fundamental goal of applied epidemiology is to determine whether a relationship between 2 factors is causal. For example, the primary purpose of an outbreak investigation is to identify what factor(s) "caused" the problem, and the purpose of the Study of the Efficacy of Nosocomial Infection Control (SENIC Project) was to measure the effect of infection control and prevention programs on rates of nosocomial infections (ie, whether strong programs were associated with or "caused" a reduction in infections).[1] There have been a number of historic efforts to formulate methods for valid causal inference.[2-5] To the layperson, association between 2 variables is often assumed to be causal, but from the epidemiologic point of view, this is not the case. For example, in studies concerning health and hygiene, associations have been found between hand hygiene interventions and a decline in diarrhea and between laundering practices and prevalence of infectious disease in the home.[6-9] Such associations, however, must properly reflect an underlying causal mechanism and must have been investigated with use of rigorous methodologic procedures.[10-12] Our purpose is to discuss 1 of the historic methods devised for inferring causation and the epidemiologic concepts that relate to causal inference, with the example of health and hygiene.

From Joseph Mailman School of Public Health, Columbia University,[a] and Columbia University School of Nursing.[b]

Funded in part by the Soap and Detergent Association, Washington, DC.

Reprint requests: Elaine L. Larson, PhD, RN, FAAN, CIC, Columbia University, School of Nursing, 630 W 168 St Mailbox 6, New York, NY 10032.

Am J Infect Control 2002;30:503-10.

0196-6553/2002/$35.00 + 0 **17/52/124585**

doi:10.1067/mic.2002.124585

HILL'S CAUSAL CRITERIA

In 1965, Hill[4] suggested a set of criteria for clarifying the difference between a noncausal and causal association. These criteria were first applied to the relationship between smoking and lung cancer and have since been applied to many other exposure and disease relationships. Although the following have been dubbed "criteria," they should be considered as simplified guidelines for assessing causation, since there are situations in which they may not be tenable. For example, it is possible to fulfill all of the criteria even when there is no underlying causal relationship because the association is actually explained by a third variable (see confounding in Fig 1). Conversely, it is possible to have an underlying causal relationship even when none of the criteria except temporality is fulfilled. The guidelines are summarized in the following, and examples related to hygiene practice and improved health are presented in Table 1.

Strength

What is the magnitude of the association between 2 variables? What are the changes in frequency of disease as indicated by the point estimates known as the *risk ratio (RR), incidence density ratio (IDR)*, or *odds ratio (OR)* along with the respective *confidence interval (CI)*? The RR and OR are estimates that measure association between 2 variables. The IDR is an estimate that measures the rate with which new cases of disease occur in the exposed relative to the unexposed. The RR is a ratio of the following 2 probabilities: the probability of disease in the exposed group as compared with that in the unexposed group (eg, the risk of developing diarrhea among those who were exposed to an egg

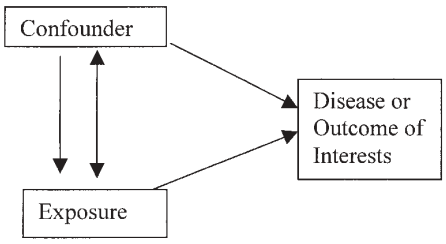
Type of Relationship	Example
Relationship of the Exposure and Effect 	Consumption of unwashed strawberries contaminated with hepatitis A is related to infection with hepatitis A in a susceptible individual.
Confounding 	The relationship between a mothers method of food preparation as a risk factor for diarrheal disease in a child may be explained by economic status. Economic status is a confounder of the relationship because it is an independent risk factor for diarrheal disease and is associated with the mother's method of preparing food. In addition, it is not in the causal pathway from exposure to disease. Failure to control adequately for the confounder may result in a false association, exaggeration or attenuation of the association.

Fig 1. Confounding, mediation, and effect modification examples in the relationship between hygiene and health. *Continued on next page.*

salad contaminated with *Salmonella* compared with the risk among those who were not exposed). The OR, generated when the study design used is a case-control study or when multivariate logistic regression is used to adjust for potential confounding, is defined as the odds of disease among the exposed as compared with those of the unexposed. The CI provides information on the precision of the point estimate (RR, IDR, or OR) and the magnitude of the association.

Consistency

Do the studies concerning health and hygiene show a persistent association with different study designs, populations, at different times, or geographic areas and according to various study investigators? Finding consistent results may reduce the possibility that the association is due to chance or error. For example, many studies conducted in various geographic regions and with differing study methods report a reduced rate of infection associated with implementation of hygiene intervention programs, although the magnitude of the effect estimates vary.[13,14]

Specificity of the association

Is there a 1-to-1 relationship between "cause" and "effect?" Does 1 cause have a single effect? In 1890, Koch[15] demonstrated a causal association between *Bacillus anthracis* and anthrax with use of the following 3 postulates: (1) the organism is present in every cases of disease; (2) the organism does not occur in other diseases as a nonpathogenic agent; and (3) the organism can be isolated, cultured, and exhibit subsequent infectivity. It is now known that

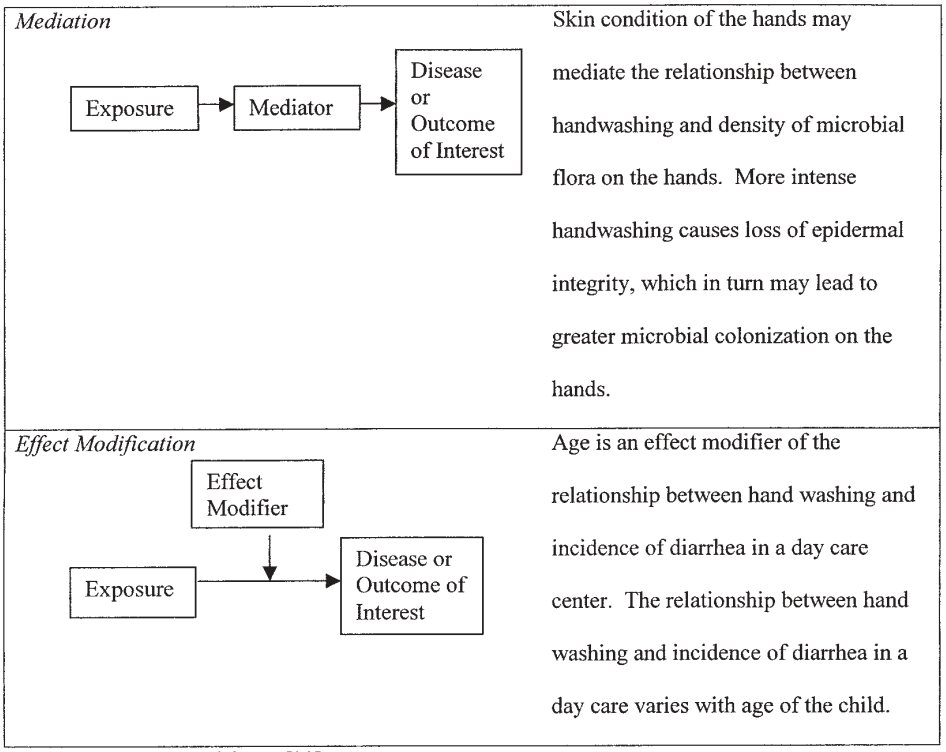


Fig 1. *Continued from previous page.*

these postulates do not apply to many infectious organisms. For example, a person may be a carrier of an agent without ever manifesting overt disease, thereby violating the second postulate. In addition, some organisms cannot be cultured and may not always exhibit subsequent infectivity. A specific 1-to-1 relationship is rare since numerous factors must work together to cause disease.

Temporality

Which came first? Did the "outcome" occur after the "cause" was introduced? A temporal relationship is most easily examined in an intervention study because the investigator is assured that the exposure preceded the outcome. There are numerous studies suggesting reductions in diarrheal disease after implementation of appropriate handwashing measures.[6,7,13,16-20] For example, Khan[20] reported that handwashing with soap and water after defecation and before ingesting food was followed by an 84% reduction in shigellosis. Hence, the intervention (handwashing) preceded the reduced incidence of shigellosis. Even with the temporal sequence that a hygiene intervention demonstrates, it is possible that the reduction was attributed to some other factors acting contemporaneously with the reduction. Therefore, it is important to hold these other factors constant (by matching, stratifying, or applying multivariate regression) to attribute the reduction to the hygiene intervention.

Plausibility

Is there a plausible and reasonable biologic or behavioral rationale to explain how 1 factor might "cause" another? For example, does it make biologic sense that a lessened microbial load in the kitchen environment would reduce risk of infectious disease transmission? Buchanan and Whiting[21] examined the risk of *Salmonella* colonization on raw food after different storage and cooking temperatures. Their risk modeling shows that reducing the number of organisms affects the probability of infection. Given that our current understanding of disease pathogenesis is often limited, the lack of scientific plausibility alone should not deter us from inferring causality.

Biologic gradient

Is there evidence of a dose-response relationship? Many, but not all, organisms exhibit a biologic gradient upon exposure. In the case of hygiene, is the risk of infection reduced as hygiene is improved? For

Table 1. Hill's causal criteria applied to hygiene and infection

Criterion	Description	Example questions
Temporality	Hygiene practice must precede reductions in infection	Was increased handwashing during an outbreak associated with subsequent reductions in infections?
Strength of association	Risk of infection among those in which the hygiene practice is present must be less than the risk of infection among those in which the practice is absent	In an elementary school, what is the rate of absenteeism due to infections among children who participate in a hand hygiene education program as compared with children who do not participate in such a program? What is the risk of diarrheal disease among families who have soap available and those who do not?
Consistency of association	Association between reduced risk of infection and the hygiene practice is present in various communities, across various studies, and in a variety of circumstances	Is the association between hygiene practices and reduced infection present in the following circumstances: In developed as well as developing countries? Across settings such as long-term care, schools, homes, hospitals, schools? Across various product brands?
Specificity of association	A unique association between hygiene and risk of infection is present	Do people practicing good hygiene ever develop infections? Are infections present only in those with poor hygiene?
Plausibility	There is a biologic explanation or pathophysiologic mechanism for the association between hygiene and infection	Is it biologically feasible that a reduced microbial load on food preparation surfaces would reduce the risk of foodborne infections? Is there a biologic mechanism by which contaminated hands could increase the spread of infection?
Biologic gradient	Dose response between hygiene practices and infections is present (ie, more hygiene results in fewer infections, less hygiene results in more infections)	What is the relative risk of diarrheal disease among communities in which various levels of hygiene can be measured? Do rates of infection vary among children in preschools with varying levels of cleaning procedures? What is the impact on skin infection rates of a daily as compared with a 3-times weekly bathing routine in a long-term care facility?
Experimental evidence	Evidence is present from studies that test the effect of a hygiene intervention on subsequent rates of infection	Is the introduction of a waterless hand rinse in refugee camps associated with reduced rates of infection?

example, Cabelli et al[22] showed a linear relationship between the amount of fecal-indicator bacteria, such as enterococci and *Escherichia coli*, in swimming water and swimming-associated gastroenteritis. Further, there is a reciprocal relationship between host susceptibility and agent dose; the weaker and more immunocompromised the host, the fewer the microbes needed to generate an infection. Hence, the presence or absence of a dose-response effect may not necessarily be evidence of causality.

Experimental evidence

Are there changes in health when hygiene interventions are introduced? Carabin et al[23] reported a considerable reduction in incidence of diarrhea and respiratory illness after implementation of a hygiene program in toddlers attending daycare centers. A handwashing intervention study[7] in 1996 showed a large reduction in diarrheal episodes in an economically depressed village in Dhaka City, Bangladesh. Caution in interpreting results is advised, even in experiments, because although a change in the outcome may have been preceded by the intervention, it is possible that the change could be the result of another exposure altogether. For example, in the 1800s foul odors were considered to be the causative agent of disease in the home. Upon removal of the foul odors by construction of appropriate plumbing, rates of many infectious diseases were reduced. The reduction was incorrectly attributed to the removal of the odor, rather than protection from fecal contamination afforded by higher-quality sanitation facilities.

METHODOLOGIC ISSUES IN CAUSAL INFERENCE

Clearly the elements of causality may be helpful for inferring causation for some associations, but they should not be applied to studies that are method-

Table 2. Examples of methodologic issues in hygiene and health

Issue: Inadequate sample size or lack of power
Example in hygiene and health: A study investigator interested in conducting an intervention trial of hand hygiene to reduce the incidence of diarrhea in a community may calculate an appropriate sample size to have enough power to detect a statistical difference in the rate of diarrhea. If the target sample size was not met for the study and the study resulted in no statistically significant reduction in diarrhea, then one cannot ascertain whether the observed lack of association is due to low statistical power or truly reflects no association.
Ways to minimize problem in study design: Use sample size calculation and recruit enough subjects for adequate power.
Issue: Selection bias
Example in hygiene and health:
Case-control studies: Controls in a case-control study must represent the same source population from which the cases arose. For example, if controls do not represent the source population of *Salmonella* cases that had been sampled from a school-wide outbreak due to consumption of a contaminated salad in the school lunch, any association between the cases and controls may be biased because of confounding.
Cohort studies: Loss to follow-up is a type of selection bias in a cohort study. An association will be biased if the proportion of people leaving a study is different among the exposed and unexposed and is also dependent on their disease status. For example, in a cohort study of the relationship between antibacterial soap use and incidence of diarrhea, if individuals who do not use antibacterial soap leave the study when they develop diarrhea, the observed association would be biased.
Ways to minimize bias in study design:
Case-control studies: Sample controls from the same source population from which the cases arise. Control group should be representative of the exposure distribution in the source population.
Cohort studies: Avoid loss to follow-up. Use equivalent methods for following-up all study subjects equally.
Issue: Recall bias
Example in hygiene and health: Subjects may selectively recall events. For example, the mother of a child with a severe case of diarrhea may be more likely to remember specific home hygiene practices. If this recall occurred more in the mothers who have a child with a case of diarrhea than in those who do not, the association may appear larger than it actually is.
Ways to minimize bias in study design: Use reliable and structured interview methods. Avoid questions that require lengthy recall.
Issue: Detection bias
Example in hygiene and health: The individual who is diagnosing subjects who have been randomized to the group lacking a hygiene intervention may inadvertently increase intensity of the search for symptoms of an infectious illness. If subjects who lack the hygiene intervention were diagnosed with related illness simply because it was identified more often, the association would appear larger than it actually is.
Ways to minimize bias in study design: Use equivalent detection methods for exposed and unexposed subjects. Use blinding of exposure status for interviewers.
Issue: Interviewer bias
Example in hygiene and health: If the study interviewers are not blinded to the intervention, exposure, or disease status, they may have a preconceived opinion and coax the subjects or search for answers more intensely in 1 group than another. For example, in a study of the relationship between housecleaning and childhood rotavirus infection, the interviewer may probe for deficiencies in housecleaning practices in families who have experienced higher rates of infection with rotavirus than those who have not. In this example, the association would appear larger than it actually is.
Ways to minimize bias in study design: Whenever possible, use blinding of exposure and disease status of subjects. Train interviewers to obtain responses in an unbiased manner.
Issue: Nonresponse bias
Example in hygiene and health: Individuals who choose not to participate in a study may differ from those who do participate with regard to characteristics that are related to the relationship being studied. For example, mothers with children that are stricken more often with diarrhea may be more likely to allow their child to participate in a hygiene intervention study at their daycare center. If these children are systematically different from children who do not participate on the basis of characteristics that are related to the study outcome, such as immune status, sex, or age, the association will be biased.
Ways to minimize bias in study design: The study should elicit subjects in a way that allows for a nondifferential response rate. Nonresponders should be followed-up and compared with responders to determine how they differ.

ologically flawed by inadequate sample size, confounding, and systematic bias. As previously mentioned, the causal criteria become untenable when numerous studies that show a strong and consistent association are assessed but are riddled with unrecognized biases and unmeasured confounders (see some examples in Table 2).

At the onset of a study, a methodologic problem that can be avoided is that of sampling error. It results

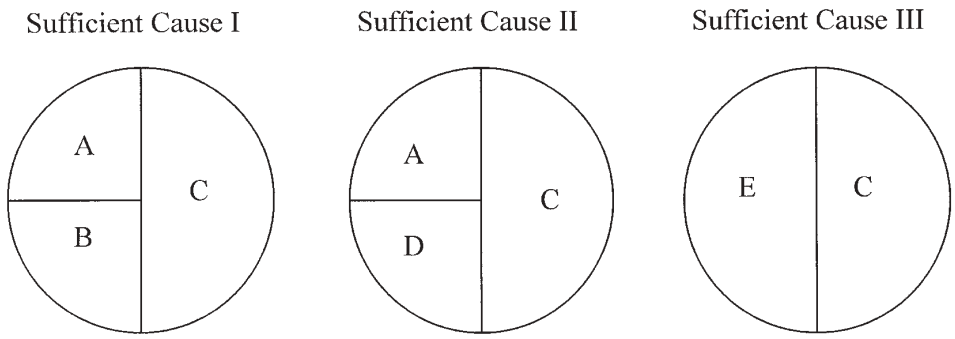


Fig 2. Three sufficient causes of shigellosis in a daycare center. *Sufficient cause I* consists of the following causal risk factors: **(A)** inadequate hand hygiene after staff changes diapers of child infected with the *Shigella* organism, **(B)** subsequent preparation and contamination of snacks for the children in daycare, and **(C)** susceptible child ingests *Shigella*-contaminated snack. *Sufficient cause II* consists of the following causal risk factors: **(A)** Inadequate hand hygiene after staff changes diaper of child infected with *Shigella* organisms, **(D)** staff-to-child hand contact, and **(C)** susceptible child puts *Shigella*-contaminated hands into mouth and ingests organisms. *Sufficient cause III* consists of the following causal risk factors: **(E)** child infected with *Shigella* organisms spreads feces from diaper onto toys and **(C)** susceptible child puts *Shigella*-contaminated toys into mouth and ingests organisms.

from inadequate sample size and can be reduced by calculating and obtaining an appropriate target sample size at the beginning of the study. If the required sample size is not met, it is difficult to determine whether a lack of a statistically significant observed association was due to chance (ie, type II statistical error). Biases, on the other hand, are systematic errors that may go unnoticed by the study investigator. Hence, it is best to minimize them with rigorous study design methods (see Table 2).

Although most diseases are the result of interacting multiple factors, epidemiologic methods are designed to isolate the effects of a particular risk factor. To examine a specific factor, investigators usually adjust for other factors that may act as confounders during the analysis. Properties of a confounder include the following: (1) being an independent risk factor for the disease of interest, (2) being associated with the exposure of interest in the source population, and (3) not being in the causal pathway from the exposure to the disease (ie, not a mediating factor). Confounding can either suppress or inflate an association. In fact, an uncontrolled confounder may be responsible for an entire association when none truly exists. Common methods for identifying and reducing confounding include randomization and statistical methods such as stratification or regression analyses.

Mediation and effect modification are methodologic phenomena that should be assessed to more clearly understand the relationship between a risk factor of interest and an outcome. Mediation is a condition in which the exposure under investigation exerts its effect through another risk factor that is in the pathway from the exposure to the disease. Effect modification is related to statistical interaction and is present when the exposure and disease relationship varies, depending on the level of another factor. By assessing effect modification, the investigator is able to consider how risk factors work together to cause disease.[24] See Fig 1 for examples of confounding, mediation, and effect modification in the relationship between hygiene and health.

To further understand causal inference, clinical investigators and discerning readers of the medical literature must also consider how multiple risk factors work together to cause disease and the conceptual meaning of "magnitude of association."

SUFFICIENT AND COMPONENT CAUSES

Rothman[24] devised a theory that demonstrates that "strength of association" and "consistency" are dependent on the prevalence of other risk factors in a given population and therefore do not represent the biologic impact of a risk factor. His pictorial model of causal pies, termed "sufficient causes," and the inner pieces of the sufficient cause, termed "component causes," elucidates how factors work together to cause disease and why removing any 1 risk factor can reduce illness in a population.[24]

For a simplified example of Rothman's causal pies, one could propose that in a given daycare center there are only 3 sufficient causes of shigellosis (Fig 2). Each sufficient cause consists of multiple risk factors (see Fig 2, *A-E*) and represents a potential pathway to shigellosis for each child in the daycare center. Therefore, the risk for an individual child during a given period is the probability of his or her sufficient cause being completed. Hence, this causal schema assumes that the outcome of 1 child is independent of any other children in the daycare.

Any single risk factor alone is not sufficient to cause disease and must interact with the other risk factors to complete the sufficient cause of disease.[24] Further, one does not need to identify all of the causal risk factors to prevent a sufficient cause of a disease. This is analogous to the concept of breaking the chain of transmission: it does not matter which risk factor is removed since removal of any 1 of the risk factors would eliminate a given pathway to disease. For example, if *E* (see Fig 2) was removed, then *sufficient cause III* (see Fig 2) could not be a mechanism for spread of shigellosis; only *sufficient causes I* and *II* (see Fig 2) would be responsible for disease occurrence in the daycare. Removing the necessary causal risk factor *C* (ingesting organisms)(see Fig 2) would eliminate all sufficient causes of disease in this simplified example.

The schema also illustrates that the strength of a risk factor as measured by the RR or OR is determined on the basis of the prevalence of the causal partners. The greater the prevalence of the causal partners in a sufficient cause, the stronger the risk factor appears. For example, if one is concerned with risk factor *A* (inadequate hand hygiene after staff changes a diaper on a child infected with the *Shigella* organism) in *sufficient cause I* (see Fig 2) and the prevalence of the component causes *B* (subsequent preparation and contamination of snacks for the children in the day care) and *C* (susceptible child ingests *Shigella*-contaminated snack) are very high in the study population, then the strength (RR) of the risk factor (component cause *A*) would appear to be very large. Consequently, the magnitude of the RR depends on the prevalence of the causal partners in the population and is not a biologic representation of the risk factor and disease relationship of interest.[24]

This concept can be demonstrated in greater detail with use of the example of rotavirus infection in daycare centers. For example, staff in 1 center change diapers frequently without adherence to proper sanitary measures and most of the children attending the daycare have previously experienced rotavirus infection (only a few children are still susceptible). In this instance, a history of no previous rotavirus infection would appear to be a strong risk factor for infection and frequent diaper change without sanitary precautions would appear to be a weak risk factor. On the other hand, in another daycare center where most of the children are susceptible to rotavirus infection and diapers are usually changed with adherence to proper sanitary measures, susceptibility would appear to be a weak risk factor and diaper changing without proper sanitary measures would appear to be a strong risk factor.

This concept demonstrates that a strong association is not necessarily greater evidence that a relationship is causal and that a weak association does not prove that the relationship is less likely to be causal. A weak association may just indicate that the prevalence of the partnering causal risk factors (other than the risk factor of interest) is very low in the study population. Since the magnitude of association is dependent on the prevalence of the background risk factors, an inconsistent association among a number of studies in health and hygiene may nevertheless be causally related but vary with respect to the prevalence of the causal risk factors among differing study populations.

Although Rothman's multiple risk factor model represents a major advance in applied epidemiologic methods, there are limitations because of the inherent assumption that the outcome for 1 individual is independent of the outcome for another individual. For most infectious diseases, individual risks are interdependent, a consequence of dynamic environmental and interpersonal interactions. A daycare center is just 1 example of a setting in which there is constant interchange between the environment and its inhabitants. Some have attempted to devise methods to analyze dependent happenings (eg, if 1 child falls ill with shigellosis in a daycare center, he or she may pass on the disease to another child in the same daycare center so that the risks are now considered "dependent") and feed-back loops, such as commonly occurs in infectious disease relationships.[25-27] Current research on these specialized methods may shed light on the dynamic relationships between hygiene interventions, infectious agents, host susceptibility, and the environment.

In summary, the application of causal inference is important for enhancing the public health decision-

making process. During the past few decades, our ability to apply causal modeling to issues such as the association between hygiene and health has been enriched by Hill's classic articulation of criteria for causation and Rothman's demonstration that it is possible to reduce risk of disease by intervening in various "component causes," even when the entire "sufficient cause" is not fully understood. The field of applied epidemiology will continue to evolve and facilitate our efforts to reduce the public's burden of disease.

We thank Ulka B. Campbell at Joseph L. Mailman School of Public Health for carefully reading the manuscript and providing many thoughtful comments.

References

1. Haley RW, Culver DH, White JW, et al. The efficacy of infection surveillance and control programs in preventing nosocomial infections in US hospitals. Am J Epidemiol 1985;121:182-205.
2. DHHS. US Surgeons General's Advisory Committee on Smoking and Health. Washington (DC): US Department of Health, Education and Welfare, Public Health Service; 1964.
3. Evans AS. Causation and disease: the Henle-Koch postulates revisited. Yale J Biol Med 1976;49:175-95.
4. Hill A. The environment and disease: association or causation? Proc R Soc Med 1965;58:295-300.
5. Susser M. Criteria of judgement. Causal thinking in the health sciences: concepts and strategies in epidemiology. New York: Oxford University Press; 1973. p. 140-62.
6. Han AM, Hlaing T. Prevention of diarrhoea and dysentery by hand washing. Trans R Soc Trop Med Hyg 1989;83:128-31.
7. Shahid NS, Greenough WB III, Samadi AR, Huq MI, Rahman N. Hand washing with soap reduces diarrhoea and spread of bacterial pathogens in a Bangladesh village. J Diarrhoeal Dis Res 1996; 14:85-9.
8. Wilson JM, Chandler GN, Muslihatun J. Hand-washing reduces diarrhoea episodes: a study in Lombok, Indonesia. Trans R Soc Trop Med Hyg 1991;85:819-21.
9. Larson E, Gomez Duarte C. Home hygiene practices and infectious disease symptoms among household members. Public Health Nurs 2001;18:116-27.
10. Buck C. Popper's philosophy for epidemiologists. Int J Epidemiol 1975;4:159-68.
11. Susser M. What is a cause and how do we know one? A grammar for pragmatic epidemiology. Am J Epidemiol 1991;133:635-48.
12. Cook T, Campbell D. Quasi-experimentation: design and analysis issues for field settings. Boston: Houghton Mifflin Co; 1979.
13. Aiello AE, Larson EL. What is the evidence for a causal link between hygiene and infections? Lancet Infect Dis 2002;2:103-10.
14. Huskins WC. Transmission and control of infections in out-of-home child care. Pediatr Infect Dis J 2000;19:S106-10
15. Koch R. Ueber bacteriologische Forschung. In Verh. In: X. International Medical Congress. Berlin, 1890.
16. Hoque BA, Juncker T, Sack RB, Ali M, Aziz KM. Sustainability of a water, sanitation and hygiene education project in rural Bangladesh: a 5-year follow-up. Bull World Health Organ 1996;74:431-7.
17. Pinfold JV, Horan NJ. Measuring the effect of a hygiene behaviour intervention by indicators of behaviour and diarrhoeal disease. Trans R Soc Trop Med Hyg 1996;90:366-71.
18. Peterson EA, Roberts L, Toole MJ, Peterson DE. The effect of soap distribution on diarrhoea: Nyamithuthu Refugee Camp. Int J Epidemiol 1998;27:520-4.
19. Alam N, Wojtyniak B, Henry FJ, Rahaman MM. Mothers' personal and domestic hygiene and diarrhoea incidence in young children in rural Bangladesh. Int J Epidemiol 1989;18:242-7.
20. Khan MU. Interruption of shigellosis by hand washing. Trans R Soc Trop Med Hyg 1982;76:164-8.
21. Buchanan RL, Whiting RC. Risk assessment and predictive microbiology. J Food Prot 1996S;55:509-13.
22. Cabelli VJ, Dufour AP, McCabe LJ, Levin MA. Swimming-associated gastroenteritis and water quality. Am J Epidemiol 1982;115: 606-16.
23. Carabin H, Gyorkos TW, Soto JC, Joseph L, Payment P, Collet JP. Effectiveness of a training program in reducing infections in toddlers attending day care centers. Epidemiology 1999;10:219-27.
24. Rothman KJ. Causes. Am J Epidemiol 1976;104:587-92.
25. Koopman JS, Lynch JW. Individual causal models and population system models in epidemiology. Am J Public Health 1999;89:1170-4.
26. Longini IM Jr, Koopman JS, Haber M, Cotsonis GA. Statistical inference for infectious diseases. Risk-specific household and community transmission parameters. Am J Epidemiol 1988;128:845-59.
27. Haber M, Longini IM Jr, Cotsonis GA. Models for the statistical analysis of infectious disease data. Biometrics 1988;44:163-73.
28. Susser M. Causal thinking in the health sciences. Oxford: Oxford University Press; 1973.