The epidemiological method of building causal inference

Nancy Muir Ryan, RN, MSN
Instructor
School of Nursing
The Ohio State University
Columbus, Ohio

MUCH OF NURSING research seeks to demonstrate causal relationships among variables pertinent to nursing situations.[1] But when is there enough evidence to be confident that the relationships among variables are causal and not merely casual? In epidemiology, a science known for its inquiry into etiology and causality, the process by which relationships are determined to be causal is called *building inference*.[2] Epidemiological methods can be applied to all types of diseases, conditions, and health-related events.[3] Therefore, they can be utilized for nursing research.

The purpose of nursing research is not only to improve practice but also to test theories, derive new theories, and systematically advance nursing knowledge.[4,5] The trend in nursing research has been toward independent, unconnected studies on a

The author wishes to acknowledge the contributions of Virginia Boardman, RN, PhD, an associate professor at the Case Western Reserve University, for her assistance in the preparation of this article.

0161-9268/83/0052-0073$2.00

wide variety of topics, rather than a systematic accumulation of related studies.[6] Nursing studies could advance at a more rapid rate if epidemiology's causal inference building process were used in the planning and implementation of nursing research.

THE SCIENCE AND ITS METHOD

It is important to understand the science of epidemiology before discussing its method. Epidemiology, the science, is the study of the distribution of disease in human populations and the factors that influence this distribution.[7] The epidemiological method is a "method of reasoning about disease that deals with biological inferences derived from observations of disease phenomena in population groups."[8(p89)] The key word in both definitions is *population*. Epidemiologists study groups of people and search for patterns, commonalities of experience, similar characteristics, and the relationship of the characteristics to the disease of interest.

Historically, infectious disease, like cholera, was the primary focus of epidemiology. Although the term *disease* is most frequently used in epidemiological literature, it is meant to be broadly defined to include chronic disease such as cancer and coronary heart disease, psychiatric conditions like suicide and schizophrenia, and such social problems as alcoholism, drug abuse, and child abuse. The goal is to determine the causes of health-related conditions to prevent such conditions and to promote health. The phenomena of interest to nurses may include any of the previous examples, as well as more specific clinical situations and nursing interventions. A summary of the process of building causal inference can be seen in Table 1.

Table 1. Summary of causal inference building steps and examples of how they can be applied to nursing research

		Example	
Step	**Question**	**1**	**2**
Clinical observation	What is the phenomenon? Which variables are associated with the phenomenon? What is the present knowledge concerning the variables and the phenomenon?	Phenomenon: Hospital patients who fall. Variables: age, diagnosis, medication, cognitive alertness, sensory impairment, season, assistive ambulatory devices, use/nonuse of side rails, etc.	Phenomenon: Children with psychosomatic (PS) symptoms. Variables: overprotective parents, genetic predisposition, organ vulnerability, prematurity, poor parent-child relationship (PCR), early severe illness, etc.
Survey/descriptive study	What is the scope of the phenomenon? What are the significant variables associated	Survey all hospitals in a county. What demographic variables are common to the pa-	Survey all children who enter a pediatric clinic with apparent PS symptoms. What

Table 1 *(continued)*

Step	Question	Example 1	Example 2
	with the phenomenon?	tients who fall? eg, elderly, confused, frequent transfers, certain medication.	family and child variables are most common? eg, poor PCR, early life-threatening illness.
Cross-sectional study	Which significant variables differentiate persons who do and do not experience the phenomenon at a certain time?	Monitor admissions to a hospital during one month; compare elderly patients who fell with those who did not fall concerning above variables.	Compare children with and without PS symptoms and their parents concerning above variables; use tools to assess current PCR.
Retrospective study	Did the significant variables precede the phenomenon, according to data recorded in the past?	Review incident reports and records of elderly patients who fell and those who did not fall. Which of the above variables are significant only to patients who fell?	Compare children with and without PS symptoms. Did a life-threatening illness occur during infancy? What was the PCR like in the past? Is a poor PCR a cause or effect of PS symptoms?
Prospective study	Do the significant variables precede the development of the phenomenon, according to data collected over a period of time? Is it possible to predict the development of the phenomenon, using knowledge of significant variables?	Over a period of one year, monitor the admission of elderly patients who have the significant characteristics and those who do not, and record the incidence of falls. Can one predict which patients are likely to fall, despite usual precautions?	Observe a group of infants from a well-baby clinic. Record the occurrence of previously significant variables, assess PCR at 6 months, and observe for development of PS symptoms. Can one predict which infants will develop PS symptoms later in childhood?
Experimental study	If some persons are exposed to the significant variables and others are not, will the exposed persons develop or experience the phenomenon? Which nursing interventions will prevent or promote the phenomenon?	One cannot ethically inflict confusion, certain medications, etc, on elderly patients, but one can test nursing interventions. Does use of side rails prevent or increase the likelihood of falls in the elderly?	Using knowledge of predictive variables, test various early preventive counseling methods on families of children likely to develop psychosomatic symptoms.

Selected epidemiological concepts

For clear communication of ideas, it is necessary to define the major concepts related to causal inference. A cause is an "act, event, or state of nature which initiates or permits a sequence of events leading to an effect."[9(p588)] A cause can be narrowly defined as sufficient, meaning it will inevitably lead to the effect, necessary, meaning the effect will not occur without it, or component, meaning that a single event combines with others to make up a sufficient cause.[9] It is generally agreed that biological, sociological, and psychological phenomena, and therefore nursing phenomena, have multiple causative factors, or component causes.

Causality is a condition interpreted in epidemiology by the following statement: If X is a cause of Y, then X precedes and produces a change in Y. Association means that a change in X is merely associated with a change in Y; it may or may not be causal.[10] Inferences are educated guesses or hypotheses about the relationship between a variable and the phenomenon of interest.

Building causal inference

Building strong causal inference is an inductive reasoning process that begins with clinical observations of a particular health-related phenomenon. It progresses through two general stages that focus first on groups, then on individuals.[7(pp13-15)]

Stage 1 involves the demonstration of a statistical association between a suspected causal characteristic and the phenomenon. Stage 1 studies include surveys and descriptive studies of population groups. The rates of occurrence of the phenomenon and characteristics are calculated, the characteristics of various groups are explored in relation to the phenomenon, and the most significant characteristics are identified.

Stage 2 involves the derivation of biological inferences from a pattern of statistical associations found in stage 1 studies. The focus is on individuals with and without the significant characteristics. Stage 2 studies include cross-sectional studies, retrospective studies, prospective studies, and experiments.

It is rarely possible to prove causality, but epidemiologists can build a strong case for causality through a series of stage 1 and 2 studies. Alternative inferences or hypotheses are refuted through successive, increasingly more rigidly controlled studies until only one hypothesis remains that is supported. Stage 1 and 2 studies each contribute some support to the validity of a causal inference, but the positive findings of an isolated study are less meaningful than the cumulative findings of many studies. Each type of study has important advantages, disadvantages, and threats to reliability and validity that have been adequately described in introductory research textbooks. Although the importance of the strengths and weaknesses of individual studies is not discounted, the underlying premise in building causal inference is that if similar results are found in various types

If similar results are found in various types of studies despite their weaknesses, then the observed relationships must be more than coincidental.

of studies despite their weaknesses, then the observed relationships must be more than coincidental.

Causal criteria

Five criteria are widely used to evaluate the likelihood that an association between variables is causal, and not spurious or indirect.[3]

1. Strength of association: When comparing the ratio of disease rates for persons with and without the hypothesized characteristic, the larger the ratio, the greater the strength. Also, strength increases if a dose-response effect can be demonstrated. That is, with increasing levels of exposure to the characteristic, a corresponding rise in occurrence of the disease or phenomenon is found.
2. Consistency of association: If an association noted in one study appears in other studies with different designs and different populations, the strength of association increases.
3. Temporally correct association: Exposure to the variable or characteristic must predate the onset of the disease or phenomenon and allow for the necessary latency period. This criterion is difficult to establish in chronic illnesses, which often have a long latency period between exposure to the factor and manifestation of symptoms.
4. Specificity of the association: Strength increases if the occurrence of one variable can be used to predict the occurrence of another. This criterion is not crucial, since it is known that one variable can result in more than one disease, and many different variables can result in the same disease.
5. Coherence with existing information: If the association noted is biologically plausible or is consistent with current knowledge, then it is more likely to be accepted. Depending on the state of scientific knowledge at the time, this criterion may not be a crucial one.

PROCESS OF BUILDING CAUSAL INFERENCE

Clinical observations

Like a pyramid, the process of building causal inference must rely on a strong base (see Fig 1). Clinical observations and the need for scientifically based nursing actions provide the focus and impetus for systematically planned and implemented research. The following steps are involved in beginning the process of building causal inference in the clinical area:

- Observe the phenomenon in a clinical situation and explicitly describe it.
- Identify all variables that seem to be related to the phenomenon through clinical observation and a review of the available literature.
- Evaluate present knowledge about the phenomenon and identified variables. Is there theoretical support for the relationship between variables noted in the clinical observation? Is there empirical support for the clinically observed phenomenon? What studies have already been conducted? Evaluate completed studies and build on their results. If little has been done in an organized way, begin with a survey or descriptive study.

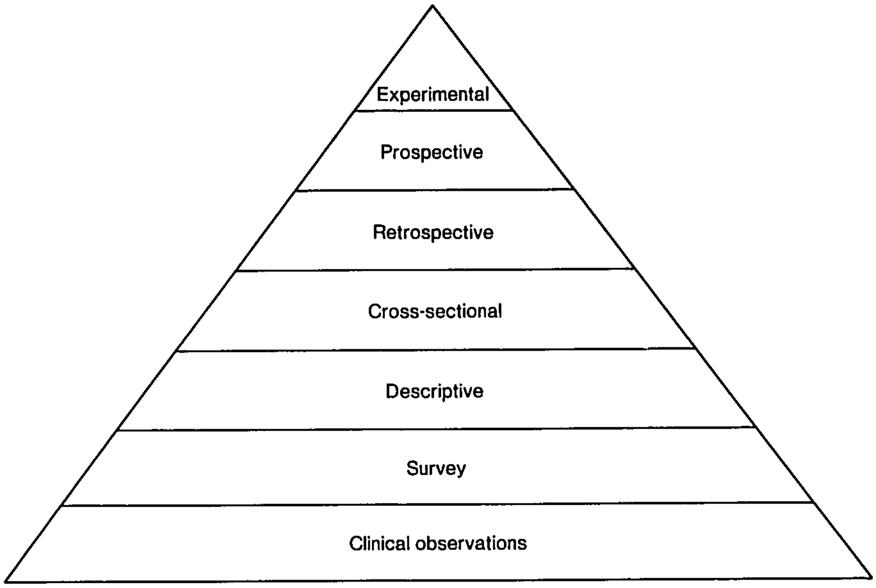

Fig 1. Causal inference conceptualized as a pyramid.

Survey and descriptive study

A survey or descriptive study reveals patterns of disease or phenomenon occurrence in selected geographic areas or populations according to demographic variables related to person, place, and time.[3,7,11] Because a survey samples a much larger geographic area or population than a descriptive study, it may be advisable to conduct a survey first, especially if the phenomenon of interest is relatively rare, to avoid bias associated with small selected samples. Data for a survey and other types of studies can be obtained directly with questionnaries, interviews, home visits, and/or physical examinations, or indirectly through legal documents such as birth or death certificates, census records, or hospital charts.

Results of surveys and descriptive studies can be summarized as frequencies, rates, and proportions, in maps, diagrams, or graphs, and as statistical associations between each demographic variable and the phenomenon of interest. If a statistical association is found, tests are then made to determine if the association is significant, ie, not likely due to chance. The association must not be interpreted as causal in this early stage of the causal inference building process, but to increase confidence in their significance, the associations must be further tested in surveys of other geographic areas or populations and in other types of studies.

Cross-sectional study

In a cross-sectional study, which further narrows the scope of the phenomenon, one obtains a large group of subjects and determines which subjects have and have not experienced the phenomenon during a particular period of time. The two groups are then compared according to the presence of the significant variables obtained in previous surveys and/or descriptive studies. If the frequency of the phenomenon is higher in subjects with a particular variable, then an association is inferred. If the frequency of a particular variable is higher in persons experiencing the phenomenon, then an association is also inferred.

Associations resulting from cross-sectional studies, though significant, are not necessarily causal for several reasons, depending on the phenomenon of interest.[7,12] Existing cases may not be representative of all cases of the phenomenon because cases of long duration are overrepresented. Cases of short duration may not survive long enough to be included at all, and their characteristics may differ markedly from the long duration cases. Certain types of cases may leave the community (as in migration) or simply not live in the community sampled. In addition, one cannot be sure if the variable occurred before or after the phenomenon, especially if the variable is a physiological or biochemical one. Similar significant results obtained in several cross-sectional studies strengthen the confidence in observed associations and narrow the scope for future studies, particularly the retrospective study.

Retrospective study

One criterion for inferring causality between a variable and a phenomenon is that the variable precedes the occurrence of a phenomenon.[3] A retrospective study is useful to determine if variables that were significant in surveys, descriptive studies, and cross-sectional studies are still significant using data about the variables that have been recorded in the past.[3,7]

Subjects are identified and accepted into the study according to whether or not they have recently experienced the phenomenon. "Cases" have experienced the phenomenon, whereas "controls" closely resemble cases except for the phenomenon. Previously recorded data are reviewed to determine if and when a particular variable existed prior to the phenomenon. Relative risk rates (RR) can be calculated to indirectly measure the degree of association between the variable and the phenomenon when frequencies are recorded in the following manner (see Table 2):

$$RR = \frac{a/a+c}{b/b+d}.$$

If $a/a+c$ is significantly greater than $b/b+d$, then a significant association exists between the variable and the phenomenon, and the inference regarding a causal association becomes stronger. However, retrospective studies have many weaknesses that limit confidence in the apparent causal association. Significant associations should then be tested in a prospective study, which is the next step in the causal inference building process.

Table 2. Relative risk matrix

	Phenomenon	
Variable	Cases	Controls
Present	a	b
Not present	c	d

Prospective study

A prospective study provides the most direct measurement of the risk of experiencing a phenomenon, the incidence rate (IR).[7,12] IR is the number of subjects developing the phenomenon divided by the population at risk during a specific period of time. The prospective study begins with groups of subjects with and without the inferred causal variable. The subjects are followed directly or indirectly for a preset length of time and are observed periodically for development of the phenomenon. If the incidence rate is significantly greater for subjects with the variable than without, one can be reasonably confident that the relationship is causal. The variable can be used as a predictor for the phenomenon, a valuable tool if prevention of the phenomenon is the ultimate goal.

If the causal inference building process is employed and a particular variable continues to be significantly associated with a phenomenon, then all five causal criteria previously mentioned will likely be met and the relationship can be described as causal. In other words, the relationship has strength, consistency, temporal correctness, specificity, and coherence. To be certain that intervening variables have not confounded the true relationship, a further step in the inference building process may be employed: the experiment.

If a particular variable continues to be significantly associated with a phenomenon and all five causal criteria are met, the relationship can be described as causal.

Experimental study

An experimental study involves the analysis of the impact on phenomenon development of exposing and not exposing some subjects to the suspected causal variable.[3] The strength of the experimental method depends upon the investigator's ability to control potentially confounding variables. The experiment is a valuable addition to the inference building process if a significant number of exposed subjects experience or develop the phenomenon.[7] However, an experiment requires that the variable be under the investigator's control, which, with human subjects, may not be possible. If it is possible, it may not be ethical.[13] Experimental tests of causal inferences are desirable but not crucial to support the statement that a variable is causally related to a particular phenomenon. The experimental stage of research is useful, though, to test nursing interventions related to the phenomenon, based on knowledge of its component causes.

• • •

Like a pyramid (see Fig 1), each step in building causal inference rests on the findings of the previous steps. One begins with several potentially causal variables, but for efficiency, the investigator must narrow his/her research to just a few variables. Significant findings for these variables do not imply that they are the only or the most important causal variables; they merely are component causes of a complex phenomenon. The same clinical observation, survey, and descriptive study can provide the base for numerous pyramids.

Nursing actions based on findings of a logically developed series of studies are more scientifically sound than nursing actions based on the findings of an independent, isolated study. Nurse researchers need to collaborate, communicate both positive and negative findings, and build on those findings to advance nursing science in the most efficient and useful manner.

REFERENCES

1. Field M: Causal inference in behavioral research. *Adv Nurs Sci* 1979; 2(1):81-93.
2. Susser M: Judgement and causal inference criteria in epidemiologic studies. *Am J Epidemiol* 1977; 105(Jan):1-15.
3. Mausner JS, Bahn AK: *Epidemiology.* Philadelphia, WB Saunders Co, 1974.
4. Gortner SR. Research for a practice profession. *Nurs Res* 1975; 24(3):193-197.
5. Fawcett J: The relationship between theory and research: a double helix. *Adv Nurs Sci* 1978; 1(1):49-62.
6. Ellis R: Fallibilities, fragments, and frames: contemplation on 25 years of research in medical-surgical nursing. *Nurs Res* 1977; 26(3):177-182.
7. Lilienfeld AM: *Foundations of Epidemiology.* New York, Oxford University Press, 1976.
8. Lilienfeld DE: Definitions of epidemiology. *Am J Epidemiol* 1978; 107(2):87-90.
9. Rothman KJ: Causes. *Am J Epidemiol* 1976; 104(6):587-592.
10. Blalock HM, Jr: *Causal Inferences in Non-Experimental Research.* Chapel Hill, University of North Carolina Press, 1964.
11. Barnard K: Research designs: descriptive method. *Am J Maternal Child Nurs* 1981; 6:243.
12. Friedman GD: *Primer of Epidemiology.* New York, McGraw-Hill, 1980.
13. Polit DF, Hungler BP: *Nursing Research: Principles and Methods.* Philadelphia, JB Lippincott, 1978.