THE EPIDEMIOLOGICAL STUDY OF SUICIDAL BEHAVIOR

David Lester

The Richard Stockton College of New Jersey
1997

1: INTRODUCTION[1]

Epidemiology is concerned with the study of the occurrence of diseases in human populations, that is, in groups rather than separate individuals. Epidemiologists look for disease patterns in populations -- communities, regions or nations.

Epidemiologists distinguish between the host (the person who has the disease), the agent (the cause of the disease) and the environment (such as social and climatic conditions). For suicidal behavior, the host is the suicidal person, and the environment constitutes the social milieu in which the suicidal people find themselves. The agent, however, is not quite so simple as in some medical diseases where there are bacteria or viruses known to cause the disease. For suicidal behavior, the "causes" are not well understood.

However, Friedman (1987) noted that all persons, for example, infected with the beta-hemolytic streptococcus do not develop the disease it causes, rheumatic fever. Some hosts are more susceptible to the disease than others. This finds an analogy in suicidal behavior where it is known that factors such as a major depressive disorder or experience of suicidal behavior in family members and peers may increase the probability of suicidal behavior in some individuals. It may be possible to view such factors as agents to which not all hosts are susceptible.

Diseases are not easy to define. Some disease names merely describe the appearance of a person (externally or internally, such as colitis) or a subjective sensation (such as headache). However, some disease names do imply a causal element (such as pneumococcal pneumonia). Typically, with increasing knowledge, disease names move from description to causation. As regards suicidal behavior, suicidologists are still working at the descriptive phase, with a great deal of effort currently being put into devising a set of descriptive terms which all suicidologists can accept and work with (e.g., Canetto and Lester, 1995; O'Carroll, et al., 1995).

2: MEASUREMENT IN EPIDEMIOLOGY

Epidemiology is a quantitative science, and there are several standard terms employed.

Prevalence

[1] This discussion of the general principles of epidemiology draws heavily from the textbook on epidemiology by Friedman (1987).

Prevalence refers to how many people in the group have the disease at a particular point in time:

$$\text{prevalence rate} = \frac{\text{number of persons with the disease}}{\text{total number in group}}$$

The point in time can refer, for example, to a particular day or to an event which happens to different people at different times, such as during a physical or psychiatric examination or during the first bleeding day of the menstrual cycle.

Prevalence rates for fatal suicide are rarely used, though occasionally an author tells how many people killed themselves in one day in the nation or in the world. A screening of people in a community for current suicidal ideation would give a point-of-time prevalence for suicidal ideation. Asking consecutive psychiatric patients admitted to an inpatient clinic about their current suicidal ideation would also provide a prevalence rate, as would a count of the number of women who engage in nonfatal suicidal behavior on the first bleeding day of the menstrual cycle.

Period prevalence rates measure the number of people who develop the disease during a period of time, such a month or a year. A yearly prevalence rate would include all of those persons in the group who had the disease at the beginning of the year plus all of those who developed the disease during the year.

Incidence

Incidence rates describe the rate of development of a disease in a group over a time period.

$$\text{incidence rate} = \frac{\text{number of persons developing the disease}}{\text{total number at risk}} \text{ per unit of time}$$

This is the most common measure used for fatal suicide behavior, for example, the suicide rate of the United States in 1990 was 12.4 per 100,000 per year.

Mortality rates are simply incidence rates of death from particular causes. *Age-specific rates* are incidence rates for those in a particular age range, such as the mortality rate from suicide for those aged 15-25 years of age. *Case fatality rates* are the proportion of people with a disease who die in a given period:

$$\text{case fatality rate} = \frac{\text{number of people dying from a disease}}{\text{total number with the disease}} \text{ per unit of time}$$

Strictly speaking, case fatality rates cannot be calculated for suicidal behavior. Suicidal behavior is not a continuous disease which ends for some in death. We do calculate ratios such

as the proportion of those who have engaged in nonfatal suicidal behavior who eventually die from suicide or the proportion of those diagnosed as schizophrenics who eventually die from suicide, but these are really incidence rates per unit of time for specific groups of individuals defined by their behavior or their psychiatric diagnosis.

Attributable Risk

If we calculate the incidence of a disease in one group and the incidence in a second group, then the *attributable risk* is the difference between the two incidences. For example, the fatal suicide rate for Native Americans in 1980 was 13.3 per 100,000 per year and the fatal suicide rate for white Americans was 13.2 per 100,000 per year (Lester, 1994); the attributable risk of suicide attributable to being Native American was 0.1 per 100,000 per year (that is, 13.3 minus 13.2).

We might want to adjust for the age differences in the two populations in calculating these incidences and for any other variables which appear to be important.

The *attributable fraction* or *etiological fraction* is defined as follows:

attributable fraction =

$$\frac{\text{incidence in group A - incidence in group B}}{\text{incidence in group A}}$$

For example, in the example above, the attributable risk for suicide in Native Americans would be (13.3-13.2)/13.3 = = 0.0075 or 0.75%.

Relative Risk

Two rates may be compared by simply examining the ratio of the two rates -- in the example above 13.3/13.2 = 1.0075. Native Americans had 1.0075 times the risk of dying from suicide as white Americans. This ratio is called the *relative risk*. If the two incidence rates are instantaneous incidence rates, then the ratio is called the *rate ratio*; if they are cumulative incidence rates over a specific period of time, the ratio is called the *risk ratio*. For fatal suicide, the risk ratio would be termed a *mortality risk ratio*.

3: THE OBSERVATIONS MADE IN EPIDEMIOLOGY

Observations or measurements should be *reliable*, that is, whatever is being measured, the measurements should be consistent. Reliability has several operational measures. Test-retest reliability means that measuring the variable on a subsequent occasion will provide the same score as the previous measurement. Parallel forms reliability means that measuring the variable with two different measuring instruments on one occasion will provide the same score.[2] For example, research on the assignment of psychiatric diagnoses to patients shows that a great deal of disagreement exists between mental health professionals who examine the same patients, and

[2] There are other measures of reliability which can applied to specific situations, such as when using a multi-item inventory.

so psychiatric diagnosis may be quite unreliable. In a study of the reliability of diagnoses using DSM-III, Strober, et al. (1981) found that clinicians interviewing the same 95 adolescent psychiatric inpatients could agree on the major diagnosis (that is, using thirteen major categories) only 77 percent of the time. There are several sources for this unreliability. For example, different interviewers will ask different questions of the patient, and the patient may feel differently about the interviewers and respond differently to the same question. However, even when professionals observe a video-tape of an interview (which eliminates these two sources of possible disagreement), they still come to different diagnoses.

Validity means that the measuring instrument measures what the researchers think it is measuring. Intelligence tests are highly reliable tests by psychological standards; however, scholars debate with great emotion the validity of the tests, that is, they disagree over what intelligence tests measure. In suicide research, the distinction between self-mutilation and nonfatal suicidal behavior may be hard to draw (Lester, 1972). For example, it may be difficult to decide whether a chronic wrist-slasher is a self-mutilator or making "suicidal gestures."

Not all variations and fluctuations in measurements are due to unreliable or invalid instruments. Variations in the current level of suicidal ideation of populations may appear because of both differences among subgroups (such as men and women, the young and the elderly, or whites and blacks) and differences within each individual (suicidal ideation in a person may vary from day to day, and even from hour to hour). In biochemical assays of the bodily fluids of suicidal individuals (such as the cerebrospinal fluid) there be measurement error, such as inaccuracies in the biochemical measuring devices.

Another set of errors comes from sampling. The use of inferential statistics requires that we use random samples. However, the difficulty of obtaining truly random samples (that is, everyone in the population of interest must have an equal probability of being selected for the study) means that researchers typically use "available" samples, such as the patients in a particular clinic or the students in a particular class. These samples may not be representative of the population as a whole. Even random samples may not be representative of the population of interest, for *sampling variation* may result in deviant samples purely by chance. Increasing the sample size reduces the risk of sampling variation.

Under-reporting of the disease or behavior in question may occur for several reasons. In estimating the incidence of nonfatal suicidal behavior, people may decline to report self-involvement for fear of ridicule or stigma. In the reporting of fatal suicidal behavior, professionals may decline to certify the death accurately for a variety of reasons such as sparing the family stigma or a fear of legal problems if the family disagrees with the assigned cause. In some nations, such as Great Britain, many suicidal deaths are classified as undetermined whereas in other nations they would be classified as suicides. The recent increase in suicide rates in Roman Catholic countries such as Ireland is probably due in part to more valid classification of deaths. In a study in the United States, Farberow, et al. (1977) found that counties where lawyers, acting as coroners, certified deaths had lower suicide rates than counties where pathologists, acting as medical examiners, certified death.

Much research in suicidology uses the responses of patients to self-report questionnaires such as depression inventories. Several sources of error and bias can occur using such tests. Typically, a number of people refuse to participate in the study, while others fail to respond to particular questions (sometimes to complete sections of questions), and the omission of these patients from the sample introduces bias since these nonresponders may differ in critical ways from those who respond (Lester, 1969), as well as reducing the sample size. Research reports should always indicate the percentage of nonresponders in the study.

Other responders sometimes answer inconsistently, have response sets (for example, they tend to agree with every item), and lie (responders can fake "bad" or "good"), and a few psychological tests have special scales built in to detect this. However, most psychological inventories do not have such scales.

The problems introduced by these issues of reliability and validity of the data can sometimes be overcome by the use of large samples, when the unreliability of the data appears as "noise" in the data, increasing the standard deviations of the measures or reducing the level of statistical significance of the results, but not destroying completely the relationships sought.

4: BASIC METHODS OF STUDY

The two major approaches in epidemiological research are observational studies and experimental studies. In *observational studies*, the researcher merely observes the phenomena as they occur. In *experimental studies*, the researcher actively intervenes to change one of the variables in the research. For example, a description of the rates of suicide in a nation, the methods used and the variation over time is an observational study. A comparison of suicide notes written by fatal and nonfatal suicidal people is an observational study. However, a study in which half of a group of suicidal clients is assigned at random to cognitive therapy, while the other half is assigned to simple crisis intervention would be an experimental study since the researcher is determining which client gets each of the two forms of treatment. Experimental studies have the advantage of providing evidence for *cause-and-effect*, that is, whether in our example, the different therapeutic interventions *caused* the difference in the response of the two groups of suicidal clients. Observational studies are correlational in nature, and correlational studies do not provide strong evidence for cause-and-effect relationships.

Because experimental studies are difficult to carry out, especially in suicidology, researchers use two tactics to improve observational studies. First, *natural experiments* sometimes occur. For example, a nation may detoxify natural gas slowly over a period of years (as it switches from very toxic coal gas to less toxic natural gas), and the effects of this change on the suicide rate can be monitored. Second, *statistical controls* for other possible causal factors can be incorporated into the research design and data analysis, permitting a multivariate analysis of the data (that is, using many variables), and this can strengthen our certainty that a particular variable causes suicidal behavior.

A STUDY USING STATISTICAL CONTROLS

Zimmerman (1990) examined the association between the state spending levels of hospitals and the suicide rates in the states of America in several years. Looking at the Pearson correlations for 1960 and 1970 between suicide rates and several social variables, she found the following:

	1960	1970
suicide rates and		
hospital spending	-0.23	-0.45*
divorce rates	0.74*	0.69*
population change	0.50*	0.56*
population density	-0.55*	-0.61*
income	0.30*	-0.02
% blacks	-0.26*	-0.32*

An "*" indicates a statistically significant correlation coefficient. Zimmerman then placed all of these variables in a multiple regression, so that their combined effects could be examined. The standardized regression coefficients were as follows:

	1960	1970
suicide rates and		
hospital spending	0.07	-0.15
divorce rates	0.56*	0.21
population change	0.07	0.36*
population density	-0.36*	-0.38*
income	0.17	0.02
% of blacks	-0.07	-0.17

In the regression analyses, only divorce rates and population density contributed significantly to the prediction of suicide rates in 1960, and only population change and population density contributed significantly to the prediction of suicide rates in 1970. Thus, for these two years, the regression analysis, which controls for the other social variables, failed to find any impact from hospital spending, whereas the simple correlation coefficients (which do not control for the other social variables) suggested a negative association.

Zimmerman's study also illustrates the importance of replicating results on other samples or in other years to see whether the results are generalizable. In fact, Zimmerman replicated her study using 1980 and 1984 also, obtaining slightly different results for those years.

Observational Studies

Observational studies can be *descriptive* or *analytic*. Descriptive studies simply describe the phenomenon whereas analytic studies seek to explain the phenomenon.

Descriptive Studies

Descriptive studies describe the patterns of disease in the population. They study the association of the disease by such variables as age, gender, marital status, race, occupation, social class, geographic location and time. This information identifies groups at high risk for a disease, assists the planning of services to respond to those with the disease, and provides clues to the etiology of the disease which may stimulate future analytical studies.

A DESCRIPTIVE STUDY

Lester and Wilson (1988) obtained raw data on the individual cases of fatal suicides in Zimbabwe for the period 1983-1986. They calculated suicide rates of 6.9 per 100,000 per year for Africans, 17.6 for Europeans and 9.7 for Asians/Coloreds. For the Africans, the suicide rate for African men was 10.5 and for women 3.4. By age, the suicide rate for both men and women peaked for those aged 60-69, at 33.4 and 10.2 respectively. The most common method for Africans committing suicide was hanging (76% of the men and 55% of the women chose this method), followed by poison (chosen by 16% of the men and 35% of the women).

Age variations in the prevalence or incidence of a disease can be presented in two ways. A current or cross-sectional presentation shows the suicide rate in each age group in one year -- different people are involved in each age group. A cohort presentation shows the suicide rate of a cohort over time as it ages. These two presentations can give quite different results. For example, the suicide rate of the Canadian male cohort born in 1911-1915 was

0.7 per 100,000 per year when they were aged 10-14,
3.2 when they were aged 15-19,
7.6 when they were aged 20-24,
12.0 when they were aged 25-29,
8.9 when they were aged 30-34,
14.9 when they were aged 35-39,
19.2 when they were aged 40-44,
22.6 when they were aged 45-49,
27.9 when they were aged 50-54,
32.5 when they were aged 55-59,
29.8 when they were aged 60-64 and
26.0 when they were aged 65-69 in 1976-1980.

Thus the suicide rate for this cohort peaked when they were 55-59 years old (Lester, 1988a). In contrast, the suicide rate by age in Canada in 1980 was:

10-14	1.5
15-19	18.9
20-24	29.1
25-29	30.8
30-34	25.0
35-39	24.2

40-44 23.5
45-49 30.7
50-54 29.2
55-59 28.1
60-64 27.9
65-69 28.2

with a peak for those aged 25-29.

Variation by place is of interest because it may provide clues to etiology. In the case of suicide, rates have typically been very high in Hungary. Hungarians have been quite concerned with their high national suicide rates, and they see themselves as a highly depressed people in general. Several explanations are possible, including physiological differences between Hungarians and other national groups, differences in child rearing practices, or differences in social expectations (that is, Hungarians are aware of their high suicide rate and, therefore, suicide becomes more of an option to Hungarians when they are in crisis).

Robinson (1950) warned of the dangers of assuming that associations between variables over geographic regions can be generalized to individuals. For example, if suicide rates and church attendance are found to be correlated over the states of America, we cannot assume without further research that these two variables are associated over individuals. Robinson called this inappropriate generalization the *ecological fallacy*.

Variations over time can be *short-term* or *long-term*. Among the short-term effects, *epidemics* (or *outbreaks*) are of special interest. An epidemic is an occurrence of the behavior in a population in excess of the number of cases expected. The disease in an epidemic may affect only those who are susceptible. Others may be immune or resistant as a result of inherent factors. After a person is exposed to the disease, there is an incubation period, and, once the person has the disease, he or she may be enter a communicable period during which they can pass on the disease to others. An epidemic typically shows an increasing incidence over time to a maximum, followed by a steady diminution until it disappears almost completely as the supply of susceptible individuals is exhausted. For suicide, the notion of resistance and incubation period may have some relevance.

For example, Taiminen, et al. (1992) reported on eight inpatient suicides in a three month period in their clinic in Finland. Six of the patients had close relationships with one another, and Taiminen was able to document the influence of suggestion and identification on the occurrence of and the methods chosen by these suicides.

There may also be *recurrent* or *periodic* time trends -- suicide shows variations by time of day, day of week and month of the year. For example, Phillips and Wills (1987) examined fatal suicides in the United States from 1973 to 1979, and found that suicide rates were above average on New Year's Day, July 4th and Labor Day and for the five following days, but below average on the five days prior to the national holidays. In contrast, suicide rates were below average before, during and after Memorial Day, Thanksgiving and Christmas.

Long-term trends are also called *secular* trends and extend over years or decades. For example, the Hungarian suicide rate rose steadily for twenty years, from 1965 to 1984, after which it has steadily declined. Araki and Murata (1987) reported on secular trends for suicide in Japan for the 33 years following World War Two (1950 to 1982). The suicide rate peaked for both men and women in the mid-1950s, dropped to lows in late 1960s, and diverged in the 1970s, increasing steadily for men and increasing and then decreasing for women. Looking at these changes, Araki and Murata suggested that suicide rates decreased during times of economic prosperity and increased in the years prior to economic depressions. However, they did not test this hypothesis, leaving later investigators to do so.[3]

Analytic Studies

Analytic studies, on the other hand, start with an hypothesis about the causes of suicide, and the data test this hypothesis. For example, Clarke and Lester (1989) hypothesized that the availability of methods for suicide would affect the suicide rate. They found that, as domestic gas was detoxified in England, the use of domestic gas for suicide declined dramatically, and the overall English suicide rate dropped by almost a third. Although this study was observational (the researchers themselves did not manipulate any of the variables), the results supported a particular hypothesis about the causes of suicidal behavior.

AN ANALYTIC STUDY

Lester (1993) examined the changes in the suicide rates of the 48 contiguous continental states of America between 1970 and 1980 and correlated these changes (both absolute and percentage) with the absolute number of suicide prevention centers in the states, the number per capita and the number per unit area in 1970. The correlations were:

	centers:		
	absolute number	per unit area	per capita
absolute change	-0.40*	-0.11	-0.22
percentage change	-0.30*	-0.12	-0.25*

The correlations marked with an "*" were statistically significant. All were negative, indicating a preventive effect on suicide from suicide prevention centers, but those for the absolute number of centers were larger than the others and statistically significant. Thus, Lester's study provided some support for the preventive effect of suicide prevention centers on fatal suicide.

Cross-Sectional Studies

[3] Lester, et al. (1992) found that unemployment rates were associated with the Japanese suicide rate from 1953 to 1982 even with controls for growth in the gross national product per capita, the divorce rate and female labor force participation.

Cross-sectional or prevalence studies examine the relationship between suicide and other variables of interest in a defined population at one point in time. For example, Lester (1994) compared the rates and methods of suicide in Chinese Americans, Japanese Americans and Filipino Americans in the United States in 1980. A cross-sectional study can examine the focal behavior (in this case suicide) in the different subgroups of the population or examine the presence or absence of a variable in those who engage in the behavior (in this case suicide) and those who do not.

A CROSS-SECTIONAL STUDY

Goldney, et al. (1955) studied 3130 school students from twelve randomly chosen metropolitan schools in Adelaide, Australia, in 1980. The students had a mean age of 15.6 years. These students were followed up by mail each year, and by 1988 472 remained in the study, with data available for 432 subjects.

Each subject was asked about their employment status and whether they had had thoughts of suicide at some point in the life. The results were as follows:

employed/dissatisfied	n=31	48% lifetime suicidal ideation
unemployed	n=21	38% lifetime suicidal ideation
students	n=35	37% lifetime suicidal ideation
employed/satisfied	n=345	20% lifetime suicidal ideation

Previous research had suggested an association between suicidal behavior and unemployment; Goldney's study suggests an association also between suicidal behavior and unhappy employment.

(Goldney labeled his study as a prospective longitudinal study -- it was not. Although this study was collected on a cohort of subjects, the data presented in Goldney's paper is a cross-sectional study on this cohort using data collected in 1988 -- it is not a cohort study.)

In cross-sectional studies there are several important methodological issues in choosing the sample. Statistical tests require that the samples be *random*, that is, everyone in the population should have an equal probability of being chosen. This is rarely done. A *stratified random sample* is one in which the population is classified on the basis of some social variables, such as age and gender, and the numbers of each age-by-gender group desired determined by the researcher. Then, within each age-by-gender group, the subjects chosen are chosen randomly. Sometimes the population forms natural clusters, such as classes in a school or cities in a region. Again, within each cluster, the subjects are chosen randomly. A *systematic sample* lists the population and chooses every tenth or one hundredth person on the list. However, in actuality, random samples are rarely taken -- rather the researcher uses the subjects available, either in a college class or a psychiatric clinic. Despite this, it must be borne in mind that the statistical analyses used by researchers require truly random samples.

A PREVALENCE STUDY

It was difficult to find a good prevalence study of suicidal ideation because most studies investigate "lifetime" experience of suicidal ideation. However, Vega, et al. (1993) reported a retrospective six-month prevalence study on a sample of 6,760 7th and 8th grade male school children in Miami. Among the questions, the students were asked about the presence of suicidal ideation in the past six months. Of the 5303 respondents, 18.2% reported such suicidal ideation. By ethnic group, the percentages were as follows:

Cuban Americans	17.3%
Other Hispanics	17.8%
Nicaraguans	16.5%
African Americans	20.5%
white non-Hispanics	19.3%
Haitians	19.9%
Caribbean blacks	16.0%

Case-Control Studies

Case-control studies are cross-sectional studies in which the variables of interest are examined in a sample of people with the focal behavior and those who do not have the behavior, such as suicidal patients versus nonsuicidal patients. The nonsuicidal patients comprise the *control* group or *comparison* group. The better studies match the members of the control group with the members of the focal group on important variables -- for example, they may all be psychiatric patients hospitalized with schizophrenia. The matching can be done on a group basis (that is, the two groups resemble each other as a whole) or case by case (that is, each member of the focal group is matched for certain variables with a member of the control group).

The first step in a case-control study is to identify the cases to be studied. The criteria for diagnosis and for inclusion in the study must be clearly spelled out and reported in the published report. Cases are sometimes identified by a community-wide search, but more often limited to clients seen at one or more clinics or hospitals in a limited period of time.

The control group sometimes involves important choices. The control group can be matched with the group of interest on some characteristics, and, as mentioned above, this matching can be done so that the groups are matched overall or paired so that one or more controls are chosen for each case. It is often helpful to have a very large control group, up to five or six times as many subjects, for this reduces the variability in the data and increases the chances of identifying statistically significant results.

Examples of poor choices for control groups are easy to find. For example, Neuringer (1964) studied the rigidity in the thinking of nonfatal suicides and compared them with psychosomatic patients. However, he did not demonstrate that the psychosomatic patients were as psychiatrically disturbed as the nonfatal suicides. It is common to use patients receiving minor

medical treatment or who are normal as controls in such studies (e.g., Wilson, et al., 1995), and this comparison group is of even less value. For example, a great deal of research on suicidal adolescents compares high school students who have engaged in nonfatal suicidal behavior with other students in the school (e.g., Bjarnason and Thorlindsson, 1994). Since the suicidal students are in all probability more psychologically disturbed or distressed than the average students, any differences identified could be a result of the distress rather than the suicidality of the students.

To take another example, Lester (1988b) criticized the use of simulated suicide notes, written by nonsuicidal individuals, as a comparison group for genuine suicide notes. He argued that such studies do not inform us about suicide; rather they inform us about how well people can fake a suicide note and provide clues as to popular conceptions about suicide.

Because of bias (conscious and unconscious) on the part of the researchers, it is best if those collecting and recording the data do not know whether the subjects are cases or controls.

In case-control studies, an *odds ratio* (or relative odds) can be calculated. This is defined as follows:

$$\text{odds ratio} = \frac{\text{(number of cases with behavior) x (number of controls without behavior)}}{\text{(number of cases without behavior) x (number of controls with behavior)}}$$

If the cases and controls are not matched, the odds ratio is calculated as follows:

	Cases	Controls
behavior present	a	c
behavior absent	b	d

(a, b, c, and d are numbers of subjects)

odds ratio = ad/bc

An example here comes from Norman, et al. (1990) who studied the relationship between response to the dexamethasone suppression test by depressed psychiatric patients and subsequent fatal suicide:

	Suppressors	Nonsuppressors
fatal suicides	6	7
not fatal suicides	43	10

The odds ratio for these data is (6x10)/(43x7) = 60/301 = 0.20, indicating that suppressors are less likely to subsequently kill themselves than are nonsuppressors.

If the cases are matched with the controls on an individual basis, the odds ratio is calculated as follows:

behavior in the controls

		present	absent
	present	a	c
behavior in the cases			
	absent	b	d

(a, b, c, and d are numbers of pairs)
odds ratio = c/b

If case-control studies focus on new cases that develop during the data collection phase rather than cases with the existing disease, then the research gives better clues to the *development* of the disease, and the odds ratio comes closer to estimating the rate ratio.

A CASE-CONTROL STUDY

Séguin, et al. (1995) identified all deaths of men aged 18 to 35 years in Montreal and Quebec City from suicide or car accidents. Families were contacted and asked to participate in a study. Eventually, 30 families of suicides and 30 families of car accident victims were recruited for the study. The groups were not matched, but simple demographic characteristics were reported for the two groups, and they were quite similar.

The parents of the suicides were significantly more depressed six months after their son's death than were the parents of the accident victims (mean scores 17.3 and 8.9) but not nine months after their son's death (mean scores 5.6 and 4.8). The impact on the family was much more often negative for the families of the suicides than for the families of the accident victims (47% versus 7%).

Cohort Studies

Cohort (incidence) studies explore the development of a focal behavior. A population free of the disease is identified, the so-called cohort, and followed for a period of time. Those with some attribute are compared to those without the attribute for the later development of the disease. Cohort studies are the best approach for answering the question of whether some attribute *predisposes* people to a disease. The distinction between case-control studies and cohort studies is lessened if the cross-sectional study includes questions about the history of the responders.[4]

A COHORT STUDY

Lester (1991a) examined the childhood experiences of 1528 gifted children identified in 1921 by Lewis Terman in California and their subsequent suicidal behavior. These children had

[4] Cost-saving approaches include the *nested case control study* in which only a subset of the disease-free patients are used as controls, and the *case-cohort study* in which, rather than using disease-free patients for controls, a sample of the entire cohort is chosen (which may, therefore, include a few patients with the disease).

intelligence test scores of 130 or higher, and were about ten years old when the study was initiated. As of 1960, the researchers had lost contact with only 1.7% of the sample, an amazing achievement.

Lester matched each of the 15 suicides identified by 1987 with a control subject on the basis of a mental health rating assigned in 1950. Thus, any subject committing suicide prior to 1950 could not be included in the study.

At the time of entry into the study, each child was rated by the parents on 25 personality traits. Only one predicted subsequent suicide -- the suicides were rated as less conscientious by their parents in 1921. The finding of only one significant difference in 25 possibilities is to be expected on the basis of chance given the statistical criteria used in the study. Lester concluded, therefore, that, parent ratings of their children did not predict subsequent suicide, once mental health was taken into account.

Studies over time can be *prospective* or *retrospective*, that is, following-up participants to see what will happens versus asking people about their past to see what happened. Prospective studies (as in Lester's [1991a] study reviewed above) are typically cohort studies, whereas retrospective studies are typically cross-sectional studies. In prospective studies, the researcher has better control over the data collected; retrospective studies often are forced to use data collected by others, sometimes for different purposes, and the data are sometimes incomplete and from out-dated inventories.[5] However, though these terms are commonly used, Friedman (1987) urged that they be discarded.

A RETROSPECTIVE STUDY

Lester (1991b) conducted a retrospective study of nonfatal suicide and childhood experiences of punishment. Using a data set for 441 male and 13 female prisoners in Vermont, the prisoners' lifetime occurrence of nonfatal suicide was examined for any association with their self-reports of how their parents punished them.

Significantly more of the prisoners who had engaged in nonfatal suicidal behavior at some point in their life (92 of the 454 prisoners -- 20.3%) were punished physically by their father than those who had not (49% versus 30%). There were no significant differences in maternal punishment. Lester examined the effect of experience of psychiatric hospitalization as a way of controlling for the degree of psychiatric disturbance in the two groups of prisoners, but this variable did not affect the association he had found between physical punishment by the father and suicidal behavior.

---

[5] One of the few prospective studies in progress is by Caroline Thomas (e.g., Graves and Thomas, 1991) in which 1046 entering medical students at Johns Hopkins University Medical School were tested on a variety of measures and followed up for the development of diseases and behaviors, including suicide.

Cohort studies provide the best evidence about the risk of disease development but, carried out prospectively, they are expensive and time-consuming. For example, the Terman study of gifted adolescents started in the 1920s and has employed staff to follow-up the 1500 cases ever since.

Cohort studies must first define a population -- a general population group or a specialized population which is more easily followed such as a group of insured workers or members of a profession. The cases can be followed up for the same period of time or for varying periods of time. For example, if the cohort is made up of consecutive patients seen at a clinic over a period of years, then the follow-up period is obviously longer for the patients seen early in the sequence than for those seen toward the end. The use of *person-years* controls for this problem -- a case followed-up for ten years contributes ten person-years, a case followed up for one year contributes only one person-year. However, this technique assumes that the risk of the disease remains roughly constant over the follow-up period.[6]

An extra source of bias in cohort studies, in addition to those mentioned in Section 3, is that study subjects are typically lost as the study progresses -- typically the longer the follow-up the harder it is to track down all of the subjects in the cohort.

Experimental Studies

Experimental studies involve some manipulation or intervention by the researcher. The subjects who undergo this intervention are called the *experimental group* while those who do not undergo this intervention are called the *control group*. If the researcher determines which subjects undergo the intervention and if these subjects are chosen at random, then the researcher can conclude that there is a cause-and-effect relationship between the intervention and the outcome.

Experimental research involves ethical problems. The information about subjects must be kept confidential, and there must be no serious harm to subjects in either the experimental group or the control group as a result of participation in the research. Most institutions which conduct research on humans have an ethics committee to review proposed research for how it deals with these ethical issues. For example, in research to test whether lithium is a useful medication for bipolar affective disorder, experimental research necessitated that the control subjects be given medication that resembled the lithium in appearance but which was actually inert. Many psychiatrists refused to participate in this research in Europe because they felt that it was unethical to withhold a potentially useful medication from the control subjects.

Experimental epidemiology is primarily concerned with testing procedures to prevent the disease. As noted above, the experimental and control subjects should be chosen randomly from the same pool, but this is often not feasible. For example, in studies to explore whether school

[6] *Clinical epidemiology* has come to mean the study of groups of patients with a disease. Studies of *the natural history of disease* are analogous to descriptive studies in epidemiology (especially cohort studies). Such studies enable us to predict the patient's future, that is, their *prognosis*. Therapeutic trials of medication can be viewed as experimental research in clinical epidemiology.

suicide prevention programs are effective, sometimes the program is given to some of the classes and not given to other classes. The classes can be chosen at random, but not the subjects.[7]

Ideally, experiments should be double-blind, that is, both the assistants involved in the study do not know to which group the subjects belong (experimental or control) and those involved in assessing, coding and analyzing the data do not know to which groups the subjects belong.

AN EXPERIMENTAL STUDY

Kalafat and Elias (1994) conducted a school-based suicide awareness program in a school for the 10th graders. Half of the students were assigned (by the school, not by the researchers) to health classes (in which the suicide awareness program was presented) and half to physical education by the school for one marking period and switched for the second marking period. Half the experimental and control students were given both a pre-test and a post-test questionnaire about suicide and, because taking a pre-test can affect responses on the post-test, half of the experimental and control students were given only the post-test.

The inferential statistics on the post-test scores indicated that the pre-test had no significant effect on subsequent knowledge about suicide, whereas the group (experimental versus control) had a significant impact on knowledge, with the experimental group having more accurate knowledge. Unfortunately, Kalafat and Elias did not report the mean "accurate knowledge" scores for the two groups.[8]

5: SOME FINAL ISSUES

Comparing behaviors in different populations is made difficult by the fact that these populations may differ in some crucial variable. For example, in comparing national suicide rates, we know that different national populations may have different age compositions. Thus, for suicide rates, it is common to *standardize* the suicide rates for age based on a standard population. However, suicide rates could also be standardized for other variables, though this is rarely done.

In Section 3, we discussed issues of reliability, but this referred to the *internal reliability* of the study. If the results of a study on one population generalize to other populations, then the results have external reliability. Results should never be assumed to have external reliability until they have been replicated by researchers independent of the original researchers.

Let us assume that we have identified a variable that is associated with, and perhaps increases the risk of, a behavior. It is important to ascertain whether this variable is *specific* to the disease or behavior we are studying, or whether it has a more general effect. For example, in

[7] Sometimes it is possible to use the subjects as their own control by making before-and-after comparisons.
[8] Kalafat and Elias also studied the impact of the program on attitudes toward suicide and responses to suicidal peers.

the research which indicates an association between sexual abuse in childhood and later suicidality, none of the research reviewed by Lester (1992) indicated specificity. Indeed, those who have been sexually abused in childhood show an increased incidence of all kinds of psychiatric and psychological problems in adulthood, only one of which is suicidal behavior.

Final Comment

Clearly, this brief discussion of epidemiology and its applications to suicidal behavior has not exhausted all of the concepts and information that could be presented. The aim has been merely to introduce the reader to some of the issues involved in epidemiology and to present some examples. It is hoped that interested readers will pursue the topic in more depth on their own.

REFERENCES

Araki, S., & Murata, K. (1987). Suicide in Japan. *Suicide & Life-Threatening Behavior*, 17, 64-71.

Bjarnason, T., & Thorlindsson, T. (1994). Manifest predictors of past suicide attempts in a population of Icelandic adolescents. *Suicide & Life-Threatening Behavior*, 24, 350-358.

Canetto, S. S., & Lester, D. (Eds.) (1995). *Women and suicidal behavior*. New York: Springer.

Clarke, R. V., & Lester, D. (1989). *Suicide: closing the exits*. New York: Springer-Verlag.

Farberow, N. L., MacKinnon, D., & Nelson, F. (1977). Suicide. *Public Health Reports*, 92, 223-232.

Friedman, G. D. (1987). *Primer of epidemiology*. New York: McGraw-Hill.

Goldney, R. D., Winefield, A. H., Tiggemann, M., & Winefield, H. R. (1995). Suicidal ideation and unemployment. *Archives of Suicide Research*, 1, 175-184.

Graves, P. L., & Thomas, C. B. (1991). Habits of nervous tension and suicide. *Suicide & Life-Threatening Behavior*, 21, 91-105.

Kalafat, J., & Elias, M. (1994). An evaluation of a school-based suicide awareness intervention. *Suicide & Life-Threatening Behavior*, 24, 224-233.

Lester, D. (1969). The subject as a source of bias in psychological research. *Journal of General Psychology*, 81, 237-248.

Lester, D. (1972). Self-mutilating behavior. *Psychological Bulletin*, 78, 119-128.

Lester, D. (1988a). An analysis of the suicide rates of birth cohorts in Canada. *Suicide & Life-Threatening Behavior*, 18, 372-378.

Lester, D. (1988b). What does the study of simulated suicide notes tell us? *Psychological Reports*, 62, 962.

Lester, D. (1991a). Completed suicide in the gifted. *Journal of Abnormal Psychology*, 100, 604-606.

Lester, D. (1991b). Physical abuse and physical punishment as precursors of suicidal behavior. *Stress Medicine*, 7, 255-256.

Lester, D. (1992). *Why people kill themselves*. Springfield, IL: Charles Thomas.

Lester, D. (1993). The effectiveness of suicide prevention centers. *Suicide & Life-Threatening Behavior*, 23, 263-267.

Lester, D. Differences in the epidemiology of suicide in Asian Americans by nation of origin. Omega, 1994, 29, 89-93.

Lester, D., & Wilson, C. (1988). Suicide in Zimbabwe. *Central African Journal of Medicine*, 34, 147-149.

Lester, D., Motohashi, Y., & Yang, B. (1992). The impact of the economy on suicide and homicide rates in Japan and the United States. *International Journal of Social Psychiatry*, 38, 314-317.

Neuringer, C. (1964). Rigid thinking in suicidal individuals. *Journal of Consulting Psychology*, 28, 54-58.

Norman, W. H., Brown, W. A., Miller, I. W., Keitner, G. I., & Overholser, J. C. (1990). The dexamethasone suppression test and completed suicide. *Acta Psychiatrica Scandinavica*, 81, 120-125.

O'Carroll, P. W., Berman, A. L., Maris, R., Moscicki, E., Tanney, B., & Silverman, M. (1995). Beyond the Tower of Babel: a nomenclature for suicidology. In D. Lester (Ed.) *Suicide '95*, pp. 136-138. Washington, DC: American Association of Suicidology.

Phillips, D. P., & Wills, J. S. (1987). A drop in suicides around national holidays. *Suicide & Life-Threatening Behavior*, 17, 1-12.

Robinson, W. S. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review*, 15, 351-357.

Séguin, M., Lesage, A., & Kiely, M. C. (1995). Parental bereavement after suicide and accident. *Suicide & Life-Threatening Behavior*, 25, 489-498.

Strober, M., Green, J., & Carlson, G. (1981). Reliability of psychiatric diagnosis in hospitalized adolescents. *Archives of General Psychiatry*, 38, 141-145.

Taiminen, T., Salmenperä, T., & Lehtinen, K. (1992). A suicide epidemic in a psychiatric hospital. *Suicide & Life-Threatening Behavior*, 22, 350-363.

Vega, W. A., Gil, A., Warheit, G., Apospori, E., & Zimmerman, R. (1993). The relationship of drug use to suicide ideation and attempts among African American, Hispanic, and white non-Hispanic male adolescents. *Suicide & Life-Threatening Behavior*, 23, 110-119.

Wilson, K. G. Stelzer, J., Bergman, J. N., Kral, M. J. Inayatullah, M., & Elliott, C. A. (1995). Problem-solving, stress, and coping in adolescent suicide attempts. *Suicide & Life-Threatening Behavior*, 25, 241-252.

Zimmerman, S. L. (1990). The connection between macro and micro levels. *Suicide & Life-Threatening Behavior*, 20, 31-55.