The Language of Health

Barbara K. Garland

Rural Health Institute

Introduction

Language: The special vocabulary and usages of a scientific, professional or other group. *Fourth definition, American Heritage Dictionary, 1991*

Health: Health is a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity. *World Health Organization, 2000*

The language of health consists of the terms and concepts that underlie beliefs and practices shared among those in the health professions. This common language ensures communication and understanding among health professionals from multiple disciplines and across many cultures.

The selected concepts presented here influence the way health professionals think about health and frame their views on how healthy people in healthy communities can be achieved. Knowing this language and how it is used is essential if we are to identify effective intervention points or successfully collaborate with those working in the health field.

Section 1: Vital Statistics

We are going to first look at vital statistics. These are the numbers collected in every county and transmitted to state and national data centers where they are aggregated (merged) to provide the big picture on how we are doing overall in health. These statistics include information on births, deaths, marriages and divorces. Of particular interest is cause of death published as leading causes of mortality. One use of this information is to determine where to concentrate resources in order to decrease excessive mortality and morbidity (illness). Making informed decisions about health programming requires being familiar with the appropriate vital statistics. Annual vital statistics publications can be obtained from your state office of vital statistics, state health department.

Vital statistics are provided as numbers and rates. Knowing only the total number of births or deaths in an area or from a specific cause has little meaning unless you are the hospital administrator, florist or undertaker and need to know how many beds, flowers and caskets to keep on hand. Numbers of events alone are referred to as numerator data.

Knowing the denominator or size of the group allows a rate to be calculated for comparison across groups and over time. A rate is the number of events in a given period of time divided by the number of people at risk of experiencing that event during that period. The rates used in reports of vital statistics may be calculated for the county, state or nation. The time period is usually one calendar year, but three and five year rates are often used when events are rare.

It is believed that the one most important vital statistic reflecting a society's ability to provide for the well-being of its population is the infant mortality rate. The infant mortality rate is the number of infant deaths for a specific year divided by all live births during that year. Infancy is considered to be from birth to one year. Rates are most often expressed as per 1,000 or per 100,000 depending on the frequency of the event. The infant mortality rate is given as the number of infant deaths per thousand live births.

Example

Infant Mortality Rate, 1999, Robbins County, U.S.

$$\frac{\text{68 deaths}}{\text{8743 live births}} = 0.007777 \times 1{,}000 = 7.8 \text{ or } 7.8/1000$$

This figure can be compared to infant mortality rates for the state, nation and other countries.

In 1998 the U.S. infant mortality rate was 7 per 1000. Thirty-one countries or territories had infant mortality rates between 3.2 and 6.7 per 1000. One explanation given for the higher U.S. rates is that we have a multi-cultural society while most other countries have more homogenous populations. A second explanation is the economic disparities and associated differences in access to health care that exist among various segments of the U.S. population. Obviously, these two explanations are related.

Other rates are calculated in much the same manner. Rates may be age, sex, race or disease specific, or they may take all of these factors into consideration. The more stratification (classifying), the larger the population needed for a meaningful rate. A problem in comparing rates is that in small populations very few excess deaths can have a drastic effect on the rate for a single year and distort the true picture. To compensate for the effect of small numbers, a three or five year rate may used. This rate is calculated by using the average annual resident deaths divided by the population mid-term.

In health promotion programming, information is frequently needed on the leading causes of death. Excessive or unusually high mortality rates for a specific cause is justification for directing scarce health resources toward targeted prevention efforts. There are several ways to look at mortality rates, and the one you choose will determine how useful the rate is.

The unadjusted or crude annual death rate is the number of resident deaths divided by the number of people living in the area on July 1 of that year. It is a simple observation of annual total or disease-specific deaths.

Many times these rates are used to compare experience across geographic areas. Doing so may be misleading. The problem is that without accounting for differences in ages or other factors between areas, unadjusted rates do not tell the real story. For instance, if the population of area one is older than area two, more deaths would be expected. It would not be possible to determine if area one was experiencing excessive mortality. For this an adjusted mortality rate is used. The adjusted mortality rate takes into consideration the age, race and sex distribution of the population, thereby making it useful for comparisons across areas and between population groups. Adjustments are achieved by applying both groups' age specific mortality rates to a standard. The standard frequently used for age adjustments is the United States 1970 population.

Mortality from any specific disease or injury is relatively rare. A very small percentage of the population dies from any one condition. For instance, about .00029 women die of breast cancer each year. To make comparisons less cumbersome, rates are expressed in per 100,000. For breast cancer this would be 29 deaths per 100,000. For more common events, the total mortality rate per 1000 may be used.

Example

Calculating the adjusted mortality rate allows us to know how many deaths would occur in each group if the age distribution was the same in each group.

Using the unadjusted and adjusted rates for California and Florida, compare 1998 rates for death from diseases of the heart.

Group Exercise

Break into state groups. Using the vital statistics book from your state, find the one and five year infant mortality rates for your state and county for whites and nonwhites. Compare these rates. What do they tell you? Are there differences among counties represented on the group?

Looking at mortality from breast cancer in women for your state and county, compare five year state and county adjusted death rates. How does your county compare? Why use the five year adjusted rate? If you were allocating funding for an early detection program, which of the counties represented by your state would receive funding. Why? Which counties would receive funding if 5-year unadjusted rates were used?

Section 2: Basic Measures Used in the Study of Health, Disease and Disability in Human Populations

What we know about health, disease and disability has been learned from the study of the occurrence of disease in human populations. The science which does this is epidemiology: epi meaning upon; demos, the people; and ology, the study of. Epidemiology is used to not only learn more about health, disease and injury in groups of people but also to determine at which point in the disease or injury process an intervention can effectively prevent it or limit its effects. Planning and implementing effective health promotion and disease and injury prevention programs depends on identifying and targeting this intervention point.

Example

Let's look at breast cancer. It is predicted that there will be about 183,000 new cases of breast cancer in 2000. This disease killed over 43,000 women in 1999. We can use numerator data here because we are referring to the entire universe of women in the U.S. population and not comparing them to any other group. Besides, it is more impressive than saying 29 per 100,000. However, when we look at change over time or between countries, we have to use rates so as to account for the differences in the denominator, the total populations in each group. Since we do not know the cause of breast cancer (we do not even have a good guess at this time), we cannot prevent it. So what can we do to save the lives of women stricken with this disease? At what point can we inter vene effectively?

Epidemiologic research has shown that when breast cancer is detected at a very early stage, the probability of surviving five years is 95 percent versus less than 20 percent for tumors found in an advanced stage. Early detection then is the most effective intervention strategy for breast cancer. It is estimated that up to 30 percent of breast cancer mortality can be prevented through early detection.

What do we need to know to design effective health programs? First we need information on the distribution of the disease within the population of interest. We need to know who is effected, where they reside, and when or under what circumstances they became ill (person, place and time). This information can be combined with data from other sources such as genetics, biochemistry and microbiology to help find the etiology or cause of the disease or disorder. Only then can the appropriate intervention points be determined and effective interventions be developed. In program evaluation, rates of disease or injury are compared before and after the intervention.

These include mortality, prevalence and incidence rates. The mortality rate equals the number of people dying divided by the total number in the population over a certain time period. Mortality rates can be calculated for specific cause of death among a population-at-risk (PAR). They can also be adjusted for age, race and sex.

A prevalence rate is the number of people with a disease at a point in time divided by the total number of people in the group. If you went out and did a door-to-door survey of your county, how many people would you find that have been diagnosed with hypertension? Divide that number, the numerator, by the total population of the county on July 1, the denominator. That is the prevalence rate. The prevalence rate gives a picture of the situation at one point in time. It is useful in determining health and health-related services needed such as number of hospital beds and home health services.

The incidence rate tells different a story. The incidence rate refers to the number of new cases occurring in a population within a specified period of time. It is the number of new cases divided by the total number of people at risk of the disease within a point in time. In calculating the incidence rate, the number of the population-at-risk is used in the denominator rather than the total population because not everyone in a group may be at risk of developing the disease under consideration. For instance, some conditions such as prostate or ovarian cancer only occur in specific gender groups, or else they already have the disease (diabetes, cancer, heart disease).

Although prevalence and incidence rates are sometimes confused even by some health professionals, it is important to know the difference. Also, when examining data, beware of confusing incidence and mortality. While mortality rates are readily available, good data on incidence does not exist for most diseases. Incidence only equals mortality when 100 percent of the new cases die within the time period under consideration. This may be true for some acute conditions. It is almost true for lung cancer where 75 percent of newly diagnosed cases die within months. On the other hand, chronic diseases tend to progress slowly, leading to an accumulation of cases (greater prevalence) than new cases (incidence) over time. The relationship between incidence and prevalence is as follows: prevalence equals incidence times survival time with the disease. The incidence rate gives us a picture of the dynamics of the disease over time within a population. Is the disease increasing or decreasing? For whom?

Prevalence may thus be many times incidence if a disease leads to a chronic state of long standing. Why do we care about all this? Because not understanding the differences can lead to errors in interpreting information, unsound

Example
In the state of Zandu there were 3,660 new breast cancer cases among women in 1999. Eight hundred women died of breast cancer. A household survey identified 14,300 women who had been diagnosed with breast cancer within the past five years. There were 2.8 million women who were 18 and older. The total population was 5 million.

The Zandu 1999 breast cancer mortality rate for women = 800/2,800,000 = .0002857 x 100,000 = 28.6 per 100,000 women

The Zandu 1999 breast cancer incidence rate for women = 2,800,000 - 14,300 = 2,785,700.
3660/2,800,000 = .001307 x 100,00 = 130.7 per 100,000 women

The Zandu 1999 breast cancer prevalence rate for women = 14,300/2,800,000 = 510.7 per 100,000 women

planning and a loss of credibility among collaborators.

Incidence rates are often used to compare the experiences of two or more dif-

Example
If non-smokers have a lung cancer incidence of 19 per 100,000, and the rate is 188 per 100,000 for smokers, we can divide 188 by 19 for a risk ratio of 9.9. This measure of risk then becomes useful in predicting lung cancer rates in smoking and non-smoking groups. These risk estimates are the basis for the risk factors which have been shown to be associated with various diseases.

ferent groups. By comparing the two rates, one can determine which group is at greater risk. We can divide the two rates and obtain a risk ratio, or relative risk.
Risk factors are associated with an increased risk of becoming diseased or sustaining an injury. They may be found in the physical environment as toxins, infectious agents, drugs, or present as part of one's genetic heritage. Risk factors may be part of the social environment such as in loss of a family member, social isolation, or aspects of a particular culture. Behavioral risk factors include tobacco use, diet, inactivity, aggressive driving habits, and inappropriate use of equipment.

Risk factors are most commonly used to predict the occurrence of disease or injury. Although risk factors are associated with specific diseases, not all risk

Example
Maternal education is associated with low birth weight. The less education the mother has, the greater risk of her having a low birth weight infant. Yet it is known that smoking, poor diet and lack of prenatel care are the direct causes of low birth weight. Low educational level although indirectly associated, is not a cause of low birth weight.

Most (some say all) diseases are multi-causal. With the possible exception of some infectious diseases, there is more than one causal factor involved in the occurrence of any disease or injury. Prevention requires removal of the risk factors directly related to the disease or injury.

factors are causes. Some risk factors may be indirectly associated with a disease.

Elicit further examples from group discussion.

Section 3: Determining Abnormality

If we are going to plan and implement effective interventions within a population group, we have to know either who has the disease of interest or, in the case of interventions involving early detection, how to find out. Unfortunately, this may not be straight forward or easy. There is a spectrum of disease ranging from no disease to a subclinical state in which there is no apparent sign of disease to clinical illness and, if severe, death. The earlier in the disease process intervention occurs, the more effective it is likely to be. If intervention occurs before the disease process begins or injury occurs, it is primary prevention. If early detection is undertaken during the sub-clinical or asymptomatic period, it is secondary prevention, and if there is intervention to prevent complications from a clinical condition, it is tertiary prevention.

Primary prevention includes weight loss and exercise programs, screening for blood lipid levels, and safety programs. For those who may be at risk but have not been diagnosed, we may be asked to implement or collaborate in implementing screening programs aimed at the early detection of a disease or disorder - secondary prevention. Examples of these programs include screening mammographies for the early detection of breast cancer or PSA testing for prostate cancer. Not all of these tests are equally useful for their intended purpose. A test may be valuable in the clinical diagnoses of a disease but still not useful for screening entire populations due to complexity, acceptability or cost. What do you say when women participating in a breast cancer program demand equal time for prostate screening for their husbands? Or when some women want to know why the CA 125 blood test for ovarian cancer is not being made available to them as a screening test?

The answer often depends on the predictive value of the individual test, which

in turn depends on its reliability and its validity as shown by sensitivity and specificity. Let's look at how each of these contributes to the usefulness of a test.

First let's consider reliability. Another term used is reproducibility. These terms refer to the ability of a test to give the same result when repeated under the same circumstances. This does not mean that the test is providing a valid or accurate assessment. Validity refers to the degree to which a test provides the

Example
It has been shown that in trying to determine the dietary intake of a population, repeated surveys using 24-hour recalls provide the same general picture. For instance, pregnant women report an intake of about 1600 calories a day. However, when a group of pregnant women are observed eating throughout the day and the food served is weighed before and after meals, the calorie intake is about 2400 calories. This would suggest that 24-hour recalls are reliable but not accurate or valid.

true answer or how accurate it is. One way to remember the difference is to imagine shooting an arrow at a target. If you hit the same spot on the target each time, you're reliable. One can rely on your arrow hitting that spot. If you hit the bull's eye, your aim is accurate. The shot is valid.
How is the validity of a test determined? There are two indexes used - sensitivity and specificity. The sensitivity of a test is how often it detects the disease when in fact the disease is present; how many people with the disease have a positive test. These are the true positives. This can be calculated as a percent. No test is 100 percent sensitive. Sensitive tests are of value when the probability of the disease is relatively low and the population at high risk can be identified. The specificity of a test is how often it indicates.

Example
A mammogram, the used test for breast cancer screening, may miss from 10 to 20 percent or more of existing breast cancer. Its sensitivity ranges from about 90 percent for women age 50 and older to 80 percent in women 40 to 49 and even lower in women 35 to 40. Why are so many cancerous tumors missed? The problem may be the location of the tumor or the density of the breast. A mammogram may not detect a tumor in an obscure location. Also, since younger women tend to have breast tissue that is about the same density as the tumor embedded in it, a mammogram may not be as useful for early detection of breast cancer in younger women. The tumor cannot be seen on an X-ray. The sensitivity of a mammogram depends largely on the ratio of fat to tissue in the breast.

There are some tests that are highly sensitive and some that are highly specific. There are very few that are both. As sensitivity goes up, specificity goes down and vice versa. Most tests work best when there is reason to believe that the individual undergoing the test may have the disease, or is in the high risk group.
The use of hormonal replacement therapy further complicates the picture since it may increase the density of the breast and thus lower the sensitivity of the test in older women.

Why do we use tests with low sensitivity, specificity or both? Because this is the best that we have at this time. And that is true for every screening and diagnostic test. There are other considerations in whether to use a test, particularly for population screening. The first of these is, is finding the disease going to decrease morbidity or mortality? Do we have an effective treatment or the facilities for treatment? If not, early detection may merely increase physical and mental suffering without helping the individual. An example of an untreatable disease is Alzheimer's, where early diagnosis may only add to the patient's and family's distress with no identifiable benefit.

Secondly, how rare is the disease? Broad testing for rare conditions requires that very large populations be tested. This may be cost prohibitive. For instance, unless certain defined groups can be identified as at high risk for HIV, it would not be cost effective to test the entire population because HIV infection is so rare. A third issue is how many people would have positive results when in fact there is no disease (low specificity)? To flood the health care system with these false positives would add an additional cost burden in addition to raising the anxiety levels of the individuals affected. It may also lead to actual harm if healthy individuals are subjected to unnecessary medical procedures.

In the case of a potentially lethal disease such as breast cancer, some of these considerations may be less important than detecting the disease and saving lives where possible. For younger women, this point is still being debated. Screening mammographies are recommended despite the lack of high sensitivity in all cases and relatively low specificity. They have been shown to save lives especially in women over the age of 50.

There is currently a movement among some women's groups for widespread screening for ovarian cancer. It has been discovered that women with high levels of a tumor marker, CA 125, have a greater than average risk of developing ovarian cancer. The test, however, is neither sensitive nor specific enough for use in screening. CA 125 is not always found in women with ovarian cancer (low sensitivity). Conversely it may be found in women with other cancers or in women with non-cancerous ovarian conditions (low specificity). It is not an

accurate screening test for ovarian cancer. Presently, CA 125 testing is used as one of several tests to determine the re-occurrence of ovarian cancer in women who have been treated for the disease, a very high risk group.

And then there is prostate cancer, which at first glance may seem to parallel breast cancer as an opportunity for saving lives through screening. It is known that a substantial proportion of the older male population will develop prostate cancer and that this disease ranks second as cause of death from cancer among males. Because of this, there has been a demand for population screening. For prostate cancer the situation is not that clear. Although high levels of a prostate specific antigen (PSA) in the blood may indicate prostate cancer, PSA levels may also be higher in men with non-cancerous prostate conditions (low specificity). The situation is further complicated by the fact that it is also known that most older men will have evidence of prostate cancer at death but will have died from other causes. Since unnecessary treatment for prostate cancer due to false screening results may be harmful, and just when and how to treat the disease is still not entirely clear, population screening is not advised at this time.

Here we have been addressing screening tests. Because no screening test is 100 percent sensitive or specific, it usually takes multiple tests to determine a disease state with certainty. These tests contribute to the clinical diagnosis of a disease.

Section 4: Study Methods and Rules of Evidence

With the new and often contradictory discoveries on health and disease being made public daily, knowing what to believe, much less to convey to others, has become a major problem. How do we sort it all out?

One way is to have some understanding of where the information came from; how conclusions were arrived at. There are three major categories of health research: descriptive studies which provide information on the patterns of disease occurrence in human populations, observational studies which allow us to follow the course of disease, and experimental studies which involve an intervention.

Descriptive Studies

Descriptive studies focus on personal characteristics of diseased and non-diseased individuals including age, race or ethnic origin, gender, occupation, and

Example

Studies describing the high rates of breast cancer among women of Jewish descent living on Long Island led to the discovery of a specific gene which predisposes certain ethnic groups to breast cancer.

social class. Also important are geographic area and time. Tracking person, place and time are important for alerting the community as to where cases may be expected or when incidence is greater than expected (i.e., there is an epidemic). Descriptive studies, then, assist in health care planning and response. They can also provide clues to the etiology (cause) of disease.

Observational Studies

Observational studies are those which seek to determine cause. Observational studies involve the observation of events as they are, as they were, or as they are happening.

The first of these are prevalence studies, also called cross-sectional studies. These studies measure the prevalence of disease or injury in a population at one point in time. They also involve collecting information on certain characteristics of the population being surveyed, both those with and those without the disease. Associations or relationships can be shown from this data; cause, however, cannot because it is not possible to establish the time sequence. Did the development of the disease precede the event? Prevalence studies may suggest hypotheses for further study but do not themselves establish cause. Claims that a cause of disease has been determined from a prevalence study - a survey showed - can be dismissed out-of-hand. Prevalence studies are limited to conditions that are relatively common.

The second observational method is retrospective. That is, persons with a disease or injury are interviewed to determine what the situation was in the past. This is then compared to information on the same characteristics and events as recalled by a control group of individuals without the disease. Much of the data collected depends on the individual's ability to recall events. These case-control or retrospective studies have a number of problems, in addition to recall bias, including establishment of time sequence and selection of controls. But because case-control studies begin with a pool of patients, they are particularly appropriate for studying rare diseases. Case-control studies may generate hypotheses for further study but do not by themselves establish cause. It has been found, however, that multiple case-control studies in which findings concur can be an efficient method of approaching cause.

The third type of observational study is the prospective or cohort study, which begins with a cross-section of a population, notes the characteristics of interest, determines and drops from the study those with the disease, and then follows the initially well group over time to observe the onset of disease. Cohort studies also subjects to study and control groups. Another name given these studies is incidence studies since they are determining relative incidence between groups. These studies can determine time sequence and, despite having other problems such as potential loss of participants over time, are the only observational studies from which cause can be ascertained. The

term cohort is derived from the early Roman legions who fought to the last man. A subtype of the cohort study is the randomized trial in which participants are randomized to the exposure.

So why are cross-sectional and case control studies done more often than cohort studies? Because they are quicker, far less expensive, and more timely. While cohort studies provide answers with greater certainty, repeated case-control studies have shown to often be almost as useful in less time and at far less cost. Further, in the case of a rare disease, the time and cost of recruiting and following enough people to amass a group large enough for study would be prohibitive.

Experimental or Intervention Studies

The difference between observational studies and experimental or intervention studies is that in an experimental study the investigator intervenes so as to affect the outcome of the study. Some action is taken rather than observation only. This is usually removing or reducing the alleged cause from treating or modifying treatment in the study group. Otherwise these studies resemble cohort studies in that they are prospective, are randomized, and have control groups.

Because of the investigator's ability to manipulate the study variables, cause and effect can be demonstrated. Changes between outcomes in the study and control groups are more clearly shown. The major drawback in experimental studies, whether clinical trials or community intervention studies, is the possibility of unethical behavior on the part of the investigator. Before doing anything to groups of people, the investigator must have solid evidence that the intervention is going to be helpful and not harm the individuals involved. To protect human subjects, institutions employing researchers must have Institution Review Boards (IRBs) to review and approve any proposed research involving human subjects.

Rules of Evidence

As seen from the foregoing discussion, determining cause is a painstaking process. We first accept that all studies are inherently flawed in some way - sometimes seriously, sometimes in minor ways. Additionally, no disease or injury is likely to have a single cause.

Sometimes multiple causes interact to produce an effect greater than the effect of either of them alone or of adding their separate effects together.

An example of this is cigarette smoking, alcohol and lung cancer. Cigarette smokers have a risk of lung cancer that is nine times that of non-smokers.

Alcohol alone does not cause lung cancer. However, the combination of heavy alcohol consumption and smoking increases the risk of lung cancer to sixteen times that of non-smokers.

So how, to the best of our ability, do we determine actual cause or causes in order to identify the most appropriate intervention point for action or to determine an intervention point when cause is not known as in breast cancer? Or even to withstand the constant onslaught of breaking news on health? We look to the rules of evidence[1] that have been formulated to help us make informed decisions in the health arena. These are presented from the strongest to the weakest.

1. **Temporal relationship between cause and effect.** Cause precedes effect. Only prospective studies - cohort and experimental - can establish temporal relationships.

2. **Strength of the association.** The larger the risk ratio, the stronger the evidence for causation. Relative risks of two or more are considered significant.

3. **Dose-Response relationships.** When increases in dose or exposure to the purported cause results in a corresponding increase in effect, a dose-response relationship exists. This strengthens the argument for cause and effect. A caution here is that there might be some other factor which is related to both purported cause and the effect which is influencing both of them.

4. **Reversible associations.** If removing a risk factor results in a decreased risk of disease, i.e., the effect can be reversed, it is more likely that the factor is a cause.

5. **Consistency.** When several studies conducted at different times under different circumstances produce the same results, evidence for a causal relationship is strengthened. That does not mean, however, that several seriously flawed studies outweigh one well designed and carried out study.

6. **Biologic plausibility.** Biologic plausibility depends on our knowledge of the mechanisms of the disease under study. Is there a reasonable biological explanation for the findings of cause and effect? Sometimes these explanations do not exist at the time of the findings. Other times findings simply do not appear to make biological sense as in the case of the purported cancer cure Laetrile.

7. **Specificity.** This refers to one cause, one effect. This is more likely to be so in acute infections but not in chronic diseases. One factor may cause several diseases, and several factors may be necessary to cause a single disease.
8. **Analogy.** Finding an analogous example to compare a finding to helps to strengthen the case for cause and effect. This is the weakest of the rules of evidence.

Again, why do we care about all this? Because if Extension is going to develop and deliver health promotion programs or work with those who do, we need to be able to use a common language, interpret health information, and use this understanding to identify the appropriate intervention points for effective programming.