ERIC Notebook

Second Edition

Common Measures and Statistics in Epidemiological Literature

Second Edition Authors:

Lorraine K. Alexander, DrPH

Brettania Lopes, MPH

Kristen Ricchetti-Masterson, MSPH

Karin B. Yeatts, PhD, MS

UNC GILLINGS SCHOOL OF GLOBAL PUBLIC HEALTH

For the non-epidemiologist or non-statistician, understanding the statistical nomenclature presented in journal articles can sometimes be challenging, particularly since multiple terms are often used interchangeably, and still others are presented without definition. This notebook will provide a basic introduction to the terminology commonly found in epidemiological literature.

Measures of frequency

Measures of frequency characterize the occurrence of health outcomes, disease, or death in a population. These measures are descriptive in nature and indicate how likely one is to develop a health outcome in a specified population. The three most common measures of health outcome or frequency are risk, rate, and prevalence.

Risk

Risk, also known as incidence, cumulative incidence, incidence proportion, or attack rate (although not really a rate at all) is a measure of the probability of an unaffected individual developing a specified health outcome over a given period of time. For a given period of time (i.e.: 1 month, 5 years, lifetime):

A 5-year risk of 0.10 indicates that

$$\text{Risk} = \frac{\#\ of\ new\ cases}{total\ \#\ individuals\ at\ risk}$$

an individual at risk has a 10% chance of developing the given health outcome over a 5-year period of time.

Risk is generally measured in prospective studies as the population at risk can be defined at the start of the study and followed for the development of the health outcome. However, risk cannot be measured directly in case-control studies as the total population at risk cannot be defined. Thus, in case-control studies, a group of individuals that have the health outcome and a group of individuals that do not have the health outcome are selected, and the odds of developing the health outcome are calculated as opposed to calculating risk.

$$\text{Odds} = \frac{\#\ individuals\ with\ the\ health\ outcome}{\#\ individuals\ without\ the\ health\ outcome}$$

Rate

A rate, also known as an incidence rate or incidence density, is a measure of how quickly the health outcome is occurring in a population. The numerator is the same as in risk, but the denominator includes a measure of person-time, typically person-years. (Person-time is defined as the sum of time that each at-risk individual contributes to the study).

$$\text{Rate} = \frac{\#\ of\ new\ cases}{total\ person\text{-}time\ at\ risk}$$

Thus a rate of 0.1 case/person-years indicates that, on average, for every 10 person-years (i.e.: 10 people each followed 1 year or 2 people followed

for 5 years, etc.) contributed, 1 new case of the health outcome will develop.

Prevalence

Prevalence is the proportion of a population who has the health outcome at a given period of time. Prevalence is generally the preferred measure when it is difficult to define onset of the health outcome or disease (such as asthma), or any disease of long duration (e.g. chronic conditions such as arthritis). A limitation of the prevalence measure is that it tends to favor the inclusion of chronic diseases over acute ones. Also, inferring causality is troublesome with prevalence data, as typically both the exposure and outcome are measured at the same time. Thus it may be difficult to determine if the suspected cause precedes the outcome of interest.

$$\text{Prevalence} = \frac{\#\ of\ affected\ individuals}{total\ \#\ of\ individuals\ in\ the\ population}$$

Thus a population with a heart disease prevalence of 0.25 indicates that 25% of the population is affected by heart disease at a specified moment in time.

A final note, risk and rates can also refer to deaths in a population and are termed mortality and mortality rate, respectively.

Measures of association

Measures of association are utilized to compare the association between a specific exposure and health outcome, They can also be used to compare two or more populations, typically those with differing exposure or health outcome status, to identify factors with possible etiological roles in health outcome onset. Note that evidence of an association does not imply that the relationship is causal; the association may be artifactual or non-causal as well. Common measures of association include the risk difference, risk ratio, rate ratio and odds ratio.

Risk difference

Risk difference is defined as

$$= Risk_{exposed} - Risk_{unexposed}$$

$$\frac{\#\ cases\ in\ exposed\ group}{total\ \#\ at\ risk\ in\ exposed\ group} - \frac{\#\ cases\ in\ control\ group}{total\ \#\ at\ risk\ in\ control\ group}$$

The risk difference, also know as the attributable risk, provides the difference in risk between two groups indicating how much excess risk is due to the exposure of interest. A positive risk difference indicates excess risk due to the exposure, while a negative result indicate that the exposure of interest has a protective effect against the outcome. (Vaccinations would be a good example of an exposure with a protective effect). This measure if often utilized to determine how much risk can be prevented by an effective intervention.

Risk ratio and rate ratio

Risk ratios or rate ratios are commonly found in cohort studies and are defined as: the ratio of the risk in the exposed group to the risk in the unexposed group or the ratio of the rate in the exposed group to the rate in the unexposed group

$$\text{Risk Ratio} = Risk_{exposed}/Risk_{unexposed}$$

$$\text{Rate Ratio} = Rate_{exposed}/Rate_{unexposed}$$

Risk ratios and rate ratios are measures of the strength of the association between the exposure and the outcome. How is a risk ratio or rate ratio interpreted? A risk ratio of 1.0 indicates there is no difference in risk between the exposed and unexposed group. A risk ratio greater than 1.0 indicates a positive association, or increased risk for developing the health outcome in the exposed group. A risk ratio of 1.5 indicates that the exposed group has 1.5 times the risk of having the outcome as compared to the unexposed group. Rate ratios can be interpreted the same way but apply to rates rather than risks.

A risk ratio or rate ratio of less than 1.0 indicates a negative association between the exposure and outcome in the exposed group compared to the unexposed group. In this case, the exposure provides a protective effect. For example, a rate ratio of 0.80 where the exposed group received a vaccination for Human Papillomavirus (HPV) indicates that the exposed group (those who received the vaccine) had 0.80 times the rate of HPV compared to those who were unexposed (did not receive the vaccine).

One of the benefits the measure risk difference has over the risk ratio is that it provides the absolute difference in risk, information that is not provided by the ratio of the two. A risk ratio of 2.0 can imply both a doubling of a very small or large risk, and one cannot determine which is the case unless the individual risks are presented.

Odds ratio

Another measure of association is the odds ratio (OR). The formula for the OR is:

$$\text{Odds ratio} = \frac{odds_{exposed}}{odds_{unexposed}}$$

The odds ratio is used in place of the risk ratio or rate ratio in case-control studies. In this type of study, the underlying population at risk for developing the health outcome or disease cannot be determined because individuals are selected as either diseased or non-diseased or as having the health outcome or not having the health outcome. An odds ratio may approximate the risk ratio or rate ratio in instances where the health outcome prevalence is low (less that 10%) and specific sampling techniques are utilized, otherwise there is a tendency for the OR to overestimate the risk ratio or rate ratio.

The odds ratio is interpreted in the same manner as the risk ratio or rate ratio with an OR of 1.0 indicating no association, an OR greater than 1.0 indicating a positive association, and an OR less than 1.0 indicating a negative, or protective association.

The null value

The null value is a number corresponding to no effect, that is, no association between exposure and the health outcome. In epidemiology, the null value for a risk ratio or rate ratio is 1.0, and it is also 1.0 for odds ratios and prevalence ratios (terms you will come across). A risk ratio, rate ratio, odds ratio or prevalence ratio of 1.0 is obtained when, for a risk ratio for example, the risk of disease among the exposed is equal to the risk of disease among the unexposed.

Statistical testing focuses on the *null hypothesis*, which is a statement predicting that there will be no association between exposure and the health outcome (or between the assumed cause and its effect), i.e. that the risk ratio, rate ratio or odds ratio will equal 1.0. If the data obtained from a study provide evidence against the null hypothesis, then this hypothesis can be rejected, and an alternative hypothesis becomes more probable.

For example, a null hypothesis would say that there is no association between children having cigarette smoking mothers and the incidence of asthma in those children. If a study showed that there was a greater incidence of asthma among such children (compared with children of nonsmoking mothers), and that the risk ratio of asthma among children of smoking mothers was 2.5 with a 95% confidence interval of 1.7 to 4.0, we would reject the null hypothesis. The alternative hypothesis could be expressed in two ways: 1) children of smoking mothers will have either a higher or lower incidence of asthma than other children, or 2) children of smoking mothers will only have a higher incidence of asthma. The first alternative hypothesis involves what is called a *"two-sided test"* and is used when we simply have no basis for predicting in which direction from the null value exposure is likely to be associated with the health outcome, or, in other words, whether exposure is likely to be beneficial or harmful. The second alternative hypothesis involves a *"one-sided test"* and is used when we have a reasonable basis to assume that exposure will only be harmful (or if we were studying a therapeutic agent, that it would only be beneficial).

Measures of significance

The p-value

The "p" value is an expression of the probability that the difference between the observed value and the null value has occurred by "chance", or more precisely, has occurred simply because of sampling variability. The smaller the "p" value, the less likely the probability that sampling variability accounts for the difference. Typically, a "p" value less than 0.05, is used as the decision point, meaning that there is less than a 5% probability that the difference between the observed risk ratio, rate ratio, or odds ratio and 1.0 is due to sampling variability. If the "p" value is less than 0.05, the observed risk ratio, rate ratio, or odds ratio is often said to be "statistically significant." However, the use of 0.05 as a cut-point is arbitrary. The exclusive use of "p" values for interpreting results of epidemiologic studies has been strongly discouraged in the more recent texts and literature because research on human health is not conducted to reach a decision point (a "go" or "no go" decision), but rather to obtain evidence that there is reason for concern about certain exposures or lifestyle practices or other factors that may adversely influence the health of the public. Statistical tests of significance, (such as p-values) were developed for industrial quality-control purposes, in order to make a decision whether the manufacture of some item is achieving acceptable quality. We are not making such decisions when we interpret the results of research on human health.

The lower bound of the 95% confidence interval is also often utilized to decide whether a point estimate is statistically significant, i.e. whether the measure of effect (e.g. the ratio 2.5 with a lower bound of 1.8) is statistically different than the null value of 1.0.

Measures of precision

Confidence interval

A *confidence interval* expresses the extent of potential variation in a point estimate (the mean value or risk ratio, rate ratio, or odds ratio). This variation is attributable to the fact that our point estimate of the mean or risk ratio, rate ratio, or odds ratio is based on some sample of the population rather than on the entire population.

For example, from a clinical trial, we might conclude that a new treatment for high blood pressure is 2.5 times as effective as the standard treatment, with a 95% confidence interval of 1.8 to 3.5. 2.5 is the *point estimate* we obtain from this clinical trial. But not all subjects with high blood pressure can be included in any study, thus the estimate of effectiveness, 2.5, is based on a particular sample of people with high blood pressure. If we assume that we could draw other samples of persons from the same underlying population as the one from which subjects were obtained for this study, we would obtain a set of point estimates, not all of which would be exactly 2.5. Some samples would be likely to show an effectiveness less than 2.5, and some greater than 2.5.

The 95% CI is an interval that will contain the true, real (population) parameter value 95% of the time if you repeated the experiment/study. So if we were to repeat the experiment/study, 95 out of 100 intervals would give an interval that contains the true risk ratio, rate ratio or odds ratio value. Remember, that you can only interpret the CI in relation to talking about repeated sampling. Thus we can also say that the new treatment for high blood pressure is 2.5 times as effective as the standard treatment, but this measure could range from a low of 1.8 to a high of 3.5.

The confidence interval also provides information about how precise an estimate is. The tighter, or narrower, the confidence interval, the more precise the estimate. Typically, larger sample sizes will provide a more precise estimate. Estimates with wide confidence intervals should be interpreted with caution.

Other terms

Crude and adjusted values

There are often two types of estimates presented in research articles, *crude* and *adjusted* values. Crude estimates refer to simple measures that do not account for other factors that may be driving the estimate. For instance, a crude death rate would simply be the number of deaths in a calendar year divided by the average population for that year. This may be an appropriate measure in certain circumstances but could become problematic if you want to compare two or more populations that vary on specific factors known to contribute to the death rate. For example, you may want to compare the death rate for two populations, one of which is located in a high air pollution area, to determine if air pollution levels affect the death rate. The high air pollution population may have a higher death rate, but you also determine that it is a much older population. As older individuals are more likely to die, age may be driving the death rate rather than the pollution level. To account for the difference in age distribution of the populations, one would want to calculate an *adjusted* death rate that adjusts for the age structure of the two groups. This would remove the effect of age from the effect of air pollution on mortality.

Adjusted estimates are a means of controlling for confounders or accounting for effect modifiers in analyses. Some factors that are commonly adjusted for include gender, race, socioeconomic status, smoking status, and family history.

References

Dr. Carl M. Shy, Epidemiology 160/600 Introduction to Epidemiology for Public Health course lectures, 1994-2001, The University of North Carolina at Chapel Hill, Department of Epidemiology

Rothman KJ, Greenland S. Modern Epidemiology. Second Edition. Philadelphia: Lippincott Williams and Wilkins, 1998.

The University of North Carolina at Chapel Hill, Department of Epidemiology Courses: Epidemiology 710, Fundamentals of Epidemiology course lectures, 2009-2013, and Epidemiology 718, Epidemiologic Analysis of Binary Data course lectures, 2009-2013.

Acknowledgement

The authors of the Second Edition of the ERIC Notebook would like to acknowledge the authors of the ERIC Notebook, First Edition: Michel Ibrahim, MD, PhD, Lorraine Alexander, DrPH, Carl Shy, MD, DrPH and Sherry Farr, GRA, Department of Epidemiology at the University of North Carolina at Chapel Hill. The First Edition of the ERIC Notebook was produced by the Educational Arm of the Epidemiologic Research and Information Center at Durham, NC. The funding for the ERIC Notebook First Edition was provided by the Department of Veterans Affairs (DVA), Veterans Health Administration (VHA), Cooperative Studies Program (CSP) to promote the strategic growth of the epidemiologic capacity of the DVA.