ERIC Notebook

Second Edition

Cross-sectional Studies

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

Like cohort studies, cross-sectional studies conceptually begin with a population base. But unlike cohort studies, in cross-sectional studies we do not follow individuals over time. Instead, we only look at the prevalence of disease and/or exposure at one moment in time. These studies take a "snapshot" of the proportion of individuals in the population that are, for example, diseased and non-diseased at one point in time. Other health outcomes besides diseases may also be studied. Cross-sectional studies also differ from cohort studies in the populations that are studied. Cohort studies begin by selecting a population of persons who are at risk of for a specific disease or health outcome; cross-sectional studies begin by selecting a sample population and then obtaining data to classify all individuals in the sample as either having or not having the health outcome.

	Cohort	Cross-sectional
Study group	Population-at-risk	Entire population (or a sample)
Common Measures	Risks and Rates	Prevalence

Ways to use cross-sectional studies

Cross-sectional studies are used both descriptively and analytically. *Descriptive cross-sectional studies* simply characterize the prevalence of a health outcome in a specified population. Prevalence can be assessed at either one point in time (*point prevalence*) or over a defined period of time (*period prevalence*). Period prevalence is required when it takes time to accumulate sufficient information on a disease in a population, e.g. what proportion of persons served by a public health clinic over a year have hypertension. These prevalence measures are commonly used in public health; often the point or period aspect is not specified.

In *analytical cross-sectional studies,* data on the prevalence of both exposure and a health outcome are obtained for the purpose of comparing health outcome differences between exposed and unexposed.

Analytical studies attempt to describe the prevalence of, for example, disease or non-disease by first beginning with a population base. These studies differ from solely descriptive cross-sectional studies in that they compare the proportion of

exposed persons who are diseased (a/(a+b)) with the proportion of non-exposed persons who are diseased (c/(c+d)).

Calculating prevalence

The prevalence of a health outcome is simply the proportion of individuals with the health outcome in a population.

Prevalence = cases / total population

For the following example, two different sub-measures of prevalence can be calculated: the prevalence of coronary heart disease (CHD) among the exposed (people who are not active) and the prevalence of CHD among the unexposed.

Example:

	Present CHD		Absent CHD		Total
Not active	50	a	b	200	250
Active	50	c	d	700	750
Total	100			900	1000

P_1= a/a+b= 50/250 = 20.0% prevalence of CHD among people who are not active.

P_0= c/c+d = 50/750 = 6.7% prevalence of CHD among people who are active.

The prevalence odds ratio

The prevalence odds ratio (POR) is calculated in the same manner as the odds ratio.

POR = ad / bc

The prevalence ratio

The prevalence ratio (PR) is analogous to the risk ratio (RR) of cohort studies. The denominators for both ratios are fixed populations – fixed at the start of the study in the case of a cohort study, and fixed at the point or period of time for the case-control study. The PR is similar to a RR when the outcome occurs over a short period of time. For example, one would calculate a prevalence ratio for an acute outbreak of tuberculosis in a prison population. This is in contrast to calculating the overall prevalence of positive tuberculin skin tests among the prisoners.

The prevalence ratio can also be calculated from the information on CHD and physical activity. It is preferable to calculate the prevalence odds ratio when the period for being at risk of developing the outcome extends over a considerable time (months to years) as it does in this example:

PR = (a/N1) / (c/N0)

PR= (50/250) / (50/750) = 3.0

In this case, a prevalence ratio of 3.0 can be interpreted to mean that the proportion of people with CHD is 3-fold greater if a person is not physically active.

POR vs. PR

For chronic disease studies or studies of long-lasting risk factors, POR is the preferred measure of association in cross-sectional studies. For acute disease studies, PR is the preferred measure of association. If the prevalence of disease is low, i.e. 10% or less in exposed and unexposed populations, POR = PR. Since cross-sectional studies are particularly useful for investigating chronic diseases (e.g. prevalence of AIDS) where the onset of disease is difficult to determine, or for studying long lasting risk factors (such as smoking, hypertension, and high fat diets), the prevalence odds ratio will generally be the preferred measure of association.

Limitations of cross-sectional studies to evaluate risk

Recall that, under steady conditions, the prevalence of disease is influenced both by incidence and duration of disease (or survival with disease).

Prevalence = Rate x Average Duration of Disease

Persons who survive longer with a disease will have a higher probability of being counted in the numerator of a prevalence proportion. Short-term survivors will be less likely to be counted as a case. Incidence is influenced only by exposure, whereas prevalence is influenced both by exposure and duration of disease.

If exposure influences survival time, then the POR or PR will not provide a valid estimate of the risk ratio or rate ratio. Thus, the interpretation of the POR or PR is subject to survival bias.

Even if incidence remains constant, either an improvement in disease treatment (that results in higher cure rates) or increased lethality (resulting in a higher case fatality rate) will result in decreased prevalence. The disease itself or the threat of developing the disease may cause outmigration of cases from an environment perceived as causing disease, e.g. workers affected by toxic exposures in a plant may quit, while more resistant workers will stay. This selective migration can bias measures of prevalence.

Other problems with interpretation of cross-sectional studies

Cross-sectional studies as well as case-control studies are affected by the *antecedent-consequent bias*, similar to the chicken and egg question (i.e. "which came first?"). This bias occurs when it cannot be determined that exposure preceded disease, since both are ascertained at the same time (unlike cohort studies or clinical trials). Antecedent-consequent bias does not affect cohort studies because subjects in cohort studies are selected for study because they are disease-free. Exposure is actually observed to precede disease only in a cohort design, including randomized trials.

Uses of cross-sectional studies

Descriptive studies are an important method to evaluate the proportion of a population with disease or with risk factors for disease, such as the prevalence of asthma in children or the prevalence of elevated blood lead in toddlers.

Descriptive cross-sectional studies are widely used to estimate the occurrence of risk factors in segments of the population characterized by age, sex, race or socioeconomic status (SES). National examples of cross-sectional studies of great importance are the *decennial census* and the National Health and Nutrition Surveys (NHANES). Opinion polls and political polls are basically cross-sectional studies. Surveillance of changes in smoking habits or of other behavioral risk factors are sequential cross-sectional studies. The US National Health and Nutrition Examination Survey (NHANES) is one such example Similarly, surveillance of long lasting diseases such as AIDS is cross-sectional. Descriptive cross-sectional studies are useful for planning or administering preventive or health care services, surveillance programs, and surveys and polls.

Descriptive/analytical cross-sectional studies are useful for establishing preliminary evidence for a causal relationship. These studies are also useful for examining the association between exposure and disease onset for chronic diseases where researchers lack information on time of onset. Examples might include diet and arthritis, smoking and chronic bronchitis, and asthma and exposure to air pollution. Interpretation requires caution regarding potential association of duration of disease with exposure status *(survival bias)*.

Survival bias may be minimized if information can be obtained on exposures that clearly preceded the first symptoms of a chronic disease such as arthritis, diabetes, or chronic bronchitis. This depends on access to medical records before the onset of a chronic disease. In addition, it may be necessary to have historical records on an individual's exposure status prior to these first medical visits, e.g. where the person lived or where the person was employed.

Terminology

Antecedent-consequent bias: occurs in cross-sectional studies when it cannot be determined if exposure preceded disease.

Prevalence: the proportion of diseased individuals in a population.

Survival bias: occurs in cross-sectional studies when the exposure influences survival time, and the distribution of that exposure will be distorted among a sample of survivors. (a.k.a. Neyman bias, incidence-prevalence bias, or selective survival bias)

Resources:

Delgado-Rodriguez M and Llorca J. J Epidemiol Community Health. 2004;58:635–641.

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, Gayle Shimokura, MSPH 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.