

ERIC Notebook

July 1999

Issue 3



Michel Ibrahim, MD, PhD
Director of Education
Program

Lorraine Alexander, Dr PH

Carl Shy, MD, DrPH

Sherry Farr, Graduate
Research Assistant



Ron Horner, PhD
ERIC Director

The ERIC Notebook is funded 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.

Cohort Studies

A cohort study is a type of epidemiological study in which a large group of people with a common characteristic is followed over time to find how many reach a certain health outcome of interest (disease, death, or a change in health status or behavior). A cohort is defined as a group of persons, usually 100 or more in size, who share a common characteristic, e.g. smokers, workers in a lead smelter. Cohort studies compare an exposed group of individuals to an unexposed group of individuals to determine if the outcome of interest is associated with exposure. There are two types of cohort studies: prospective and retrospective or historical cohort. Prospective studies follow a cohort into the future for a health outcome, while retrospective studies trace the cohort back in time for exposure information after the outcome has occurred. Both types of cohort studies are also referred to as longitudinal or follow-up studies.

Establishing the cohort. The investigator controls the selection of the cohort. The investigator may choose a cohort based on age, on exposure to a certain working environment, or on some other common characteristic. Cohorts may be selected on the basis of exposures known at baseline, e.g. smokers vs. nonsmokers. Alternatively, cohorts may be divided into exposure categories once baseline measurements of a defined population are made. For example, the Framingham Cardiovascular Disease Study (CVD) used baseline measurements to divide the population into categories of CVD risk factors.

For instance, an investigator wants to study whether exposure to military aircraft engine noise is a risk factor for hearing loss. The cohort this investigator would want to establish should be composed of two groups of military personnel: one exposed to engine aircraft noise (the group under study) and the other unexposed to engine aircraft noise (a comparison group). The unexposed group should be representative of the exposed group on all factors except exposure.

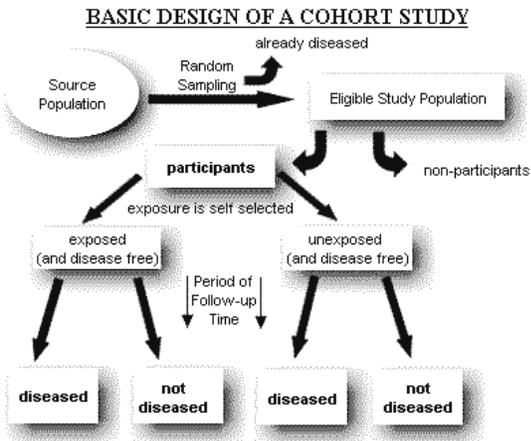
The cohort at baseline. After the cohort of study subjects is established, their individual exposures of interest are identified at baseline (through interviews, bioassays, medical records, etc.). Subjects with the outcome of interest at baseline are excluded. Therefore, all members of the cohort are at risk of developing the outcome at the beginning of observation.

- Following the last example, anyone in the cohort of military personnel with a specified hearing loss at baseline would be excluded from the cohort and would not be followed.

Following the cohort. The cohort is then followed over time for new occurrences of the outcome of interest, in the above example, hearing loss. In a prospective, or concurrent, cohort study baseline exposure is assessed at the beginning of the study and the cohort is followed into the future. In a retrospective, or historical cohort study, baseline exposure is assessed at some point in the past through historical records, e.g. health records for a cohort of factory workers may provide exposure and outcome information up to the present.

Cohort	Baseline exposure	Follow-up
Prospective	Assessed at beginning of study	Followed into the future for outcome
Retrospective	Assessed at some point in the past via historical records	Outcome has already occurred and is assessed via historical records

Cohorts are followed over time to the end of follow-up. Occurrence of the outcome of interest is determined via interviews with members of the cohort and/or family members, or by viewing health and/or work records to conclude the study.



The basic design of a cohort study from beginning of the study to end of follow-up.

Evaluation of the results. During the follow-up period the investigator counts the number of subjects who develop the outcome of interest. This count becomes the numerator for an incidence calculation. The number of persons at risk at baseline becomes the denominator for a **cumulative incidence (CI) calculation**. The CI measures an individual's risk of developing the outcome of interest.

$$\text{Cumulative incidence} = \frac{\text{new occurrences of the outcome}}{\text{population-at-risk at baseline}}$$

Two CIs can be compared to provide a **cumulative incidence ratio (CIR)**, also known as the **relative risk**. The reference group is a comparable unexposed cohort. The index group is the exposed cohort. The CIR is computed by dividing the CI in the index group by the CI in the reference group. The CIR gives a relative measure of the increase or decrease in incidence between the index and reference groups.

$$\text{Cumulative incidence ratio or Relative risk} = \frac{CI_{\text{exposed}}}{CI_{\text{unexposed}}}$$

As with CI, an **incidence rate measure (IR)** uses new occurrences of the outcome as the numerator. However, in an IR calculation the person-time (days, months, or years) at risk during follow-up becomes the denominator. Person-time is measured by summing the total time each member of the cohort was free of the outcome of interest and thus contributed to person-time-at-risk during the follow-up period. The IR measures the rapidity of occurrence of new disease in the population.

$$\text{Incidence rate} = \frac{\text{new occurrences of the outcome}}{\text{person-time at risk}}$$

Two IRs may also be compared to find the relative increase or decrease in the rate of disease occurrence between the exposed and unexposed groups. This relative measure is called the **incidence rate ratio (IRR)**.

$$\text{Incidence rate ratio} = \frac{IR_{\text{exposed}}}{IR_{\text{unexposed}}}$$

Incidence measures between exposed and unexposed cohorts can also be subtracted from one another to find the difference between the two measures. This measure is referred to as the **rate difference**.

$$\text{Rate difference} = I_{\text{exposed}} - I_{\text{unexposed}}$$

Exposure may be a causative risk factor or a preventive factor in the development of the outcome of interest. When exposure is preventive the CIR or IRR, depending on which measure is computed, will be less than one.

Advantages of a cohort study. A cohort study allows a direct estimate of risk (cumulative incidence) and rate of disease occurrence over time (incidence rate). Cohort studies are an efficient means of studying rare exposures (e.g. gasoline fumes, as discussed in next paragraph), in contrast to case-control studies, which tend to be better for rare outcomes. Cohort studies also allow the investigator to assess multiple outcomes of a single exposure.

- A cohort study would be the most efficient means of studying the effects of long-term exposure to gasoline fumes. The cohort would consist of individuals who are exposed daily to gasoline fumes (auto mechanics, gas station attendants, sea crewman on tankers, etc.). By studying this group of individuals, the investigator can better determine the direct effects of long-term, regular gasoline inhalation. Also, by conducting a cohort study, an investigator could determine if gasoline inhalation causes many different health outcomes (e.g., different types of cancer and respiratory illnesses).

Additional advantages of cohort studies. Cohort studies establish temporal relationships between exposure and outcome. Exposure clearly precedes the outcome because the population under study at baseline is free of the outcome of interest. Cohort studies also avoid recall bias (as the exposure is determined before the outcome, one's disease state won't affect how

accurately one recalls exposure levels), as well as, survival bias (duration of disease influencing exposure measurements). Therefore, cohort studies are the best observational design in order to establish cause and effect relationships.

Disadvantages of cohort studies. Cohort studies often require large sample sizes, especially when the outcome is rare, defined as less than 1 event per 1000 person-years (e.g., all specific cancers). Therefore, cohort studies tend to be expensive and time-consuming. When there are losses to follow-up (individuals who leave the cohort before the end of follow-up) biases may occur. Thus, individuals who leave the cohort prematurely may have a different baseline risk than the members who remain in the cohort throughout the entire length of follow-up. Therefore, the study may not be generalizable to the original target population, but only to those who remained under investigation throughout the length of the study. Also, any differences in the quality of measurement of exposure or disease between exposed and non-exposed cohorts may introduce information bias and thereby distort the results.

Self-evaluation

Q1: An investigator wants to discover whether or not being overweight in adolescence increases the risk of cardiovascular mortality in adulthood.

- Assuming historical records are available, would a prospective or retrospective study be more practical?
- Who would comprise the investigator's cohort under study?
- Who would comprise the investigator's exposed and unexposed groups in this cohort?

Q2: An investigator conducts a retrospective cohort study to explore the relationship between perimenopausal exogenous estrogen use and the risk of coronary heart disease (CHD). A total of 5000 exposed and 5000 unexposed women are enrolled and followed for 15 years for the development of myocardial infarction (MI). A total of 200 estrogen users and 300 nonusers had MIs.

- The risk (CI) of a MI among estrogen users is:
- The risk (CI) of a MI among nonusers of estrogen is:
- The relative risk (CIR) for MI is:
- Based on the results of this study is estrogen use a causative or protective factor for MI?

Answers

1.a A retrospective study would be more practical due to the long follow-up time necessary in a prospective study.

b. The cohort under study should be comprised of adolescents of known height and weight (to obtain a measure of body mass index) all of whom lived in the same geographic area during the 1950s and who were within a certain number of years of age from each other. These data might be obtained from historical high school records or similar sources.

c. The exposed group would be comprised of adolescents whose body mass index was 25% or more above the average, while the unexposed group would be comprised of adolescents within 10% of average body mass index, all of whom lived in the same geographic area during adolescence and are within the same age range. The cohort of exposed and unexposed adolescents would be followed over historical time to determine whether the mortality rate from cardiovascular disease was greater in the exposed than in the unexposed adolescents.

2. a. Risk in exposed = $CI_{\text{exposed}} = 200 / 5000 = 0.04$ cases per person, or 4 cases per 100 persons.

b. Risk in unexposed = $CI_{\text{unexposed}} = 300 / 5000 = 0.06$ cases per person, or 6 cases per 100 persons.

c. Relative risk = $CIR = CI_{\text{exposed}} / CI_{\text{unexposed}} = 0.04 / 0.06 = 0.67$. Therefore, the risk of MI among perimenopausal users of estrogen is two-thirds of the risk among nonusers.

d. In this study, estrogen use is a protective factor in the development of MI. Women who use estrogen have an absolute decrease of 0.02 cases per person in risk ($RD = 0.06 - 0.04$). The exposed cohort's risk of MI is 33% less than that of nonusers (1.0 vs. 0.67).

Special Announcements

Internet Epidemiology Course for Fall 1999

This 16-week course, taught entirely on the internet is designed to require problem-based learning of basic epidemiological concepts and methods. A student who successfully completes this course will receive 3 credit hours from the University of North Carolina at Chapel Hill. Registration is now open! For more information, please visit our website at http://cdlhc.sph.unc.edu/dl_courses/epid160ERIC/index.html (Copy and paste this URL into the location box of your browser)

Missing an ERIC Notebook?

If you want to get past issues of the ERIC Notebook, please visit our website at: http://cdlhc.sph.unc.edu/dl_courses/epid160ERIC/index.html We have a complete archive of downloadable Adobe Acrobat © files of past issues of the ERIC Notebook.

If you would like to receive ERIC Notebook please fill out the form below:

Name: _____

Degree(s): _____

Address: _____

City, State, Zip: _____

Telephone Number: _____

Fax Number: _____

E-mail Address: _____

Please fax to: 919-966-2089 – Attn: Angie Tuck **Or**

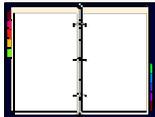
Mail to : UNC-CH , School of Public Health , Attn: Angie Tuck
Department of Epidemiology , Campus Box 7400, Chapel Hill, NC
27599

Upcoming Topics

Incidence Measures in Cohort Studies

Assessment of Diagnostic and Screening Tests

Please let Angie Tuck know which topics are of special interest to you so that we can include them in a future issue.



ERIC Notebook

UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL
SCHOOL OF PUBLIC HEALTH
DEPARTMENT OF EPIDEMIOLOGY
CAMPUS BOX 7400
CHAPEL HILL, NC 27599