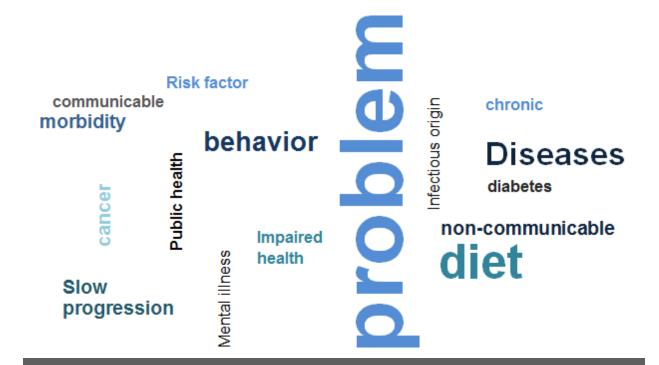
PARTICIPANT GUIDE



Descriptive and Analytic Studies

Created: 2013





Table of Contents

ESCRIPTIVE AND ANALYTIC STUDIES	
LEARNING OBJECTIVES	3
ESTIMATED COMPLETION TIME	
REFERENCES AND RESOURCES	
Module Content	
SKILL ASSESSMENT	30

Descriptive and Analytic Studies

LEARNING OBJECTIVES

At the end of the training, participants will be able to:

- Identify the following for an NCD problem:
 - Type of study to conduct
 - o Sampling methods to use
 - o Measure of association to calculate for a particular study
- Interpret the results of descriptive and analytic studies.

ESTIMATED COMPLETION TIME

• 5 hours, 30 minutes (4 hours for interactive lecture; 1 hour, 30 minutes for Skill Assessment).

REFERENCES AND RESOURCES

- Aschengrau A, Seage GR. Essentials of Epidemiology in Public Health, 2nd edition. Sudbury, Massachusetts: Jones and Bartlett Publishers; 2008.
- Aschengrau A, Seage GR. Essentials of Epidemiology in Public Health, 2nd edition. Case-control Studies PowerPoint™ Slides: http://publichealth.jbpub.com/aschengrau/ppts/case-control%20studies.ppt accessed on February 22, 2011
- Gordis, L. Epidemiology, 2nd edition. Philadelphia, PA: W.B. Saunders Company; 2000.
- Herold JM and Peavy JV. Surveys and Sampling. Field Epidemiology,
 2nd ed. Ed. Gregg M. New York: Oxford University Press, 2002.
- Oleckno WA. Essential epidemiology: principles and applications.
 Prospect Heights, IL 2002;108.
- Remington RP, Brownson RC, Wegner MV, ed. Chronic Disease Epidemiology and Control. 3rd ed. Washington DC: American Public Health Association; 2010
- Rothman K.J., Greenland S. Modern Epidemiology, Second edition, Philadelphia, PA, 1998.
- Stehr-Green, J and Stehr-Green P, Survey Design Part 1: Sampling.
 North Carolina Center for Public Health Preparedness Training Website.
 Accessed on January 31, 2011 at
 http://cphp.sph.unc.edu/training/HEP_SDP1/certificate.php
- Stöckl H, Watts C, Kilonzo Mbwambo JK. Physical violence by a partner during pregnancy in Tanzania: prevalence and risk factors. Reprod Health Matters. 2010 Nov;18(36):171-80.

DESCRIPTIVE AND ANALYTIC STUDIES

- Stern, F, Halpern, W, Hornung, R, Ringenburg, V and McCammon, C. Heart Disease Mortality Among Bridge and Tunnel Officers Exposed to Carbon Monoxide. American Journal of Epidemiology. 1988;128:1276-1288.
- Unwin N, James P, McLarty D, Machybia H, Nkulila P, Tamin B, Nguluma M, McNally R. Rural to urban migration and changes in cardiovascular risk factors in Tanzania: a prospective cohort study. BMC Public Health. 2010 May 24;10:272.



Learning Objectives

- · Identify the following for an NCD problem:
 - · Type of study to conduct
 - · Sampling methods to use
 - Measure of association to calculate for a particular study
- Interpret the results of descriptive and analytic studies.

Descriptive and Analytic Studies

2



Lesson Overview

- · Reasons for conducting studies
- · Definition, characteristics, and analysis of:
 - · Descriptive studies
 - · Analytic studies
- · Methods of sampling

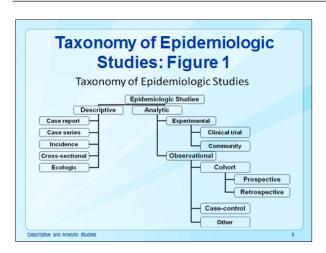
Descriptive and Analytic Studies

3

Why Conduct Studies?

To describe burden of disease or prevalence of risk factors, health behaviors, or other characteristics of a population that influences risk of disease

- To determine causes or risk factors for illness
- To determine relative effectiveness of interventions





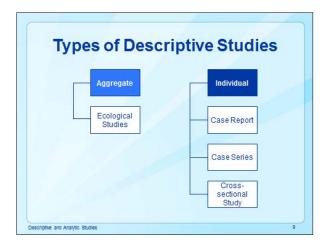


Descriptive Studies

Characterize <u>who</u>, <u>where</u>, or <u>when</u> in relation to <u>what</u> (outcome)

- Person: characteristics (age, sex, occupation) of the individuals affected by the outcome
- Place: geography (residence, work, hospital) of the affected individuals
- Time: when events (diagnosis, reporting; testing) occurred

Descriptive and Analytic Studies



Cross-Sectional Study as a Descriptive Study

Purpose: To learn about the characteristics of a population at one point in time (like a photo "snap shot")

Design: No comparison group



Population: All members of a small, defined group or a sample from a large group

Results: Produces estimates of the prevalence of the population characteristic of interest

When to Conduct a Cross-Sectional Study

- To estimate prevalence of a health condition or prevalence of a behavior, risk factor, or potential for disease
- To learn about characteristics such as knowledge, attitude and practices of individuals in a population
- To monitor trends over time with serial crosssectional studies

Descriptive and Analytic Studies

Cross-Sectional Study Measures

Prevalence of a condition:

= number of existing cases / size of population



(or population count)

Descriptive and Analytic Studies

Example: Cross-Sectional Study

Objective

To estimate the magnitude and patterns of violence against pregnant women

Study

 Population-based, household, cross-sectional study in Mbeya and Dar es Salaam, Tanzania, 2001-2002

Result

 Violence experienced by 7% in Dar es Salaam and 12% in Mbeya

Ref. Stöcki H, Watts C, Kilonzo Mowambo JK. Physical violence by a partner during pregnancy in Tanzania: prevalence and risk factors. Reprod Health Matters. 2010 Nov;18(35):171-80.

Studies to Track Trends in Newly Recognized Cases

Incidence study

- Newly reported or registered disease cases compared over time, place, or person
- Population estimates or other population group totals used as denominators

Ecological study

 Rates are linked to the level of exposure to some agent for the group as a whole

Descriptive and Analytic Studies

14

Example: Incidence Study

Objective

 To estimate the incidence and prevalence of diabetes in young persons in the United States

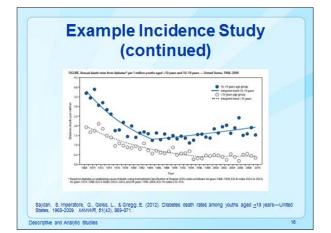
Study

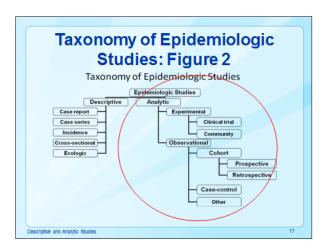
 Annual diabetes death rates among youth aged ≤19 calculated from National Vital Statistics System data from 1968-2009

Result

· Trends for diabetes death rates varied by age group

Saydah, S. Imperatore, G., Gelss, L., & Gregg, E. (2012). Diabetes death rates among youths aged \leq 19 years—United States, 1963-2009. MMMVR, 61(43), 869-871





Analytic Studies Definition

Analytic studies test hypotheses about exposureoutcome relationships

- Measure the association between exposure and outcome
- · Include a comparison group

Descriptive and Analytic Studies

Developing Hypotheses

- A hypothesis is an educated guess about an association that is testable in a scientific investigation.
- Descriptive data (Who? What? Where? When?) provide information to develop hypotheses.
- Hypotheses tend to be broad initially and are then refined to have a narrower focus.

Developing Hypotheses Example

Hypothesis: People who smoke shisha are more likely to get lung cancer than people who do not smoke shisha.

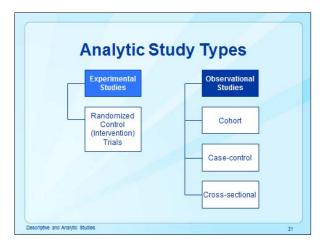
Exposure: smoking shishaOutcome: lung cancer

Hypothesis: ?

• Exposure: ?

· Outcome:?

Descriptive and Analytic Studies



Cohort Studies

What is a cohort?

A well-defined group of individuals who share a common characteristic or experience

• Example: Individuals born in the same year

What are other examples of cohorts?

Cohort Study

(longitudinal study, follow-up study)

- Participants classified according to exposure status and followed-up over time to ascertain outcome
- Can be used to find multiple outcomes from a single exposure
- Appropriate for rare exposures or defined cohorts
- Ensures temporality (exposure occurs before observed outcome)

Descriptive and Analytic Studies

Descriptive and Analytic Studies

Cohort Study Design

Disease

Exposed

No Disease

Exposure is self-selected over time

Disease

Unexposed

No Disease

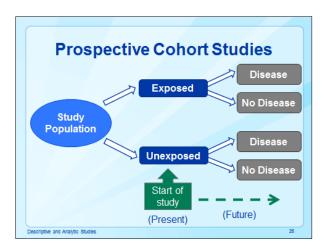
Types of Cohort Studies

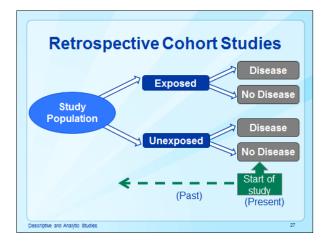
Prospective cohort studies

 Group participants according to past or current exposure and follow-up into the future to determine if outcome occurs

Retrospective cohort studies

 At the time that the study is conducted, potential exposure and outcomes have already occurred in the past





When to Conduct a Cohort Study

When the exposure is rare and the outcome is common

· Agricultural pesticide use and cancer events

To learn about multiple outcomes due to a single exposure

 Health effects of a nuclear power plant accident

Analysis of Cohort Studies

Risk

Quantifies probability of experiencing the outcome of interest in a given population

 Calculation: Number of new occurrences of outcome/population at risk

Example:

- · 29 new cases of diabetes in a community
- 100,000 people in the community at risk for diabetes
- · What is the risk of diabetes? 29/100,000

Descriptive and Analytic Studies

29

Analysis of Cohort Studies: Person-Time, Rate

Quantifies occurrence of outcome in population by time

Calculation:

number of new cases during follow-up period

Sum of time each study participant was followed and at risk of disease

Example: 1,212 tunnel workers

160 deaths among tunnel workers

24,035 person-years at risk

Mortality rate = 160 / 24,035

= 6.7 deaths per 1,000 workers per year

Ref. Stem et al. Heart Disease Mortality Among Bridge and Tunnel Officers Exposed to Carbon Monoxide. American Journal of Epidemiology;1988;128:1276-1288

Descriptive and Analytic Studies

Risk Ratio

- · Can also be called Relative Risk or RR
- Quantifies a population's risk of disease from a particular exposure
- Calculation:

Risk in the exposed group / Risk in the unexposed group

Descriptive and Analytic Studies

31

Rate Ratio

Compares the rates of disease in two groups that differ by demographic characteristics or exposure history

Calculation:

Rate for group of primary interest

Rate for comparison group

Descriptive and Analytic Studies

RR Strength Scales

RR	Strength	RR
0.71 - 0.99	Weak	1.01 - 1.50
0.41 - 0.70	Moderate	1.51 - 3.00
0.00 - 0.40	Very strong	>3.00

Oleckno WA, Essential epidemiology: principles and applications. Prospect Heights, IL 2002;108.

Descriptive and Analytic Studies

Example: Risk Ratio

Question: What is the relationship between being obese and getting type 2 diabetes?

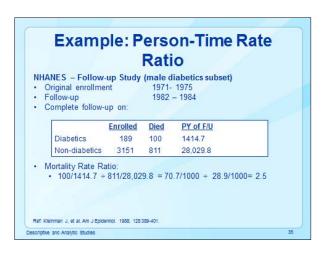
 $\frac{\text{Risk in the exposed group (obese)}}{\text{Risk in the unexposed group (non-obese)}} = \frac{.00076}{.00013} = 5.8$

Risk Ratio = 5.8

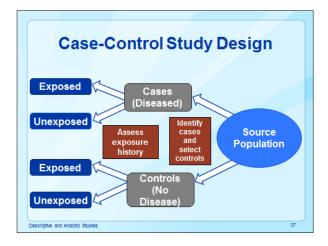
Interpretation: The risk of diabetes among those who are obese is 5.8 times the risk among those who are not obese.

Descriptive and Analytic Studies

34







When to Conduct a Case-Control Study

- · The outcome of interest is rare
- Multiple exposures may be associated with a single outcome
- · Funding or time is limited

Descriptive and Analytic Studies

Case-Control Study: Analysis Format

Exposure	Cases	Controls	
Yes	а	b	
No	С	d	

Exposure odds ratio (OR) ≈ RR when disease is rare

Odds of being exposed among the cases = a/c
Odds of being exposed among the controls = b/d

Exposure odds ratio = (a/c)/(b/d) = (a*d)/(b*c)(Cross-product ratio)

Descriptive and Analytic Studies

Example Odds Ratio

Lead Poisoning

Work in mine?	Cases	Controls	
Yes	17	13	
No	83	87	

Odds Ratio = 17/83 ÷ 13/87 = 17x87 / 13x83=1.37

Prevalence Ratio and Prevalence Odds Ratio

- · Chronic disease date of onset is unknown
- · Measure prevalence rather than incidence

RR —

PR (prevalence ratio)

POR (prevalence odds ratio)

Descriptive and Analytic Studies

Prevalence Ratio

- · Usually from a cross-sectional study
- · Similar to risk ratio from cohort study

Exposure	With disease	Without disease	Total
Exposed	а	b	a+b
Unexposed	С	d	c+d
Total	a+c	b+d	

- PR= Prevalence of disease in exposed group/ Prevalence of disease in unexposed group
 - OR
- PR= a/(a+b) / c/(c+d)

Descriptive and Analytic Studies

Prevalence Odds Ratio

- Usually from a cross-sectional study
- · Similar to odds ratio from case control study
- · Calculated same way as odds ratio:

 $POR = \underline{a*d}$ c*b

	disease	disease	
Exposed	а	b	a+b
Unexposed	С	d	c+d
	a+c	b+d	

Example: Prevalence Ratio and Prevalence Odds Ratio

Prevalence of Breast Cysts

Lifetime use of oral contraceptives	Yes Cyst	No Cyst	Total
Ever Used	124	3123	3247
Never Used	77	2557	2644
Total	201	5690	5891

Prevalence of breast cysts among ever users = 124/3247 = .038

Prevalence of breast cysts among never-users = 77/2644 = .029

Prevalence ratio = .038/.029 = 1.3

Prevalence odds ratio = <u>124 * 2557</u> 3123 * 77

= 1.3

- 1.

Practice Exercise #1

Background:

Descriptive and Analytic Studies

- NCDs such as type 2 diabetes are poorly understood and under-prioritized in many low-to-middle income countries.
- You want to determine the risk of type 2 diabetes associated with cardiovascular risk factors such as obesity and abdominal fat mass in your country.

Questions:

- 1. What type of study would you conduct and why?
- 2. What is the measure of association to calculate for this study?

Descriptive and Analytic Studies

45

Practice Exercise #2

Background:

- The prevalence of prostate cancer has increased in your country over the last 5 years.
- You want to examine the association between calcium intake and prostate cancer risk.
- · You have limited time and funding to conduct this study.

Questions

- 1. What type of study would you conduct and why?
- What is the measure of association to calculate for this study?

Descriptive and Analytic Studies

46

Practice Exercise #3

Background:

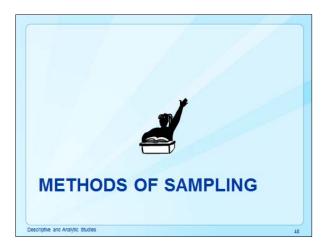
- Cardiovascular disease (CVD) is of growing concern; however your country has no recent data on the burden of this disease.
- You want to estimate the burden of cardiovascular disease in the two main cities in your country.

Questions:

- 1. What type of study would you conduct and why?
- 2. What is the measure of association to calculate for this study?

Descriptive and Analytic Studies

47



Discussion Question

Why do we use sampling?

- Cannot get information on everyone in a population
- Efficiently gets information on a large population
- · Obtains a representative sample of a population

Sampling Methods Two main types of sampling methods: Probability sampling Non-probability sampling

Probability Sampling

What are types of probability-based samples?

- · Simple random sampling
- · Systematic random sampling
- · Stratified random sampling
- · Cluster sampling



Descriptive and Analytic Studies

Descriptive and Analytic Studies

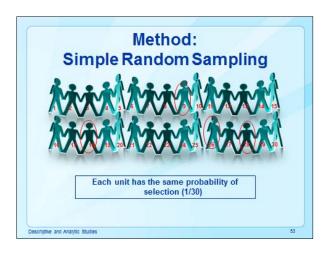
Simple Random Sample

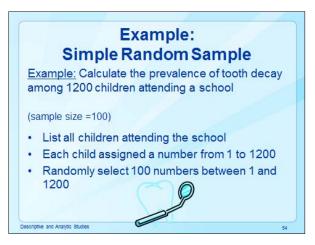
Principle

 Equal chance/probability of drawing each unit

Procedure

- · List all units (persons) in a population
- · Assign a number to each unit
- · Randomly select units





Advantages & Disadvantages: Simple Random Sample

Advantages

Simple

Disadvantages

- · Need complete list of units
- Units may be scattered and poorly accessible

Systematic Random Sample

Principle

 Select sample at regular intervals based on sampling fraction

Procedure

- · List all units (persons) in a population
- · Assign a number to each unit
- Calculate sampling fraction (population size + sample size)
- Select first unit at random based on sampling fraction
- · Subsequent units are chosen at equal intervals

Descriptive and Analytic Studies

56

Advantages & Disadvantages: Systematic Random Sample

Advantages

- · Simple
- · Can be implemented easily without software

Disadvantages

· Need complete list of units

Descriptive and Analytic Studies

57

Example: Systematic Random Sample

Example: Calculate the prevalence of tooth decay among 1200 children attending a school (sample size =100)

- · List all children attending the school
- · Randomize the list to avoid bias
- Each child assigned a number from 1 to 1200
- Sampling fraction =1200/100 = 12
- Randomly select a number between 1 and 12
 - Example: 8
- Select every 12th child, starting with child #8
 - Example: 8, 20, 32, 44...

Descriptive and Analytic Studies

58

Stratified Random Sample

Principle

 Select random samples from within homogeneous subgroups (strata)

Procedure

- · List all units (persons) in a population
- Divide the units into groups (called strata)
- Assign a number to each unit within each stratum
- · Select a random sample from each stratum
- · Combine the strata samples to form the full sample

Descriptive and Analytic Studies

Method:
Stratified Random Sample

• Sampling frame divided into groups (age, sex, socioeconomic status)

• Units in each group have the same probability of selection, but probability differs between groups

Probability:1/20

Probability: 1/15

Advantages & Disadvantages: Stratified Random Sample

Advantages

- Can get separate estimates from the whole population <u>and</u> from individual strata (if sample is large enough)
- Precision increased if less variability within strata than between strata

Disadvantages

· Can be difficult to identify strata

Class Discussion Question

What are some examples of strata that you might sample within?

- · Race/ethnicity/tribe/nationality · Smoking status
- Age group
- Occupation
- Gender
- Education
- Geographic location
- Many possibilities!
- Socioeconomic status

Descriptive and Analytic Studies

Example: **Stratified Random Sample**

Example: Calculate the prevalence of tooth decay among 1200 children attending a school, with equal representation of males and females

(sample size =100)

- · List all children attending the school
- · Divide the children into two groups
 - · 540 males and 660 females
- · Assign each child a number
 - Males: 1 to 540
 - · Females 1 to 660
- · Randomly select 50 males and 50 females

Descriptive and Analytic Studies

Cluster Sample

Principle

· Select all units within randomly selected geographic

Procedure

- Divide population into geographic groups (clusters)
- Assign a number to each cluster
- · Randomly select clusters
- · Sample all units within selected clusters OR select a random sample of units within selected clusters

Advantages & Disadvantages: Cluster Sample

Advantages

- · List of sampling units not required
- More efficient for face-to-face interviews when units are dispersed over a large area

Disadvantages

- Loss of precision due to correlation within clusters
- This correlation needs to be taken into account in sample size calculations and analysis ("design effect")

Descriptive and Analytic Studies

Non-probability Sampling

- · Probability of selection is unknown or zero
- Inexpensive
- · Results not generalizable
- · Results often biased

Common types of non-probability sampling:

- · Convenience sampling
- Snowball sampling / Respondent-driven sampling
- · Voluntary sampling

Descriptive and Analytic Studies

Choosing a Sampling Method

Consider:

- · Population to be studied
 - Size/geographic distribution
 - · Availability of list of units
 - · Heterogeneity with respect to variable
- Level of precision required
- · Resources available

Practice Exercise #4

Background: You will choose a sampling method for each of the following studies.

Questions:

What sampling method would you use for:

- The cross-sectional study on CVD described in Practice Exercise #3? Why?
- A one-time survey of citizens' attitudes toward smoking and second-hand smoke in response to proposed legislation to impose a ban on smoking in restaurants. Why?
- 3. Serosurvey of blood lead levels (or urinary arsenic levels) of prisoners entering the nation's largest prison (or pregnant women entering the nation's largest maternity ward) to determine average level of exposure in the population.

Descriptive and Analytic Studies

68



Descriptive vs. Analytic Epidemiology

Descriptive epidemiology:

· Who, What, When, and Where

Analytic epidemiology:

· Why and How

Types of Descriptive and Analytic Studies

Types of descriptive studies

- · Aggregate: Ecological study
- Individual: Case report, case series, crosssectional study

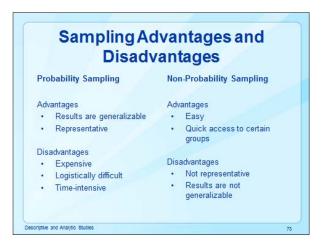
Types of analytic studies

- · Experimental: Randomized control trial
- Observational: Cohort, case-control, crosssectional

Descriptive and Analytic Studies

71

Cohort vs. Case-Control Studies Study Comparison Preferred Study Design When... Members are easily identifying entire cohort would be too costly or time consuming Members are easily accessible Exposure is rare There may be multiple diseases involved Study Group Exposed persons Comparison Group Unexposed persons Descriptive and Analytic Studies Cohort Study Case-Control Study Identifying entire cohort would be too costly or time consuming Accessing entire cohort would be too costly or time consuming Illiness is rare There may be multiple exposures involved Study Group Exposed persons Persons with illness (cases) Comparison Group Descriptive and Analytic Studies





Skill Assessment

- You will work in small groups to complete two parts of a skill assessment:
 - Identify the type of study to conduct and sampling method
 - 2. Interpret the results
- Materials and questions can be found in your Participant Guide.
- Spend approximately 1 hour completing the assessment.
- · Be prepared to share the group's work.

Descriptive and Analytic Studies

74

Instructions:

- 1. You will work in small groups to complete a three-part assessment.
- 2. Select a member of your group to record your responses.
- 3. Groups will have approximately 1 hour to complete the assessment.
- 4. At the end of the assessment, one member from the group will share your summary with the class. (20 minutes)

Part 1. Identify study to conduct and sampling method to use (30 minutes) Read the following three issues of public health concern and answer the questions that follow.

- 1. Cardiovascular disease is an increasing problem in Tanzania. The last study to examine risk factors of cardiovascular disease in Tanzania was in 1987. You want to know about the current status of risk factors for cardiovascular disease and would like to examine how the risk factors have changed since the last study in 1987.
 - a. What type of study would you conduct and why?
 - b. What sampling method would you use?
- 2. In Thailand, breast cancer incidence is increasing, but little is known about the primary risk factors for breast cancer among Thai women. Current understanding of breast cancer risk factors are from studies in high-income countries. It is unclear if identified risk factors in high-income countries are the same in Thailand. You want to learn about the risk factors for breast cancer in Thailand females.
 - a. What type of study would you conduct and why?
 - b. How would you use **simple random sampling** to select 860 controls?
 - c. How would you use systematic random sampling to select 860 controls?
 - d. How would you use stratified random sampling to select 860 controls?
- 3. Non-smoking women in China are disproportionally exposed to environmental tobacco smoke exposure in their homes and at work. You want to know if non-smoking Chinese women exposed to environmental tobacco smoke have increased mortality compared to non-smoking Chinese women who live and work in smoke-free

environments. You'd also like to know if environmental tobacco smoke exposure has an impact on the risk of cardiovascular and cancer deaths among Chinese women.

a. What type of study would you conduct and why?

Part 2. Interpreting the results (30 minutes)

Interpret the results for each study in 1-2 concise sentences. If applicable, explain whether the exposure was associated with an increased or decreased chance of experiencing the outcome.

- 1. You found that the prevalence of hypertension was 41.1% in men and 38.7% in women.
- 2. You found that women with older siblings had a statistically significant odds ratio of 0.3 for breast cancer.
- 3. You found that non-smoking women whose husbands smoked had a statistically significant risk ratio of 1.19 for mortality due to cancer.