

Epidemiological Study Designs, Survey Development & Capitalizing on EHR Data

Department of Medicine – Research Curriculum Series
February 14, 2019

Resa M. Jones, MPH, PhD
resa.jones@temple.edu



Overview

- Study Designs
 - Descriptive
 - Analytical
 - Observational
 - Experimental
- Survey Development
- EHR Data

Questions to ask yourself

- What is the population of interest?
- What is the research question?
 - Estimating population parameters (what is distribution of health-related outcome/disease)
 - Estimating causal effects of exposure on outcome

Descriptive Studies: Purpose

- Describe patterns/distribution of disease occurrence in relation to person, place and time
 - First step in search for determinants or risk factors
 - Hypothesis generating
 - Efficient allocation of resources
 - Targeting of particular populations for education/prevention programs

Analytical Studies: Purpose

- Investigate the determinants of disease
- Search for causes and effects (why and how)
- Test hypothesis about causal associations
- Quantify association between exposure and outcome
- Comparison between groups
- Measures: CI, IR, RR, RD, AR%, PAR, PAR%, OR

Types of Descriptive Studies

- Correlational (Ecological) Studies
- Case Reports
- Case Studies
- **Cross-Sectional (Prevalence) Studies**

Cross-Sectional (Prevalence) Studies

- Exposure and outcome status are assessed simultaneously among **individuals** in well-defined population at one particular time

Cross-Sectional (Prevalence) Studies,

continued

- **Strengths**

- Provides information about frequency and characteristics of a disease by furnishing a “snapshot” of health experience of population at a specified time
- Provides information on prevalence of disease or other health outcomes

Cross-Sectional (Prevalence) Studies,

continued

- **Limitations – Temporality is unknown**
 - Since exposure and outcome status are assessed at single point in time → not possible to determine whether exposure preceded or results from disease/outcome
 - Bias can result as it considers prevalent rather than incident cases → data obtained always reflect determinants of survival as well as etiology

Cross-Sectional (Prevalence) Studies,

continued

- **Exceptions to limitations**
 - Consider as a type of analytic study and can use to test hypotheses IF:
 - Current values of exposure variables unalterable over time (e.g., eye color, blood type, etc.)

Cross-Sectional (Prevalence) Studies,

continued

- When appropriate to use X-sectional
 - Common diseases with a long duration
 - Diseases where determination of onset (incidence) is difficult
 - Examination of quantitative factors over time
- When not to use X-sectional
 - Rare diseases
 - Diseases with short duration

Types of Analytical Studies

- **Observational**
 - Case-control
 - Cohort
- **Experimental**
 - Randomized Trials: Individual and Group
 - Pragmatic Trials

Case-Control Studies

- Best suited to study diseases where medical care is sought
- Suitable for diseases with a relatively short period between first appearance of symptoms and time of diagnosis and interview
- Two types of Case-Control Studies
 - Incidence density
 - Cumulative incidence

Case-Control Studies, continued

- **Conditions that must be met:**
 - Cases need exposure odds representative of exposure odds of all cases in cohort
 - Controls need exposure odds representative of exposure odds of all individuals in cohort
 - Need a conceptual source cohort or study base from which cases and controls are selected
 - Need to measure exposure status as it existed at baseline time point (before disease occurred)

Case-Control Studies, continued

- **Strengths**

- Relatively quick, easy, and less expensive
- Good to identify risk factors for diseases with long latent periods
- Good for rare diseases
- Can examine multiple etiologic factors for a single disease
- Provides odds ratio

Case-Control Studies, continued

- **Most common concerns**

- May lack accurate information on potential risk factor
- May lack accurate information on important confounders
- Cases may search for a cause of disease
- May not be able to determine whether exposure caused the disease or whether the disease caused the person to be exposed
- Identifying and assembling a case group representative of all cases may be difficult
- Identifying and assembling appropriate control group may be difficult
- Participation rates are generally low

Case-Control Studies, continued

- **Limitations**

- Inefficient for rare exposures (unless high odds ratio)
- Cannot directly compute incidence rates (unless study is population-based and sampling fractions are known)
- Temporality can be difficult to establish
- Difficult to validate exposure
- Prone to bias

Cohort Studies

- Analytic design where selection or comparison of subjects is based on exposure status
 - Select subjects according to exposure levels and follow for outcome occurrence
- Incidence of disease/outcome in exposed and unexposed members of cohort is determined and compared

Cohort Studies, continued

- **4 types of Cohort Studies**
 - Prospective
 - Retrospective
 - Ambidirectional
 - Nested Case-Control

Cohort Studies, continued

- **When to use Cohort Study:**
 - Use when clinical trial is unethical, not feasible or too expensive
 - Moderate or large effect expected
 - Little is known about exposure so can evaluate many effects of an exposure
 - Exposure is rare
 - Underlying population is fixed

Cohort Studies, continued

- **Strengths**

- Well-suited for common diseases
- Well-suited for rare exposures in general population
- Evaluate multiple health effects of one or more exposures
- Can directly calculate incidence rates, relative risk, and attributable risk
- Time sequence between exposure and disease more easily established
- Differential misclassification of exposure is unlikely
- If prospective, minimize bias in ascertainment of exposure

Cohort Studies, continued

- **Limitations**

- Not well-suited for rare diseases (unless attributable risk percent is high)
- Often need either many years of follow-up or very large cohorts
- Changes in exposure over time difficult to measure
- Cohort members become increasingly difficult to follow over time
- Expensive and time consuming to conduct (unless retrospective)

Data Collection

- Abstracting records
- Questionnaires (open-ended, close-ended, probes, yes/no, multiple choice)
- Interviews (structured, unstructured)
- Physical examinations
- Biospecimen collection
- Environmental samples (e.g., air, water, dust)
- Tracing (death records, voting lists, phone directories)
- Data capture (centralized, remote, real-time)

Principles of Good Survey Instrument Design...What to Ask About

- 1. Ask people about their first hand experience: what they have done, their current situations, their feelings and perceptions
 - Beware of asking about information that is acquired only secondhand
 - Beware of hypothetical questions
 - Beware of asking about causality
 - Beware of asking respondents about solutions to complex problems

Principles of Good Survey Instrument Design...What to Ask About

Page 12 - Heading

For each of the next questions, please mark the answer that describes you best.

Page 12 - Question 39 - Choice - One Answer (Bullets)

[Mandatory]

Have you ever tried or experimented with cigarette smoking, even a puff?

- Yes
- No [Skip to 14]
- Don't Know

Page 13 - Question 40 - Choice - One Answer (Bullets)

[Mandatory]

Have you ever smoked a whole cigarette?

- Yes
- No (If you click No, please click Submit) [Skip to 14]
- Don't Know

Page 13 - Question 41 - Choice - One Answer (Bullets)

Have you smoked at least 100 cigarettes in your life?

- Yes
- No
- Don't Know

Page 13 - Question 42 - Open Ended - One or More Lines with Prompt

During the past month, on how many days did you smoke cigarettes? If none, put 0 (zero).

Days in the last month

Page 13 - Question 43 - Open Ended - One or More Lines with Prompt

During the past month on the days you smoked cigarettes, how many cigarettes did you smoke each day? If none, put 0 (zero).

Cigarettes each day in the last month

Principles of Good Survey Instrument Design...What to Ask About, continued

- 2. Ask 1 question at a time
 - Avoid asking 2 questions at once
 - Avoid questions that impose unwarranted assumptions
 - Beware of questions that include hidden contingencies

39. Which of the following best describes you? (Mark all that apply.)

- ₁ White
- ₂ African American or Black
- ₃ Asian
- ₄ Native Hawaiian or Pacific Islander
- ₅ American Indian, Alaskan Indian, Alaskan Native
- ₆ Other (Please specify): _____

40. Are you of Hispanic or Latino origin or descent?

- ₁ Yes, I am of Hispanic or Latino origin or descent
- ₂ No, I am not of Hispanic or Latino origin or descent

Principles of Good Survey Instrument Design...Wording Questions

- 3. Question should be worded so every respondent is answering same question
 - To extent possible, words in questions should be chosen so that all respondents understand their meaning, and all respondent have same sense of meaning
 - To extent that words or terms must be used to have meanings that are likely not shared among all, definition should be provided to all respondents
 - Time period referred to by a question should be unambiguous. Questions about feelings or behaviors must refer to timeframe
 - If what is to be covered is too complex to be included in a single question, ask multiple questions

Principles of Good Survey Instrument Design...Wording Questions

The next few questions are about some different tests you may have had to look for signs of colon cancer. Colon cancer is cancer of the colon (large bowel, large intestine) or rectum.

STOOL BLOOD TEST or FECAL OCCULT BLOOD TEST (FOBT)

First, there are two types of stool blood tests also known as fecal occult blood tests (FOBT) that you might have had.

One test **IS DONE AT HOME** using a set of either 2 or 3 cards to determine whether your stool contains blood. You smear a sample of your fecal matter or stool on a card from 2 or 3 separate bowel movements and return the cards to be tested.

3. Have you ever had this stool blood test? (Mark one answer.)

- 1 No, have never had one done
- Yes:
- 2 A year ago or less
- 3 More than 1 but not more than 2 years ago
- 4 More than 2 but not more than 5 years ago
- 5 More than 5 years ago
- 6 Don't know if I have

The other test **IS DONE AT HOME** using 1 tube to determine whether your stool contains blood. You take small samples of your fecal matter or stool with a stick and place it in a tube and return the tube to be tested.

4. Have you ever had this stool blood test? (Mark one answer.)

- 1 No, have never had one done
- Yes:
- 2 A year ago or less
- 3 More than 1 but not more than 2 years ago
- 4 More than 2 but not more than 5 years ago
- 5 More than 5 years ago
- 6 Don't know if I have

Principles of Good Survey Instrument Design...Wording Questions, continued

- 4. If a survey is to be interviewer administered, wording of questions must constitute a complete and adequate script such that, when interviewers read the question as worded, respondents will be fully prepared to answer question
 - If definitions are to be given, they should be given before the question itself is asked
 - A question should end with the question itself. If there are response alternatives, they should constitute the final part of the question

Principles of Good Survey Instrument Design...Wording Questions, continued

The next two questions are about sigmoidoscopy and colonoscopy. These are two other tests to check for colon cancer. Both tests examine the colon using a narrow, lighted tube that is inserted in the rectum.

SIGMOIDOSCOPY

A sigmoidoscopy (pronounced: sig-MOY-DAHS-kuh-pee) is also referred to as flexible sigmoidoscopy or "flex sig." Sigmoidoscopy examines only the lower part of the colon. You are awake during the test, can drive yourself home, and can resume normal activities after the test.

5. Have you ever had a sigmoidoscopy? (Mark one answer.)
- 1 No, have never had one done
 - Yes:
 - 2 A year ago or less
 - 3 More than 1 but not more than 5 years ago
 - 4 More than 5 but not more than 10 years ago
 - 5 More than 10 years ago
 - 6 Don't know if I have

COLONOSCOPY

A colonoscopy (pronounced: koh-luh-NAHS-kuh-pee) is a test that uses a narrow, lighted tube to examine the entire colon. With a colonoscopy, you are given medicine through a needle in your arm to make you sleepy, you need someone to drive you home, and you may need to take the rest of the day off from normal activities

6. Have you ever had a colonoscopy? (Mark one answer.)
- 1 No, have never had one done
 - Yes:
 - 2 A year ago or less
 - 3 More than 1 but not more than 5 years ago
 - 4 More than 5 but not more than 10 years ago
 - 5 More than 10 years ago
 - 6 Don't know if I have

Principles of Good Survey Instrument Design...Wording Questions, continued

- 5. Clearly communicate to all respondents the kind of answer that constitutes an adequate answer to the question
 - Specify the number of responses to be given to questions for which more than one answer is possible

3. In terms of making decisions about your health care with your doctor, which **ONE** of the following best describes how you would like to make these decisions? (**Mark one answer.**)

I prefer to...

- Make the final selection about what I will receive
- Make the final selection after seriously considering my doctor's opinion
- Have my doctor and I share responsibility for deciding what is best
- Have my doctor make the final decision but consider my opinion
- Leave all decisions to my doctor

Principles of Good Survey Instrument Design...Formatting Survey Instruments

- 6. Design survey instruments to make the tasks of reading questions, following instructions, and recording answers as easy as possible for interviewers and respondents

Please fill in the circle that matches your answer with a black pen. Answer each question to the best of your ability. If you are unsure about how to answer a question, please give the best answer you can.

Marking Instructions:

Correct:



Incorrect:



1. Compared to other people your age, would you say that in general your health is:
(Mark one answer.)

¹○.....²○.....³○.....⁴○.....⁵○
Poor Fair Good Very Good Excellent

Principles of Good Survey Instrument Design...Formatting Survey Instruments

9. How much would you agree or disagree that the following reasons would make it difficult for you to have a colonoscopy?

As a reminder, with the colonoscopy: (pronounced: koh-luh-NAHS-kuh-pee)

- The entire colon is examined.
- You are given medicine through a needle in your arm to make you sleepy.
- You need someone to drive you home.
- You may need to take the rest of the day off from your normal activities.

It would be difficult to have a colonoscopy...

	Strongly Agree	Somewhat Agree	Neither	Somewhat Disagree	Strongly Disagree
a. ...because I am worried about what this test might find	1 <input type="radio"/>	2 <input type="radio"/>	3 <input type="radio"/>	4 <input type="radio"/>	5 <input type="radio"/>
b. ...because I am afraid of having this test	1 <input type="radio"/>	2 <input type="radio"/>	3 <input type="radio"/>	4 <input type="radio"/>	5 <input type="radio"/>
c. ...because I am worried that the test is uncomfortable or painful	1 <input type="radio"/>	2 <input type="radio"/>	3 <input type="radio"/>	4 <input type="radio"/>	5 <input type="radio"/>
d. ...because I might get injured by this test	1 <input type="radio"/>	2 <input type="radio"/>	3 <input type="radio"/>	4 <input type="radio"/>	5 <input type="radio"/>
e. ...because I do not want a tube inserted into my rectum	1 <input type="radio"/>	2 <input type="radio"/>	3 <input type="radio"/>	4 <input type="radio"/>	5 <input type="radio"/>
f. ...because I do not want to do the preparation (prep) and take laxatives	1 <input type="radio"/>	2 <input type="radio"/>	3 <input type="radio"/>	4 <input type="radio"/>	5 <input type="radio"/>
g. ...because I do not want to have anesthesia or be "put under"	1 <input type="radio"/>	2 <input type="radio"/>	3 <input type="radio"/>	4 <input type="radio"/>	5 <input type="radio"/>
h. ...because this test is too embarrassing	1 <input type="radio"/>	2 <input type="radio"/>	3 <input type="radio"/>	4 <input type="radio"/>	5 <input type="radio"/>
i. ...because getting to and from the test is difficult	1 <input type="radio"/>	2 <input type="radio"/>	3 <input type="radio"/>	4 <input type="radio"/>	5 <input type="radio"/>

Designing data forms to facilitate data processing

- Use pre-coded data forms
 - Convert responses to numeric variables
 - Some data easy to convert
 - Truly continuous variables
 - Data collected as numeric scores
 - Simple nominal variables
 - Some questions more difficult to convert
 - Check all vs. check only one
 - Numbers and units indicated separately by codes
 - Open-ended questions
 - Special issues for pre-coded forms
 - Consistency of assigned pre-codes
 - Use of special codes to mean “don’t know”, “refusal” and “missing”
 - Ample space allocated for coding open-ended questions
 - Forms which indicate positioning in data file
 - By physical location
 - By variable name

Principles of Good Survey Instrument Design...

'Training' Respondents

- 7. Measurement will be better to extent that people answering questions are oriented to the task in a consistent way

Improving Response Rates

- Survey auspices/Use of information
- Accuracy of list
- Quality of questionnaire
 - Appropriate
 - Brief, easy to complete
 - Light colored paper, professional design
- Cover letter and reminder postcard
- First class mail and stamps
- Incentives
- Tracking

Measurement

- Several important principles guide development of measurement
 - Relate to theory
 - Specific to the behavior being assessed
 - Relevant to the population among whom they will be used

Measurement Considerations

- Content Validity
 - Measure full range of factors that may influence behavior
- Validity and reliability of measures should be reexamined with each study
 - Cultural/population differences make applying scales without such examination susceptible to error
- Using multiple items for each scale reduces measurement errors and increases probability of including all relative components of each construct

Measurement: Resources

- Look to nationally administered questionnaires:
 - <http://www.cdc.gov/brfss/>
 - <https://www.cdc.gov/healthyouth/data/yrbs/index.htm>
 - <http://www.monitoringthefuture.org/>
 - <http://www.cdc.gov/nchs/nhanes.htm>
 - <http://www.cdc.gov/nchs/nhis.htm>

Measurement: Resources, continued

- Health Behavior Constructs: Theory, Measurement, and Research
 - Provides definitions of major theoretical constructs employed in health behavior research, and information about the best measures of these constructs
 - <http://dccps.cancer.gov/brp/constructs/index.html>
- Health and Psychological Instruments (HAPI)
 - Provides information on measurement instruments in the health fields, psychosocial sciences, organizational behavior, and library and information science
 - <http://web.b.ebscohost.com.libproxy.temple.edu/ehost/search/advanced?vid=0&sid=5c3be63c-67d8-41cd-a69b-567165089527%40pdc-v-sessmgr02>

Quality Assurance & Quality Control

- Maximizing reliability and validity of data collection
 - Maximize reliability
 - Repeatability across observers, across time
 - Based on formal rules and procedures
 - Maximize validity
 - Minimize extraneous variability
 - Consider control for sources of variability

Quality Assurance & Quality Control, continued

- Quality assurance
 - Steps taken to assure quality as you go along
- Quality control
 - Formal measure of the extent to which data are collected reliably

Quality Assurance & Quality Control, continued

- Typical quality assurance steps
 - Written protocols
 - Supervision
 - Training of data collectors
 - Selection
 - Content of training
 - End point training
 - Automated data recording
 - Site visits

Quality Assurance & Quality Control, continued

- Typical quality control
 - Procedures for repeat measures
 - Standards for deciding whether measures are within range
 - Procedures for correction if not within range
 - Quality control steps for laboratory analyses
 - Submission of blind duplicate samples
 - Calibration standards
 - Quality control steps for other measures
 - Repeat measurement
 - Analytic tools comparison
 - Data monitoring committees