



# Planning and conducting a research survey

ANN MOORMANN, PHD, MPH

UNIVERSITY OF MASSACHUSETTS MEDICAL SCHOOL

# Why would I do a survey?

- ▶ To measure or assess the relationship between past exposures and a disease or health outcome.
  - ▶ To create a data record for later analysis
- ▶ As a first step, define your hypothesis based on the research question and then decide which **study design** will be best suitable to answer that question.
- ▶ Is your study observational or experimental?

# Observational: Case Control Study

- ▶ A survey can be used to determine the degree of association between various risk factors and outcomes.
- ▶ **Cases** have the disease outcome of interest when enrolled in study
- ▶ **Controls** do not have the disease of interest when enrolled in study
- ▶ A survey is a data collection tool that provides a retrospective assessment of exposures.
- ▶ Exposures can be harmful or beneficial.

# Benefits and limitations of case-control studies

- ▶ Good for rare diseases that take a long time to develop
- ▶ Relatively inexpensive
- ▶ Quick – since you don't have to wait for disease to occur
  
- ▶ Limitations – recall bias or insufficient records of past exposures
- ▶ Does not prove causality but can provide evidence for a significant association between exposure and outcome.

# Survey type and data source

## Subjective sources

- ▶ Survey types used to find out which questions to ask
  - ▶ Focus group discussion
  - ▶ Key informant interviews
- ▶ Individual Questionnaire to record exposures
  - ▶ Interview study participant (or parent of minor)

## Objective sources

- ▶ Medical record extraction including clinical lab values
- ▶ Blood test (if exposure can persist in the body, such as drug levels to see if patient compliant with a treatment regimen)

# Observational: Cohort study

- ▶ Study participants are selected based on **EXPOSURE** of interest (exposed versus non-exposed)
- ▶ Both groups followed over time to determine how many develop the disease or outcome of interest.
- ▶ Longitudinal study design – length of observation time determined by disease or outcome of interest.
- ▶ Can be prospective or retrospective

# Benefits and limitations of cohort studies

- ▶ Good for common diseases or outcomes that don't take a long time to develop
- ▶ Supports a causal link between exposure and disease outcome (relative risk)
- ▶ Limitations – takes time to complete study (can be years of data collection), more expensive, needs consistent and robust data collection team to prevent operator bias.

# Survey types and data source

## Objective outcome measurement

- ▶ Diagnostic test to confirm disease outcome
- ▶ If exposure time was definitive, able to assess progression of disease (how long it takes from exposure to disease presentation)
- ▶ More than one outcome can be assessed if exposure is suspected to cause multiple diseases

# Observational: Cross-sectional study

- ▶ Measures both exposure and outcome at one time for individuals (snap shot)
- ▶ Determines prevalence of exposure and disease (scope of the problem and suspected causes)
- ▶ Quick and easy to do
- ▶ Limitation - cannot conclude cause-effect relationship between exposure and outcome

# Observational: Ecological study

- ▶ Measures both exposure and outcome at one time for a population (snap shot)
- ▶ Determines prevalence of exposure and disease (scope of the problem within a community and suspected causes)
- ▶ Quick and easy to do (large scale)
- ▶ Limitation - cannot conclude cause-effect relationship between exposure and outcome
- ▶ Beware of ecological fallacy (attributing population level data to individuals)

# Survey type and data source

## Objective sources

- ▶ Population level databases, if available (some examples)
  - ▶ cancer or other disease prevalence registry
  - ▶ pollution records
  - ▶ introduction of a product, extent of use and availability, etc.
  - ▶ Census records

# Experimental: Randomized clinical trial

- ▶ Study participants randomized to control group or experimental group
- ▶ Both groups followed over time to determine how many achieve the outcome of interest
- ▶ Longitudinal study design – length of observation time determined by outcome of interest.
- ▶ Prospective in nature – considered **GOLD STANDARD** as it proves cause-effect relationship

# Benefits and limitations of clinical trials

- ▶ Demonstrates causal link between intervention and disease outcome (i.e. survival analysis for treatment studies or protection from disease for vaccine studies)
- ▶ Avoids confounding and selection bias
- ▶ Limitations – takes time to complete study, most expensive, needs consistent and robust data collection team to prevent operator bias and monitors for severe adverse events.
- ▶ Study participant refusal, crossover, drop outs and lack of compliance

# Survey types and data source

## Objective exposure and outcome measurements

- ▶ Controlled intervention (i.e. drug treatment or vaccine)
- ▶ Robust diagnostic test to confirm disease or outcome
- ▶ Since exposure is pre-determined, able to conclusively assess how long it takes from exposure to achieve outcome
- ▶ Can determine benefit of intervention compared to no intervention on health outcomes
- ▶ Cost-benefit analysis

# Paperless data collection

- ▶ RedCap (<https://en.wikipedia.org/wiki/REDCap>)

Research Electronic Data Capture

Ensures security and protection of health data linked to a study participant

Free – but password protected

Developed by Vanderbilt in 2004

Used by 2400 institutional partners in over 115 countries, with more than 590,000 total end-users employing the software for more than 450,000 ongoing research studies.<sup>[2]</sup>



What types of data do you want to collect  
and how do you make sure it's accurate?

Designing your data collection/survey tool

# Demographic data (protected health information)

- ▶ Name
- ▶ Date of birth
- ▶ Date of survey
- ▶ Sex (different from gender)
- ▶ Home location
- ▶ Name of family members (spouse or parents of minors)

21 variables considered PHI = can this data be used to identify a study participant?

[https://en.wikipedia.org/wiki/Protected\\_health\\_information](https://en.wikipedia.org/wiki/Protected_health_information)

# sources

- ▶ <https://www.ncbi.nlm.nih.gov/books/NBK470342/>