Southeastern National Tuberculosis Center
SHARE • LEARN • CURE

Research Methods: Operational Studies in Tuberculosis

Michael Lauzardo, MD

Principal Investigator, Southeastern National Tuberculosis Center

Assistant Professor, Division of Pulmonary and Critical Care Medicine, University of Florida College of Medicine

Deputy Health Officer for TB – State of Florida

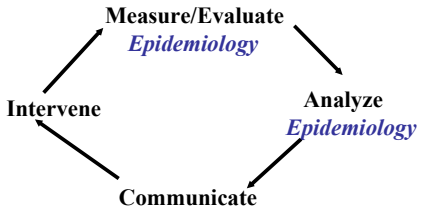
Southeastern National Tuberculosis Center
SHARE • LEARN • CURE

Introduction to Epidemiology

Epidemiology

- Is the process to study the distribution and determinants of disease frequency
- Is a discipline which approaches problems systematically and quantitatively
- Is the basic science of public health

The Public Health Cycle



Epidemiology

- Disease frequency (descriptive):
  - Quantifies the occurrence of disease [How many?]
- Disease distribution (descriptive):
  - Who gets the disease? Where is the disease occurring? When is the disease occurring?
- Disease determinants (analytic):
  - Combination of first two (Why?)

History of Epidemiology

- Original focus was infectious disease epidemics
- Current epidemiology is much broader
  - Acute and chronic diseases
  - Health events
  - Conditions
  - Behaviors

Populations

- Unit of observation is groups rather than individuals
- Clinical observations and basic science *inform* epidemiology, but epidemiology also *informs* clinical practice
  - Understanding disease trends helps clinical providers with their diagnoses

The Epidemiologist's Toolbox

- Experimental Studies
  - Investigators allocate the exposure
- Observational Studies
  - Investigators observe natural course of events

A Typical Epidemiological Approach (1)

- Determine the existence and magnitude of a problem
- Describe who has the problem:
  - Person, place, and time
- Develop hypotheses about why the problem occurs
- Test these hypotheses using appropriate analytic study designs and statistical techniques

A Typical Epidemiological Approach (2)

- Assess validity of any observed association
  - Exclude possible alternative explanations
    - Chance
    - Bias - systematic data collection/interpretation error
    - Confounding - other variables cause the observed association
- Make judgement if a true cause-effect relationship exists between factor and outcome
- Based on the findings, develop interventions
- Evaluate intervention effectiveness

Important Definitions

- *Risk*: the statistical chance of being ill if one is exposed to some factor
- *Exposure*: being in contact with, or having, a factor which may or may not be the cause of illness

[other term used = risk factor]

Southeastern National Tuberculosis Center
SHARE • LEARN • CURE

Ratios, Proportions, and Rates

Ratios

- Express the relationship of two quantities. These quantities may be related or totally independent
  - x / (x+y) or x / y
- Example: A hospital sees 4000 male TB patients and 2000 female TB patients.
  - The ratio of male to female TB patients = 4000 / 2000 = 2 / 1 = 2 to 1

Proportions

- Proportion = a ratio in which the numerator is included in denominator, often expressed as %
  - e.g. 4/100 homeless individuals have TB = 4%
- Prevalence = proportion of individuals in a population who have the disease (event)
  - *Point* prevalence: # with disease / total # in population at a given point in time
  - *Period* prevalence: # with disease any time during a given interval / total # in population at mid-interval

Rates

- A ratio with a distinct relationship between the numerator & denominator, and time is an intrinsic part of the denominator
- Incidence = quantifies the # of new events or cases of disease that develop in a population at risk during a specified time interval
  - Example:
    South Africa had 346 new TB cases per 100,000 population in 2000

Comparison of Incidence and Prevalence

- Prevalence =
  Incidence X Duration of illness
- Prevalence is the *product* of incidence and disease duration

Incidence Rates

- *Cumulative incidence rate* = # new cases during specified time period / total pop at risk at mid-interval
  [note: assumes everyone in study for whole time period]
- *Incidence density rate* = # new cases during specified time period / total person-time at-risk
  [note: accounts for different amounts of time in study]

Other Important Rates

- Morbidity: measure rate of illness/ time
- Mortality: measure rate of death/ time
  - *Crude mortality rate*: the mortality rate from all causes of death for a specified population
  - *Cause-specific mortality rate*: the mortality rate from a specified cause for a specified population
  - *Age-specific mortality rate*: a mortality rate which limits both the numerator and denominator to a particular age or age group

Estimated Annual Incidence of TB in High Burden Countries, 1999

Country	Rate x10^5	Population (thousands)	Cases (thousands)
1. India	185	998,056	1,847
2. China	103	1,266,838	1,300
7. Philippines	314	74,454	234
8. S. Africa	495	39,900	197
10. Viet Nam	189	78,705	149
13. Brazil	70	167,988	118
15. Kenya	417	29,549	123

Examples

All Roads Lead to Ratios

- Proportions, such as PREVALENCE, and Rates, such as INCIDENCE, are both *specific types* of ratios
- RELATIVE RISKS and ODDS RATIOS (to be discussed more later…) are *ratios* of *ratios*

Risk Ratio or Relative Risk

- The ratio of two ratios
- A way to compare risks in those exposed versus those not exposed
  - If 4 in 100 people who are homeless develop TB (exposure=homeless), then risk=4/100 = 0.04 =4%
  - If 1 in 100 who are not homeless develop TB, then risk = 1/100 = 0.01 = 1%
  - Then the risk ratio is 4/100 divided by 1/100 = 4/1
  - Homeless people are 4 times more likely to get TB as non-homeless

Conclusion

- Epidemiology
  - Is a basic tool for public health action
  - Provides data for decision-makers
  - Increasingly emphasizes development and evaluation of control measures
- Application of findings to improve health

Southeastern National Tuberculosis Center
SHARE • LEARN • CURE

Study Design

Questions for Study Design

- What is the purpose?
- What information is needed to intervene?
- When is it needed?
- What kind of study provides the answers?

Study Types

- Descriptive studies
- Analytical studies
  - Experimental
  - Observational

Descriptive Studies

- Describes patterns of disease
  - person, place, time
- Provides data for program planning
- Provides data for resource allocation
- Generates hypotheses

Descriptive Studies: Person (1)

- Question: Who is getting the disease/event?
- Characteristics of person must include age and sex
- Other characteristics
  - high risk behaviors: alcohol use, smoking
  - Homelessness, incarceration, income level, occupation, or others

Descriptive Studies: Person (2)

Death Rates per 100,000 from coronary disease in the US, 1981, by age and sex

Age	Men	Women
0-4	2.2	2.0
5-14	0.9	0.8
15-24	2.6	1.6
25-34	9.4	4.2
35-44	60.6	16.2
45-54	265.6	71.2
55-64	708.7	243.7
65-74	1669.9	769.4
75+	5696.0	4215.1

Descriptive Studies: Person (2)

Death Rates per 100,000 from coronary disease in the US, 1981, by age and sex

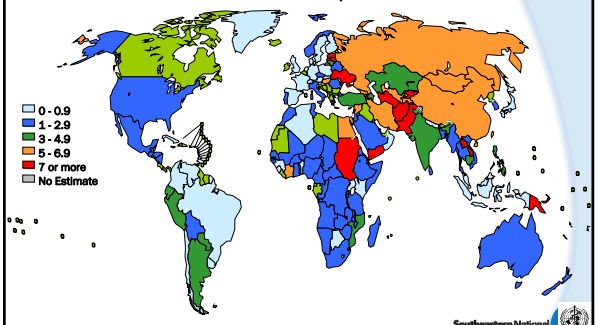
Age	Men	Women
0-4	2.2	2.0
5-14	0.9	0.8
15-24	2.6	1.6
25-34	9.4	4.2
35-44	60.6	16.2
45-54	265.6	71.2
55-64	708.7	243.7
65-74	1669.9	769.4
75+	5696.0	4215.1

Descriptive Studies: Place (1)

- Question: Where are the rates of the disease highest and lowest? Where do resources need to be targeted?
- Geographic characteristics can provide insights into disease etiology
- Geographic comparisons of disease frequency can be made
- Data can be efficiently presented in a pictorial manner

Southeastern National Tuberculosis Center

Estimated Percent of New Cases that are MDR-TB, 2000



Southeastern National Tuberculosis Center

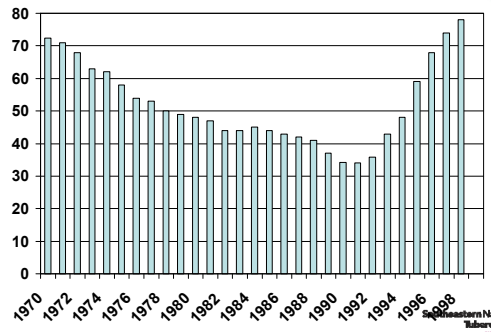
Descriptive Studies: Time (1)

- Question: When does the disease (event) occur? Is the disease frequency different now compared to the past?
- Changes in disease rates over time can signal an epidemic or introduction of the causal agent
- Cyclic changes, such as seasonal patterns, are very valuable

Southeastern National Tuberculosis Center

Tuberculosis Morbidity in Russia

per 100,000 population



Southeastern National Tuberculosis Center

Descriptive Studies: Types

- Correlational studies
- Case Reports and Case Series
- Cross-sectional Surveys

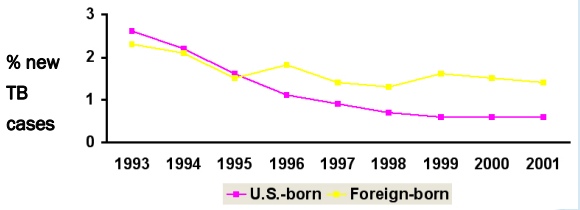
Southeastern National Tuberculosis Center

Descriptive Studies: Correlational (Ecological)

- Determines the relationship of disease and exposure in a population
  - Different groups, same time period
  - Same group, different time period
- Cannot link an exposure to the occurrence of disease in the individual

Southeastern National Tuberculosis Center

Primary MDR TB in US-born versus Foreign-born Persons, 1993-2001



Descriptive Studies: Case Reports and Case Series

- Careful, detailed report(s) by one or more clinicians
- No denominator

Descriptive Studies Cross-sectional Study

- Both exposure and disease outcome is measured simultaneously
- Generally used to determine association between disease and exposure
- Usually cannot establish if exposure came before or after

Summary: Descriptive Studies

- Characterize disease by person, place, time
- Three types: Correlational, case reports or case series, cross-sectional
- Assist in generating hypothesis, but not used to test hypothesis

Study Types

- Descriptive studies
- Analytical studies
  - Observational
  - Experimental or Intervention

Analytic Studies (1)

- Used to test epidemiologic hypotheses
- Identify risk factors
- Compare groups

Analytic Studies (2)

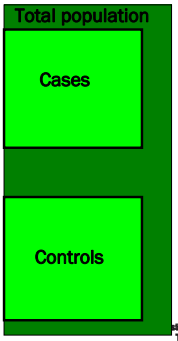
- Observational Studies
  - Investigators observe natural course of events
  - Two types
    - Case-Control Study
    - Cohort Study
- Experimental Studies
  - Investigators control the exposure

Analytic Studies: Case-Control Study

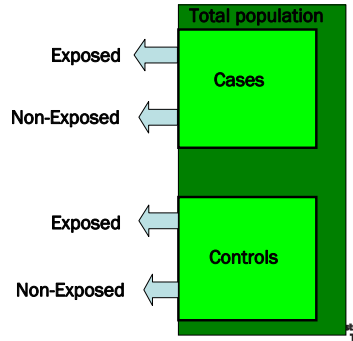
- Subjects are selected on basis of outcome status
  - Case: Subject with outcome of interest
  - Control: Subject without outcome of interest in the same population
- Exposure for two groups is determined

EXPOSURE ⟸ OUTCOME

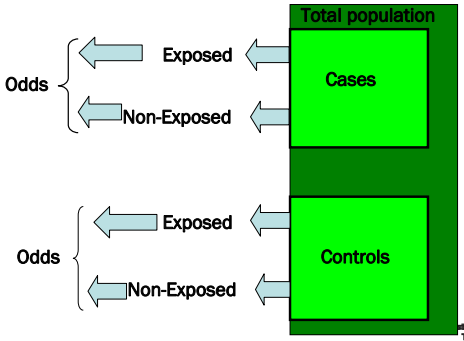
Basic Analytical Approach in a Case-Control Study



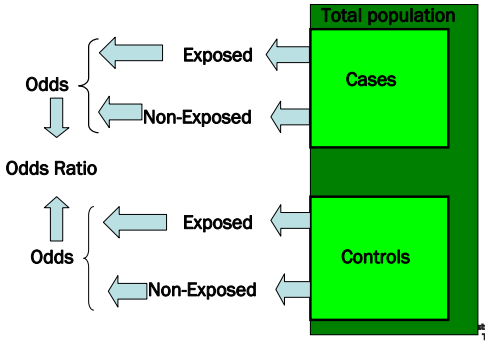
Basic Analytical Approach in a Case-Control Study



Basic Analytical Approach in a Case-Control Study



Basic Analytical Approach in a Case-Control Study



Example: Case-Control Study

- "Smoking and mortality from TB and other diseases in India: retrospective study of 43,000 adult male deaths and 35,000 controls," Gajalakshmi et al
  - Cases: men who died of disease
  - Controls: living men
  - Exposure: Smoking

Advantages and Disadvantages(1): Case-Control

- Advantages
  - Best design for rare outcome
  - Can identify more than one exposure
  - Relatively quick and inexpensive

Advantages and Disadvantages(2): Case-Control

- Disadvantages
  - Not useful for rare exposures
  - Not population-based, cannot calculate incidence rates of disease
  - Time relationship between exposure and disease may be difficult to establish
  - Prone to selection and recall bias
  - Hard to ascertain exposure accurately

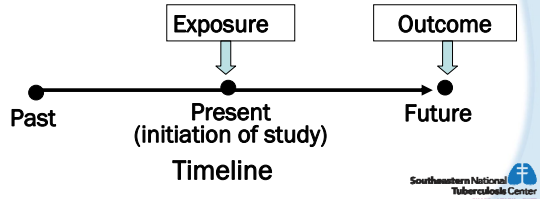
Analytic Studies: Cohort Study

- Subjects are chosen on the basis of the presence or absence of exposure
- Subjects are followed for a specified period of time to determine the development of outcome

EXPOSURE ⇨ OUTCOME

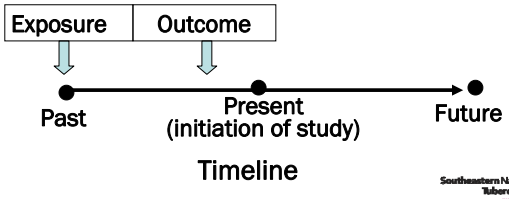
Analytic Studies: Cohort Study Timing

- Prospective: Subjects followed from time of exposure (present) to outcome (future)

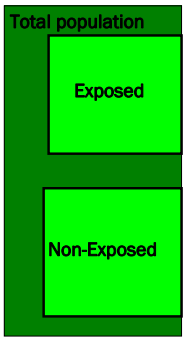


Analytic Studies: Cohort Study Timing

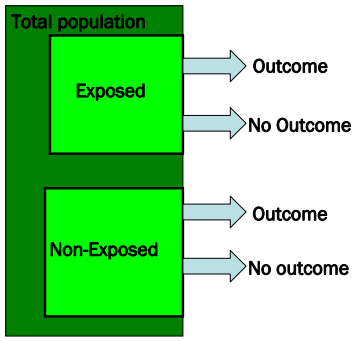
- Retrospective: Exposure and outcome occurred in the past



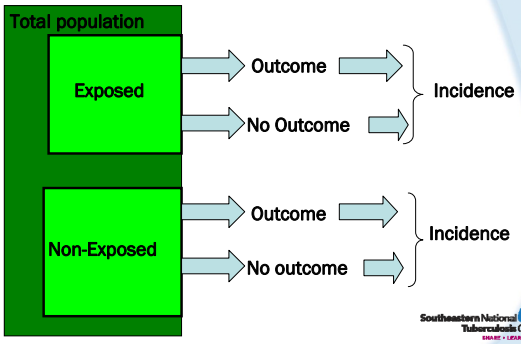
Basic Analytical Approach in a Cohort Study



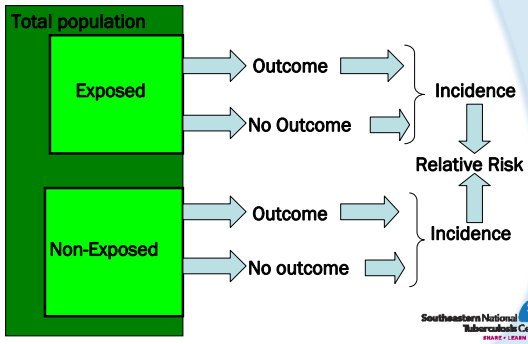
Basic Analytical Approach in a Cohort Study



Basic Analytical Approach in a Cohort Study



Basic Analytical Approach in a Cohort Study



Example: Cohort Study

- "Effect of Highly Active Antiretroviral Therapy (HAART) on incidence of tuberculosis in South Africa: a cohort study," Badri et al
  - Exposed: HIV patients receiving HAART
  - Unexposed: HIV patients without HAART
  - Outcome: active TB

Advantages and Disadvantages: Cohort (1)

- Advantages
  - Best design for studying common diseases or rare exposures
  - Can examine multiple outcomes
  - Can determine time relationship between exposure and outcome
  - Can measure incidence of outcome

Advantages and Disadvantages: Cohort (2)

- Disadvantages
  - Is inefficient for rare diseases
  - Can be expensive and time-consuming
  - Validity of the results can be seriously affected by losses to follow-up
  - Requires the availability of adequate records if retrospective

Analytic Studies Experimental/Intervention Study

- Type of cohort (ex: clinical trial)
- Exposure status is determined by the investigator
- Provides the most reliable evidence from epidemiologic research
  - randomization used to determine exposure status
- Ethical issues
- Expensive

Design Choice

- The choice of study design to use for a particular exposure-outcome relationship depends upon
  - The nature of the outcome under investigation
  - The type of exposure
  - The available resources

Conclusion

- Descriptive studies describe patterns of disease occurrence and allow the formulation of hypotheses
- Analytic studies generally test epidemiologic hypotheses
- For most epidemiologic hypotheses, it is necessary and desirable to employ both descriptive and analytic design strategies

Southeastern National Tuberculosis Center
SHARE • LEARN • CURE

Writing a Research Proposal

Steps in the research process

- Generate a research question
- Organize a team
- Draft a proposal/protocol
- Get TB Program and ethics approval
- Test on a small scale (pilot) and revise proposal
- Do study and analyze data
- Make conclusions for report and manuscript
- Disseminate information and plan action

Why write a proposal?

- To organize your own thinking and plans
- To share your ideas
  - Why, what, and how you want to do something
- To ask for funds
- To obtain ethical clearance

A good proposal…

- Focuses on a clear and specific research question
- Uses a structured format
- Is easy to read and understand
- Is detailed enough so any reader can understand
- Anticipates most questions and problems

A good proposal

- Allows planning
- Helps secure support from supervisors and funders
- Makes doing the activity easy
- Makes doing the analysis easier
- Makes drawing conclusions easier
- Makes writing the report easier

A proposal starts with an important idea

- A problem
- A question
- A hypothesis
  - Cause => Effect

Examples of TB research questions

- Why do patients interrupt TB treatment?
- Why do TB drug shortages occur at district clinics?
- Does treatment failure predict MDR TB?
- What kind of stigma is attached to TB?
- Is clinic-based DOT more cost-effective than home-based DOT?

Components of a proposal

1. Background/Rationale
2. Objectives and hypothesis
3. Methods
4. Ethics/Protection of human subjects
5. Timeline
6. Budget
7. Investigators and responsibilities
8. Results dissemination
9. Appendices

1. BACKGROUND/ RATIONALE

=> Why is it important to do this study?

Background: Components

- General background of issue
- Literature review and previous data
  - Has anyone studied this question before?
  - If yes, are the answers relevant?
- Is it ethical to do it again? (Use of resources)
- Is it essential to do it again? (Justification)
  - What will THIS study contribute?
- What is the funding source?
- What is the intended use of the findings?

Sources of Information

- Peer-reviewed books and articles, official statistics (best)
  - *Lancet, New England Journal of Medicine, JAMA, WHO Bulletin, PanAmerican Journal of Public Health, International Journal of TB and Lung Disease, etc.*
- Health services reports (good)
- Personal communications/anecdotal evidence (fair)

2. OBJECTIVES AND HYPOTHESIS

=> What is the purpose of this study?

Objectives and hypothesis

- Objectives
  - Statement of what you will do, short and clear
  - May be primary and secondary objectives
  - Must be specific and attainable
- Hypothesis:
  - Cause => Effect
  - Concise, clear, specific

Objectives: Examples

- Measure body mass changes in children on TB treatment in district X in 2004
- Identify the barriers to clinic access by TB patients in rural areas in 2004
- Compare the treatment outcomes of male and female patients at health centre Y during 2003
- Determine whether previous TB treatment is a risk factor for anti-TB drug resistance
- Compare the cost to society of home-based and clinic-based DOT in province Z

3. METHODS

=>"How are you going to do the study?"

=>"Who are you studying?"

=>"When will you do the study"

METHODS Components

Design
Intervention
Study population
Study sample
Enrollment procedures
Data collection and variables
Data analysis
Limitations
Pilot study
Training

3.1 Design

- Descriptive study?
- Analytical study? (Asking WHY?)
  - Cross-sectional
  - Case-control
  - Cohort
- Is it testing an intervention?
  - Randomized controlled trial
- Prospective or retrospective?

3.2 Intervention(s)

- Refers to the intervention received by the study participants
- May have been administered in the past (retrospective study)
- Can be more than one per study (describe all)
- One can be the common standard of care, or no care
- Description must be detailed enough

Intervention: Example

"The anti-TB drug regimens used in Country A and referred to in this study, are defined as follows:

- Standardized treatment for MDR-TB: Three months of PZA, EMB, ethionamide, kanamycin, and ciprofloxacin, followed with 15 months of PZA, EMB, ethionamide, and kanamycin. Treatment is administered daily and under direct observation.
- Individualized treatment for MDR-TB: A treatment regimen of at least 18 months duration that includes at least five drugs, to which the organism has shown *in vitro* susceptibility. It is administered daily under direct observation."

3.3 Study Population

- Exactly WHO will be studied
  - Time, place, person
- Inclusion & exclusion criteria
  - Stated exclusions – eg children, recent arrivals, prisoners, extrapulmonary TB

Study Population: Example

"The study population is made up of adult TB patients in Country A who began treatment for MDR-TB between August 1996 and March 2002. These patients are distributed throughout Country A, however a majority of them are residents of the capital.

Inclusion and exclusion: Example

Participants Inclusion Criteria

- Patients who started treatment with second-line drugs between August 1996 and March 2002. Patients with HIV disease, a history of renal insufficiency, hepatitis, or diabetes will be analyzed separately due to these medically complicating factors.

Participant Exclusion Criteria

- Children <= 18 years old, pregnant women, and/or those for whom treatment outcomes cannot be determined with certainty.

Justification of Exclusion

- Establishing a culture-confirmed treatment outcome for children is much more difficult than for adults. As this evaluation relies on knowing treatment outcomes, children have been excluded. Lastly, MDR-TB treatment and MDR-TB treatment outcomes are more complex for pregnant women, and thus pregnant women have been excluded from this evaluation.

Estimated Number of Participants: 2700

3.4 Case Definition

- Criteria for classifying subjects as cases or controls
- Does not need to be a TB case (depends on the study question)

Case Definition: Example

- For the purposes of this study, a patient who received MDR-TB treatment is a patient who was enrolled in a second-line drug regimen for at least 1 day.

3.5 Study Sample

- Describe the sampling frame
  - E.g.: clinic's patient registry
- Describe the sampling method
  - Random, systematic, cluster, etc.
- Stratification?
- Indicate the sample Size (n=xx)
  - Computer, table, formula, or statistician
  - Based on estimated proportion of what is being measured, precision and variation required

3.6 Participant enrollment

- How will eligible participants be identified
- Will study participants be assigned a study ID number?
- Who will talk to the eligible participants to ask them to participate?
- Will a form be filled at the time of enrollment?

3.7 Variables

- List each variable necessary for the analysis
  - Dependent, independent, control…
- Describe what it will look like and how it will be created
  - Type (dummy, index, scale, categorical, etc.)
  - Values
  - Source of information
  - Any recoding, calculations involved, etc.

Variables : Examples

- Socio-demographic characteristics
  - Gender, age, etc.
- Culture dates and results
- Treatment outcomes
- Drug expiry dates
- Distance between patient's house and clinic
- Number of patients per DOT worker
- Patient's accurate knowledge of TB transmission
- Quality of TB services at clinics

3.8 Data collection

- Design questionnaire/data collection forms
- Describe the data collection process
  - Who will collect/abstract the data?
  - How many data collectors?
  - Who will supervise the data collection?
- Describe data entry process
  - Who will enter the data?
  - Software used?
  - How will the data quality be checked?

Sources of data

- Records
  - Patient histories, lab registry, personnel roll, etc.
- Interviews or surveys
- Observations
- Tests/instruments

3.9 Data analysis plan

- Software used
- Statistical approach
  - Univariate – frequencies
  - Bivariate – x by y tables
  - Multivariate (regression)
- Planned tables and figures

Refer to objectives and research question!

3.10 Limitations

- Description of bias you might expect
  - Self-report bias
  - Recall bias
  - Lag-time bias
  - Selection bias
- Limitations in design that you cannot fix
- Can you estimate true cause -> effect?

3.11 Pilot Study

- Purpose: to check study procedures:
  - Understanding (survey questions)
  - Acceptability
  - Feasibility
  - Time, distances
  - Further training needs
- Describe where, when and how pilot testing will occur

3.12 Training

- Purpose of training
- Who will be trained, by whom
- What you will train them to do
- When training will occur
- Where training will take place

4. HUMAN SUBJECTS PROTECTION

- Describe all measures taken to protect study participants from harm
  - Describe possible harm – pain, risks, embarrassment, costs, costs & risks to health services
  - (How) will informed consent be obtained?
  - How will confidentiality/privacy be protected
  - What will be done if a problem arises

Informed consent

- Needed for research with human subjects
- Must follow three principles
  - Participation is voluntary
  - Participant must be able to understand the purpose of the research, in language that is understandable
  - Participation must NOT be coerced

Consent Form: Essential Components

- Describe nature of research and procedures
- Describe nature and duration of participation
- Risk & benefits (physical, psychological, social)
- Stress that participation is voluntary
- Must state that the participant can stop taking part at any time
- Describe how you will protect confidentiality
- Name of contact who can answer questions about the study

Examples of protective measures

- Writing consent forms in simple language
- Translate consent forms into participants' native language
- Not writing patients' name on data collection forms
- Keeping filled forms in locked cabinets
- Destroy pages containing identifiers once data is entered
- Training data collectors to maintain confidentiality

5. TIMELINE

- Work back from deadline
- Proposal writing — Month 1-2
- Recruiting, training of field workers & piloting — Month 3
- Finalizing interview format — Month 3
- Permission, ethical clearance, funding — Month 5
- Data collection — Month 6-8

TIMELINE Continued

- Data collection — Month 6-8
- Data entry & cleaning — Month 8, 9
- Analysis — Month 9, 10
- Report writing — Month 11
- Deadline for report — Month 12
- Community meetings — Month 12-14
- Conference presentation — Month 12-14

6. BUDGET

- What resources will be needed to conduct the study?
- Salaries: Different personnel categories
  - (Amount) x (duration) = Total
- Training (equipment, space, refreshments, etc.)
- Travel (airfare, bus fare, hotel, per-diem)
- Office accommodation and furniture
- Equipment (computer, printer, software)

BUDGET Continued

- Stationary, photocopies, printing
- Telephones, faxes, couriers, postage
- Report dissemination
  - Written
  - Conference attendance
  - Community meetings – hire of hall, refreshments
- Administrative charges

7. INVESTIGATORS and RESPONSIBILITIES

- Name, title, affiliation
- Contact information for each
  - Address, phone, fax, e-mail
- Describe in detail who will do what

8. RESULTS DISSEMINATION

- To whom will results be reported?
- Unethical NOT to report research (whatever the results)
- Variety of forums
  - conference, publication, report, community meeting

9. APPENDICES

- Information sheet and consent form
- Data collection instrument(s)
- Other relevant documents

Southeastern National Tuberculosis Center
SHARE • LEARN • CURE

Southeastern National Tuberculosis Center
SHARE • LEARN • CURE

Conclusion