How to develop a research question

Farzana Choudhury, PhD

fchoudhu@usc.edu

Assistant Professor
Division of Biostatistics

LA's BeST @USC

Objectives: develop research question

- Review the scientific method

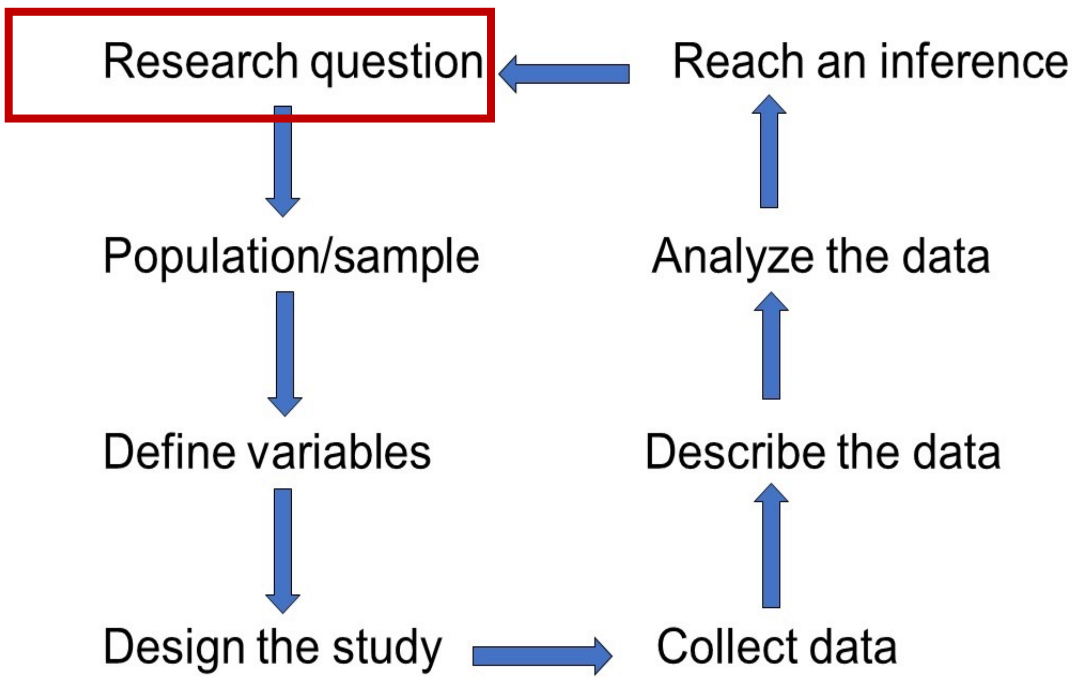
- Understand how to formulate a research question
  - What makes a well-specified research question?

- What's an appropriate/feasible study design to address the question?

- Be an informed consumer of scientific study results, and potentially plan your own!

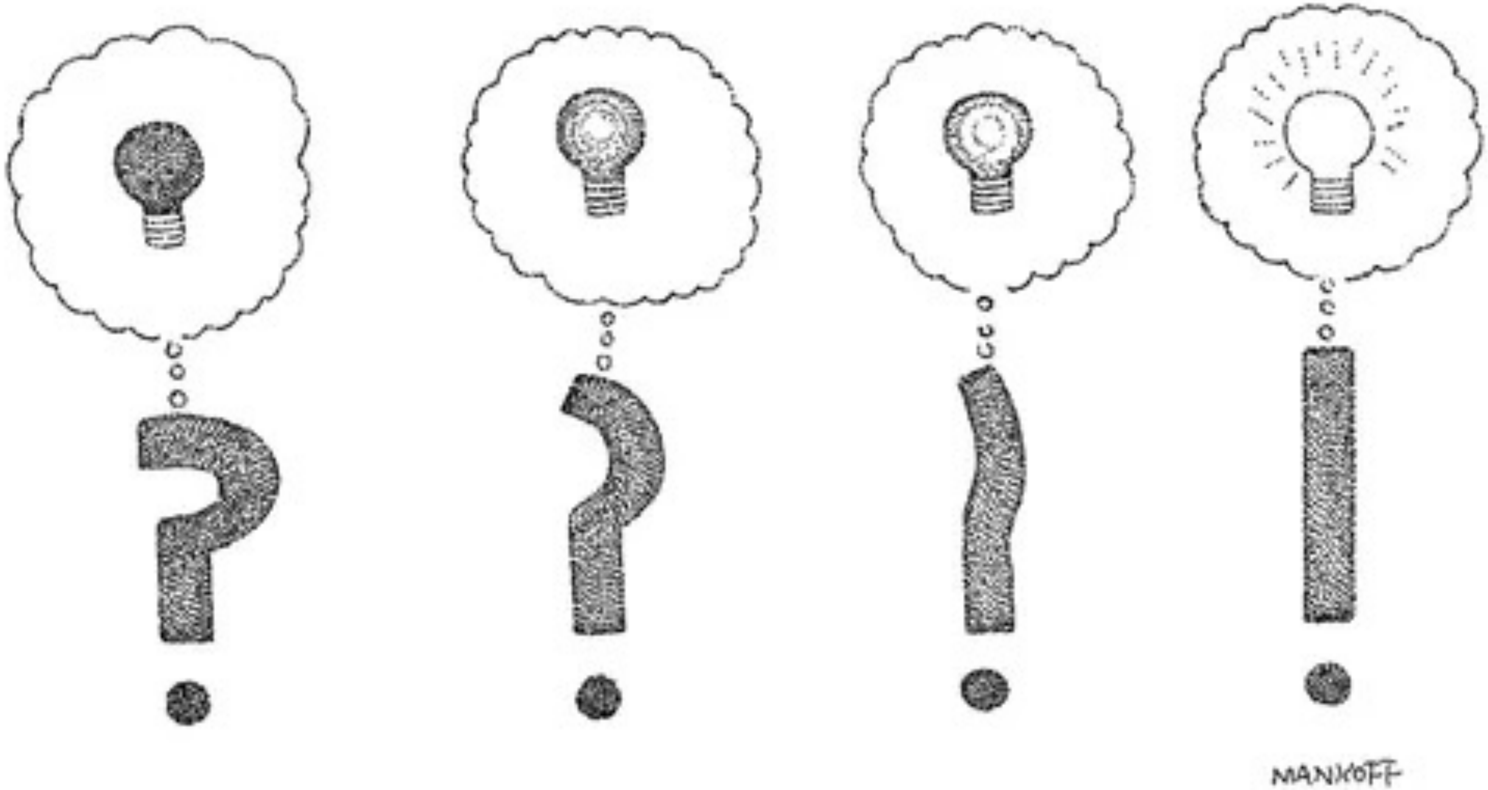
LA's BeST @USC

Research



LA's BeST @USC

Objectives: develop research question



"The scientist is not a person who gives the right answers, he is one who asks the right questions."

-Claude Levi-Strauss
(French philosopher)

LA's BeST @USC

The Scientific Method

Science is a method of gaining knowledge.

The Scientific Method

1. Identify a question.

2. Propose an answer (research hypothesis).

3. Plan a research design (method) to test the hypothesis.

   Identify population

   Identify variables

4. Collect empirical data.

5. Analyze data.

6. Based on analyzed data, confirm or denounce hypothesis.

7. Interpret results.

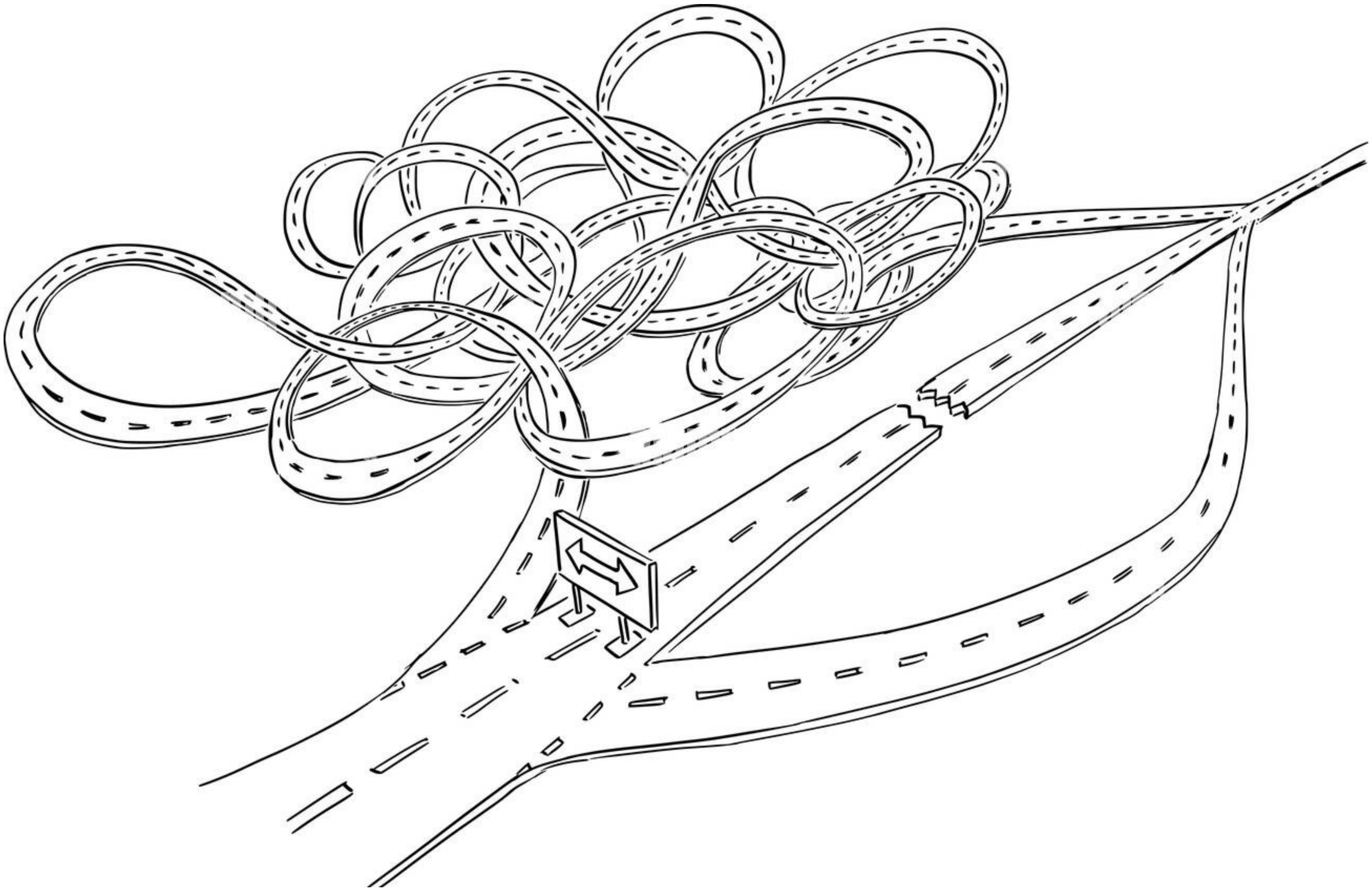
Defining the research question, translating into a statistical question

- What are the components of a good research question?
  - FINER and PICOT criteria (we'll come to these later)

- **How** do I translate my research question to a statistical question/hypothesis that I can test?
  - Be specific and quantitative!
  - Consider the **study design(s)** that can be used to answer the research question

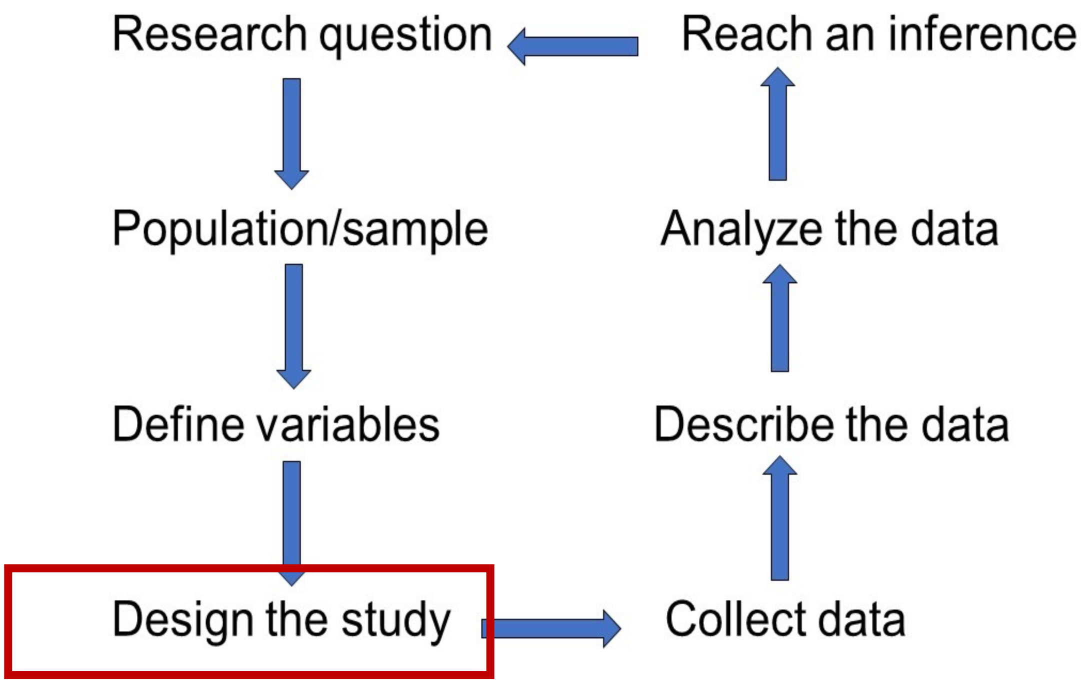
LA's BeST @USC

First, let's discuss study designs



LA's BeST @USC

Research



LA's BeST @USC

Study Designs in Medical Research

Study designs

<u>Descriptive</u>

- Describes only
- Does not examine associations
  - Case reports
  - Case series

<u>Explanatory</u>

- Examines associations
- Tests hypotheses

<u>Experimental</u>

- Investigator assigns exposures
  - Clinical trials

<u>Observational</u>

- Investigator observes without intervention
  - Ecological
  - Cross-sectional
  - Case-control
  - Cohort

LA's BeST @USC

Descriptive Studies

- Case reports
- Case series
- Registry summaries
- Surveys
- Any non-experimental study

Study Designs in Medical Research

Study designs

Descriptive
- Describes only
- Does not examine associations
  - Case reports
  - Case series

Explanatory
- Examines associations
- Tests hypotheses

Experimental
- Investigator assigns exposures
  - Clinical trials

Observational
- Investigator observes without intervention
  - Ecological
  - Cross-sectional
  - Case-control
  - Cohort

LA's BeST @USC

Observational studies - definition

- Clinicaltrials.gov: a clinical study in which participants identified as belonging to study groups are assessed for biomedical or health **outcomes**. Participants may receive diagnostic, therapeutic, or other types of interventions, but **the investigator does not assign participants to specific interventions** (as in a clinical trial).

- Associations between exposures/interventions and outcomes may be biased (confounded) by characteristics that differ between those that choose exposure vs. no exposure.

LA's BeST @USC

Observational Designs

- Ecologic
- Cross-sectional
- Case-Control (Retrospective)
- Cohort (Prospective)

Observational: Ecological study

- Unit of observation is a population or community, rather than individuals

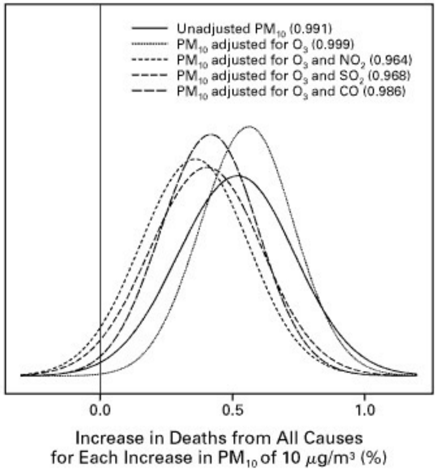
- Compare community-level outcomes/exposures across a number of communities

- Useful for monitoring population-level health

LA's BeST @USC

Ecological study: example in literature

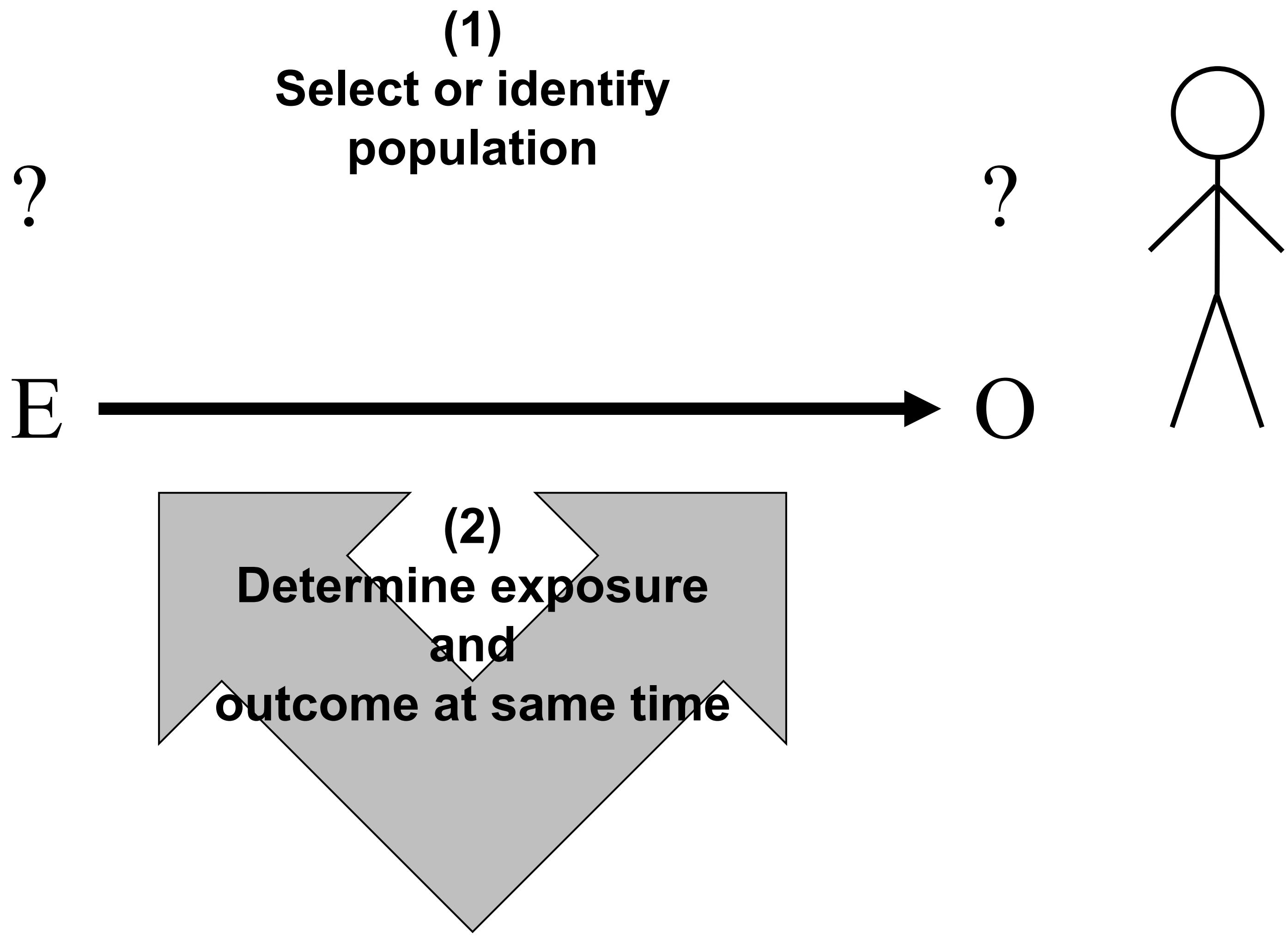
- Samet et al 2000. Fine particulate air pollution and mortality in 20 U.S. cities, 1987-1994. https://www.ncbi.nlm.nih.gov/pubmed/11114312
- **Question**: Are city-level daily air pollution levels related to daily city-level mortality counts?
- **Design/population**:
  - 20 largest cities (counties) in the US
  - Collected: administrative daily, city-level data on: deaths, air pollution, meteorology



- **Results**: "Every 10 µg/m^3 increase in PM_{10} was associated with 0.68 % increase in relative rate of death from cardiovascular/respiratory causes (95% CI: 0.20, 1.16)"

LA's BeST @USC

Cross-sectional Study



LA's BeST @USC

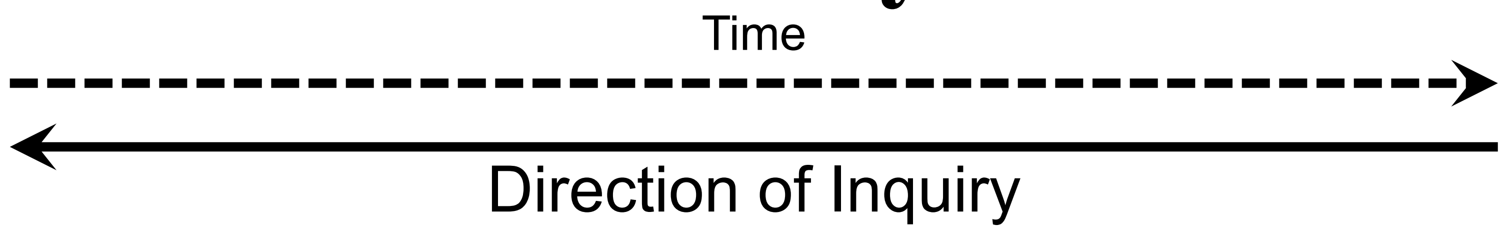
Cross-sectional: example in literature

- Chen et al 2015. **Ambient air pollution and neurotoxicity on brain structure: Evidence from women's health initiative memory study** https://www.ncbi.nlm.nih.gov/pubmed/26075655
- **Question:** Does air pollution exposure impact the brains of older women?
- **Study:**
  - Population: 1,403 community-dwelling older women without dementia in Women's Health Initiative Memory Study
  - Outcome: MRI scans of gray matter (GM) and white matter (WM) volume at ages 71-89 (2005-2006)
  - Exposure: cumulative PM2.5 exposure from 1999-2006 based on residential histories and air pollution monitoring network data

- **Results: "**Older women with greater $PM_{2.5}$ exposures had significantly smaller WM, but not GM, volumes, independent of geographical region, demographics, socioeconomic status, lifestyles, and clinical characteristics, including cardiovascular risk factors."

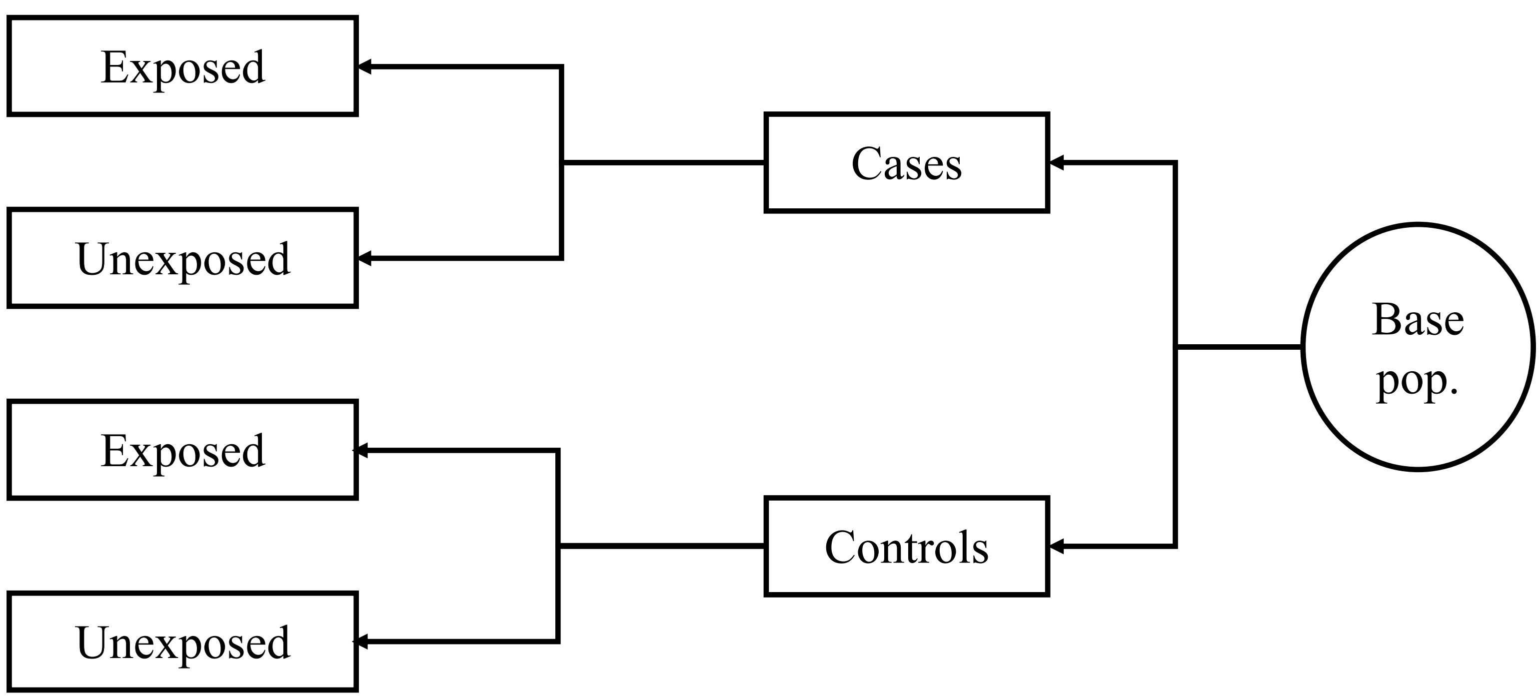
LA's BeST @USC

Case-Control Study



2. Go back and determine if they were previously exposed

1. Recruit people with disease (cases) and without disease (controls)



LA's BeST @USC

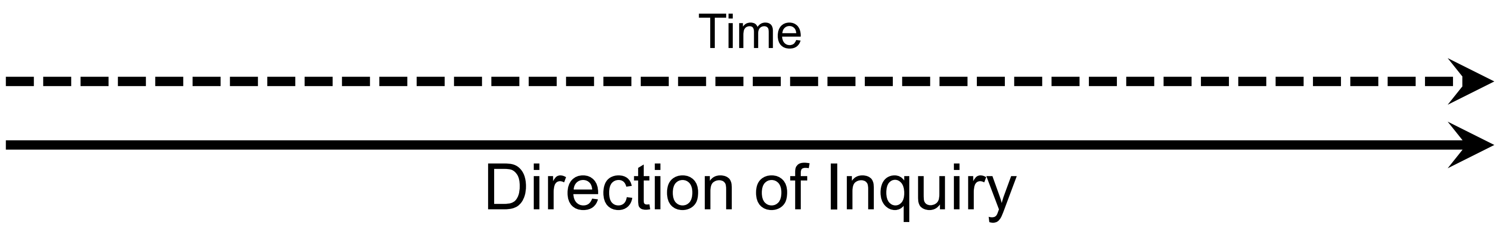
Case-control: example in literature

Volk et al 2013. **Traffic-Related Air Pollution, Particulate Matter, and Autism.** https://www.ncbi.nlm.nih.gov/pubmed/23404082

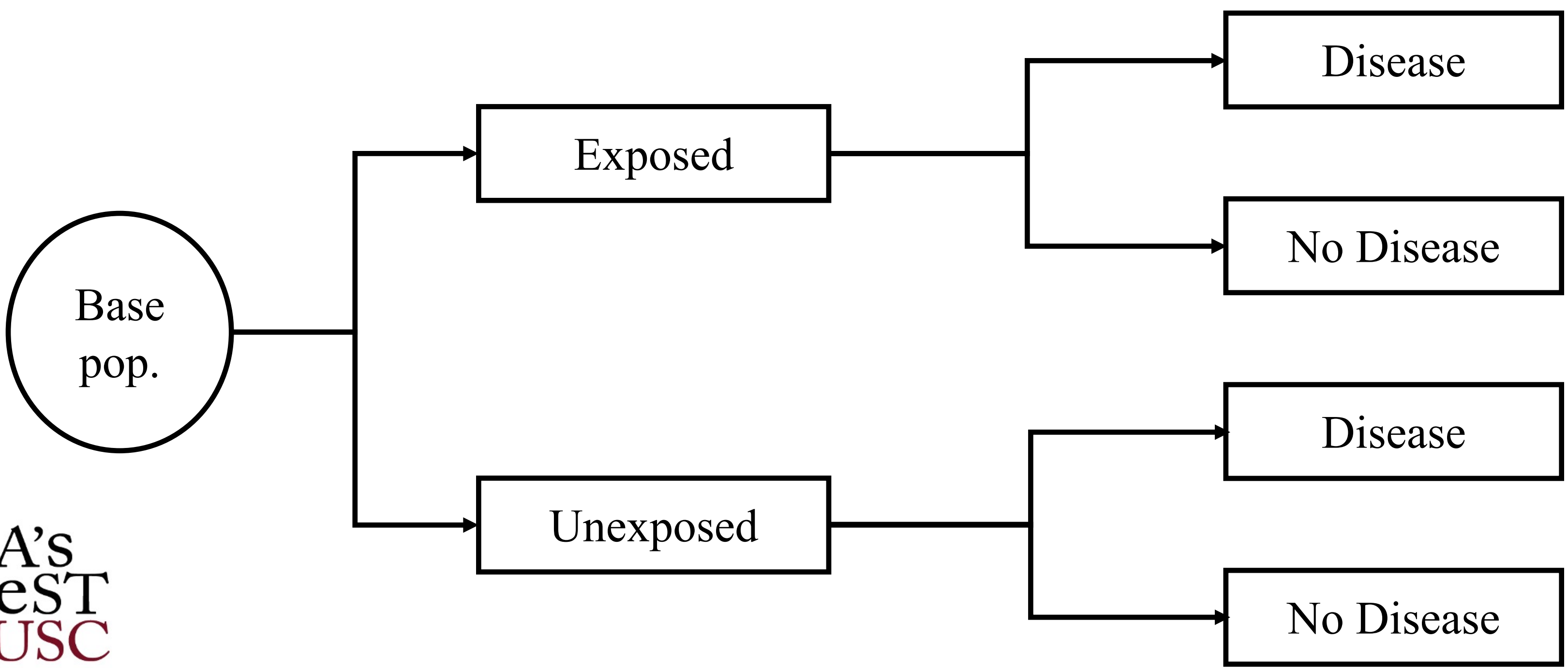
- Question: Is traffic-related pollution exposure related to the development of autism?
- Study design:
  - Cases: 279 children with autism, from CA Dept. Dev. Services
  - Controls: 245 sampled from CA births (age, sex, area matched)
  - Mother's address from birth certificate used to assign traffic exposures
- Results: "Children with autism were more likely to live at residences that had the highest quartile of exposure to traffic-related air pollution, during gestation (AOR, 1.98 [95% CI, 1.20-3.31]) and during the first year of life (AOR, 3.10 [95% CI, 1.76-5.57]), compared with control children"

LA's BeST @USC

Cohort Study



1. Determine exposure in study population
2. Follow subjects and see if they develop disease



LA's BeST @USC

Cohort study: example in literature

Cheng et al 2019. **Association between ambient air pollution and breast cancer risk: The multiethnic cohort study**
https://www.ncbi.nlm.nih.gov/pubmed/30924138

- **Question:** Is air pollution exposure associated with an increased risk of breast cancer?
- **Study design**
  - 57,589 females from the multiethnic cohort, residing mostly in LA County from recruitment through 2010
  - Outcome: breast cancer incidence (Cox model)
  - Exposure: time-varying air pollution assignments

- **Results:** "Among women who lived within 500 m of major roads, significantly increased risks were observed with NOx (hazard ratio [HR] = 1.35, 95% confidence interval [95% CI]: 1.02–1.79)… Subgroup analyses suggested stronger associations of NOx and NO2 among African Americans and Japanese
Americans."

LA's BeST @USC

Comparison of Observational Study Designs

Type of Study	Group(s) defined by	Observe	Pros	Cons
Ecological	Geography	Exposure and outcome (in a community)	• Very Inexpensive • Covers a large population • Readily available administrative data can be used	• Ecological fallacy • Outcome/disease may vary by location
Cross-sectional	Defined population	Exposure and outcome	• Quick and easy • Inexpensive • Can often use existing data • Efficient for common exposure and outcome	• Temporal ambiguity • Prevalence bias • Recall bias
Case-Control	Disease status	Exposure	• Efficient in long duration between E and O • Low time and cost • Efficient for rare outcomes	• Control selection challenging • Differential misclassification of exposure • Recall bias
Cohort	Exposure	New cases of outcome	• Efficient for rare exposure • No temporal ambiguity • Can study multiple outcomes	• Time and cost • Self-selection bias • Incomplete follow-up • Inefficient for rare outcomes

Nested Case-Control Study

- A case-control study can be “nested in” an ongoing follow-up study
  - Could be nested in a cohort study
  - Could also be nested in a clinical trial
- Combines the strengths of both case-control and follow-up studies
  - Information on exposure collected before the development of disease
  - Cost-efficient when measurement of exposure involves expensive analysis of previously collected specimens

Study Designs in Medical Research

Descriptive
- Describes only
- Does not examine associations
  - Case reports
  - Case series

Explanatory
- Examines associations
- Tests hypotheses

Experimental
- Investigator assigns exposures
  - Clinical trials

Observational
- Investigator observes without intervention
  - Ecological
  - Cross-sectional
  - Case-control
  - Cohort

LA's BeST @USC

Clinical trial - definition

- Clinicaltrials.gov: A **clinical study** in which participants are **assigned** to receive one or more **interventions** (or **no intervention**) so that researchers can evaluate the effects of the interventions on **biomedical or health-related outcomes.** The assignments are determined by the study protocol. Participants may receive diagnostic, therapeutic, or other types of interventions.

  A cohort study where persons are "assigned" to exposures (interventions) and followed for ascertainment of outcomes.

  Clinical trials are not feasible when assignment to an exposure/intervention is not ethical.

Trial designs

- **Parallel group**: Each participant is assigned to one (and only one) of the trial interventions. Standard approach for most clinical trials

- **Crossover**: Each participant receives both the experimental and comparator interventions, usually in randomized order, with a washout period between interventions

  Perfect matching – each participant acts as their own control – requires fewer subjects

  Disadvantages: Greater likelihood of dropout; must be a stable disease under study; only appropriate for interventions that wash-out and have short-term (not permanent) outcomes

LA's BeST @USC

Assignment to intervention

- **Randomization**
  Greatly reduces the possibility of systematic differences between study groups that might taint (confound) your conclusions regarding the efficacy of your experimental intervention. Key advantage over observational studies.

- **Blinded vs open-label**
  (do participants/researchers know the intervention?)
  Knowing the intervention may affect how a participant responds, outcomes.

Randomized controlled trial (RCT)

Jhun et al 2017. School environmental intervention to reduce particulate pollutant exposures for children with asthma https://www.ncbi.nlm.nih.gov/pubmed/27641483

- **Question:**
  Does classroom-based air cleaner/filter intervention reduce particulate pollutants in classrooms of children with asthma?
  Does the lung function of these children improve with improved air?
- **Design (pilot study):**
  - Population
    - 18 classroom (9 control, 9 intervention)
    - 25 kids with asthma (13 control, 12 intervention)
  - Intervention: Air filter in classrooms (randomized: filter or sham)
  - Measurements: air pollution in classrooms and lung function once before and twice after air filtration
- **Results:**
  - "The intervention group had greater reductions in $PM_{2.5}$ levels compared with the control group (2.3 μg/m^3, 95% CI, −3.5 to −1.0; *P* = .003)
    [49% reduction]
  - The intervention group had a greater improvement in [peak expiratory flow]PEF compared with the control group (0.46 L/s, 95% CI, 0.09-0.83; *P* = .03) [16% improvement]"

LA's BeST @USC

Distinguishing Study Designs

Did the investigator(s)...

1) ...examine an association (Qx: If E, then O)?
2) ...assign E to subjects and then follow them for O?
3) ...ascertain E and O for geographic areas or groups of people rather than for individual subjects?
4) ...ascertain E and O at roughly the same time on ALL eligible subjects?
5) ...identify cases of O (& non cases) before ascertaining previous E?
6) ...identify E on subjects first before following them for onset of O?

Which study design?

1) If NO, study design is Descriptive.
2) If YES, study design is Experimental.
3) If YES, study design is Ecological.
4) If YES, study design is Cross-Sectional.
5) If YES, study design is Case-Control.
6) If YES, study design is cohort.

Consider these study designs in terms of the following broad research question

Broad research question:

Does air pollution impact health?

(Note: you will hear a lot about the Southern California Children's Health Study (CHS)—led by USC investigators—in future class sessions, so here we consider *other* studies to provide perspective)

Cross-sectional study: alternatives

- **Cohort Study Alternative:**
  Obtain air pollution levels in group of children without asthma. Follow to xx (time) to ascertain development of asthma. Compare asthma rates by air pollution level.

- **Case-Control Alternative:**
  Select samples of persons with asthma (cases) and without asthma (controls). Access historical data on air pollution levels in their hometowns. Compare pre-disease air pollution in cases and controls.

…and we're back to research question!

Translating Research Question to Statistical Question

- What are the components of a good research question?
    - FINER and PICOT criteria (we'll come to these **NOW!**)

- **How** do I translate my research question to a statistical question/hypothesis that I can test?
    - Be specific and quantitative!
    - Consider the **study design(s)** that can be used to answer the research question

LA's BeST @USC

Research question → statistical question

Research Question
(general concepts)

→

Statistical Question
(statistical concepts)

Design a study and collect data to
test the statistical question
and answer the research question

LA's BeST @USC

Example: Research question → statistical question

Research Question:
Does higher systolic blood pressure (BP) increase dementia in the elderly?

Statistical Question(s):
Among persons age>70, do those who develop dementia have a higher proportion of systolic BP>120 than those who do not develop dementia?
Or…
Among persons age>70, is the incidence rate of dementia greater in those who have systolic BP>120 compared to those who did not?

1st statistical question: case-control study;
Statistical concepts: Compare proportions (with systolic BP>120)

2nd statistical question: cohort study
Statistical concepts: Compare incidence rates (developing dementia)

Other options: randomized trial?

LA's BeST @USC

The research question

- States a relationship between two or more variables, phrasing in terms of some question.
- Why is this research important? What is the research gap in our scientific understanding?
- What is the past research in this area?
- What areas need further exploration?
- Can my study help fill in these gaps or lead to greater understanding?

Refining the research question

- Identify the main concepts or keywords of your research question
- Topic too broad? Add more concrete or specific terms to your question
- Topic too narrow? Broaden content of terms

LA's BeST @USC

FINER criteria to develop the research question

FEASIBLE – F

- Required Expertise
- Research facilities
- Available financial resources
- Not too many or too ambitious research objectives

INTERESTING – I

- Answering the scientific question should entangles the scientific community
- Presents a different perspective of a problem

NOVEL – N

- Contributes new information
- Generate new hypotheses
- Confirmatory study should avoids the weaknesses of previous studies

ETHICAL – E

- Respect for patients and for lab animals
- Complies with local ethical committees
- Beneficence
- Unbiased access to the benefits of research

RELEVANT – R

- Consider how the outcomes might advance the scientific knowledge, or guide further research
- Provide an accurate answer to a specific research question

FINER criteria to develop the research question

- **F Feasible**
  Adequate number of subjects and technical expertise
  Affordable in time and money
  Manageable in scope

- **I Interesting**

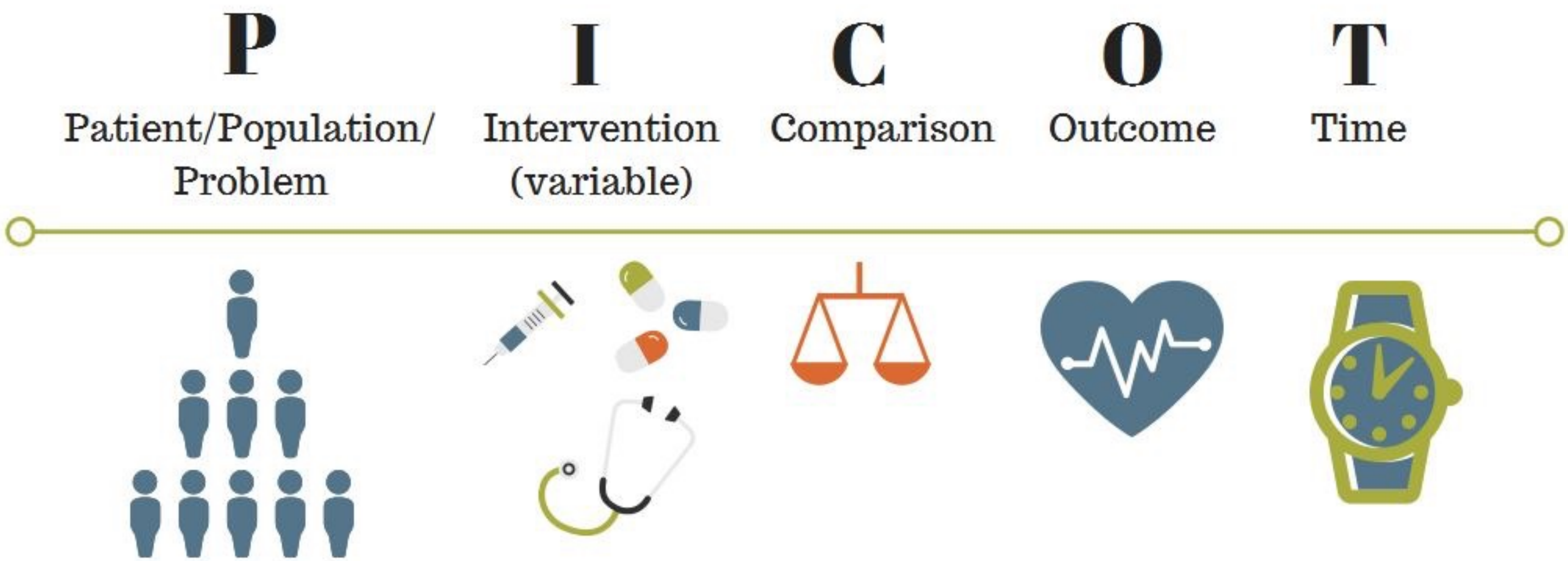
- **N Novel**
  Confirms, refutes or extends previous findings

- **E Ethical**

- **R Relevant**
  To scientific knowledge, clinical health and policy, future research

LA's BeST @USC

PICOT criteria to develop the research question



Sackett D, Richardson WS, Rosenburg W, Haynes RB. How to practice and teach evidence based medicine. 2nd ed. Churchill Livingstone; 1997.

PICOT criteria to develop the research question

- **P Population**
  What specific population will you test the intervention in?

- **I Intervention (or Exposure)**
  What is the intervention/exposure to be investigated?

- **C Comparison Group**
  What is the main comparator to judge the effect of the exposure/intervention?

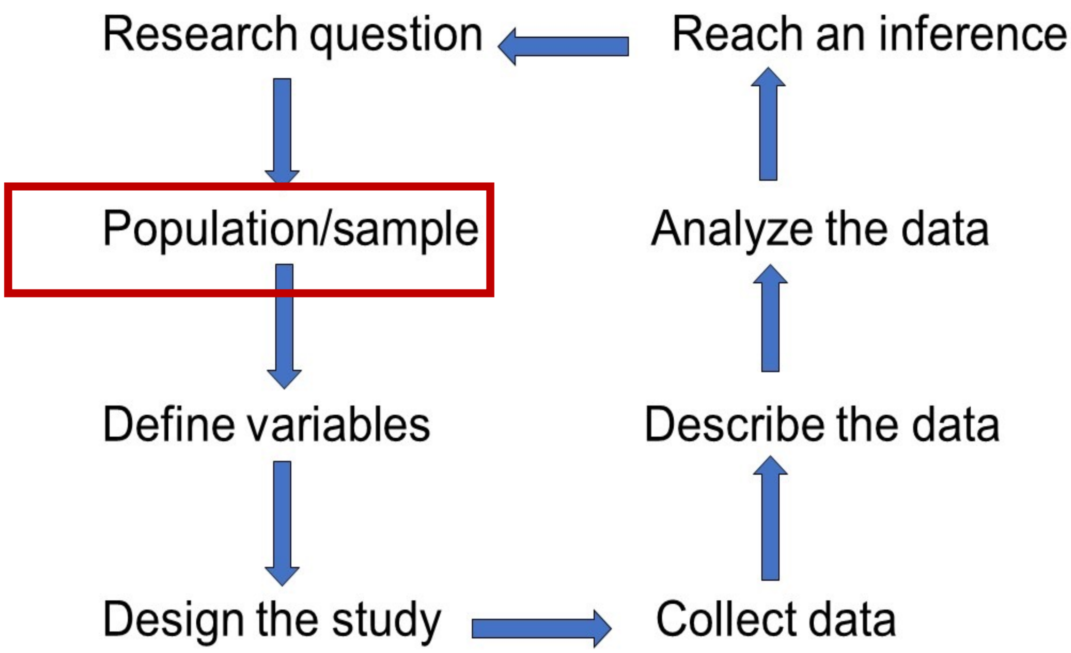
- **O Outcome**
  What will you measure, improve, affect?

- **T Time**
  Over what time period will outcome be assessed?

LA's BeST @USC

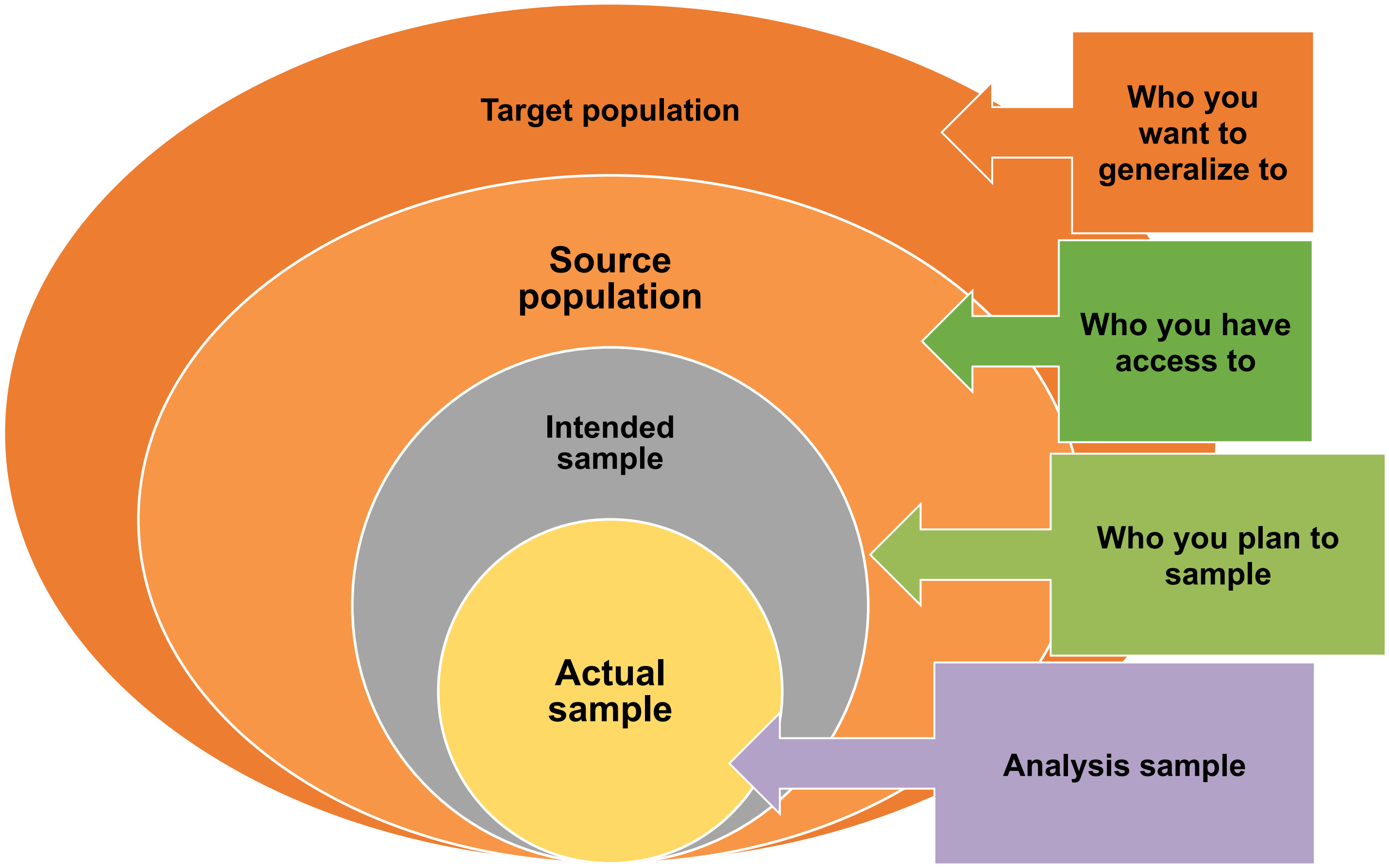
Sackett D, Richardson WS, Rosenburg W, Haynes RB. How to practice and teach evidence based medicine. 2nd ed. Churchill Livingstone; 1997.

Research



LA's BeST @USC

Defining your Population: Population to Sample



LA's BeST @USC

Defining your population and sample

- **Target populations**
  - What populations are relevant to the research question?
    - Demographics (e.g., age>70)
    - Clinical characteristics (not diagnosed with dementia)
- **Source populations**
  - What population with the characteristics of the target population are available to you?
    - Geography (residents of Los Angeles county)
    - Temporal (e.g. recruitment period 1/1/2020 – 12/31/2023)
    - Availability affects generalizability and reproducibility
- **Intended sample**
  - The part of the source population you will attempt to recruit; ideally representative of the source population
  - Prevalence of target characteristics

LA's BeST @USC

Operationalizing the general concepts in your research question

- Research Question: Does elevated systolic blood pressure (BP) increase risk of dementia in elderly persons?
- **Population:** Elderly persons
  Operationalize (who we will sample from): persons aged>70 without dementia (at baseline)
- **Intervention/exposure:** Systolic BP level
  Operationalize: Average systolic BP>120 in prior 1 year
- **Comparator (and exposed) groups:** Persons age>70 without dementia at baseline, without and with average systolic BP>120 in the prior year
- **Outcome:** Dementia
  Operationalize: new diagnosis of dementia
- **Time**: over 5 years of follow-up (time)

LA's BeST @USC

Recap: how to develop a research question

- Review the scientific method
- Understand how to formulate a research question
  - What makes a well-specified research question?
- What's an appropriate/feasible study design to address the question?
  - A broad research question can be addressed by many different study designs!

LA's BeST @USC

Activity time!

theguardian.com/environment/2021/sep/06/gen-z-climate-change-careers-jobs

This land is your land
Environment

'No point in anything else': Gen Z members flock to climate careers

Colleges offer support as young people aim to devote their lives to battling the crisis

Supported by

SEJ www.sej.org

About this content

Angela Lashbrook

Mon 6 Sep 2021 06.00 EDT

60

▲ Hundreds of protesters march to the White House calling for climate action, including a Civilian Climate Corps. Photograph: Allison Bailey/Rex/Shutterstock

California is facing a drought so devastating, some publications call it "biblical". Colorado now has "fire years" instead of "fire seasons". Miami, which sees more dramatic hurricanes each year, is contemplating building a huge seawall in one of the city's most scenic tourist districts to protect it from storm surges.

"Once you learn how damaged the world's ecosystems are, it's not really something you can unsee," says Rachel Larrivee, 23, a sustainability consultant based in Boston. "To me, there's no point in pursuing a career - or life for that matter - in any other area."

LA's BeST @USC

The New York Times

Stanford Gets $1.1 Billion for New Climate School From John Doerr

The billionaire venture capitalist said the study of climate change and sustainability would be the “new computer science.”

John Doerr, the venture capitalist who invested in tech companies such as Slack, Google and Amazon, on the Stanford campus. Carolyn Fong for The New York Times

By David Gelles
David Gelles writes about climate change and business, and has interviewed hundreds of C.E.O.s in recent years.

May 4, 2022 Updated 11:50 a.m. ET

John Doerr, one of the most successful venture capitalists in the history of Silicon Valley, is giving $1.1 billion to Stanford University to fund a school focused on climate change and sustainability.

“Climate and sustainability is going to be the new computer science,” Mr. Doerr, who made his estimated $11.3 billion fortune investing in technology companies such as Slack, Google and Amazon, said in an interview. “This is what the young people want to work on with their lives, for all the right reasons.”

“Climate and sustainability is going to be the new computer science”

https://www.nytimes.com/2022/05/04/climate/john-doerr-stanford-climate.html

LA’s BeST @USC

“Tackling climate change could be the greatest global health opportunity of the 21st century”

“Many mitigation and adaptation responses to climate change are “no-regret” options, which lead to direct reductions in the burden of ill-health, enhance community resilience, alleviate poverty, and address global inequity. Benefits are realised by ensuring that countries are unconstrained by climate change, enabling them to achieve better health and wellbeing for their populations. These strategies will also reduce pressures on national health budgets, delivering potentially large cost savings, and enable investments in stronger, more resilient health systems.”

Watts N, Adger WN, Agnolucci P, Blackstock J, Byass P, Cai W, Chaytor S, Colbourn T, Collins M, Cooper A, Cox PM. Health and climate change: policy responses to protect public health. The Lancet. 2015 Nov 7;386(10006):1861-914.

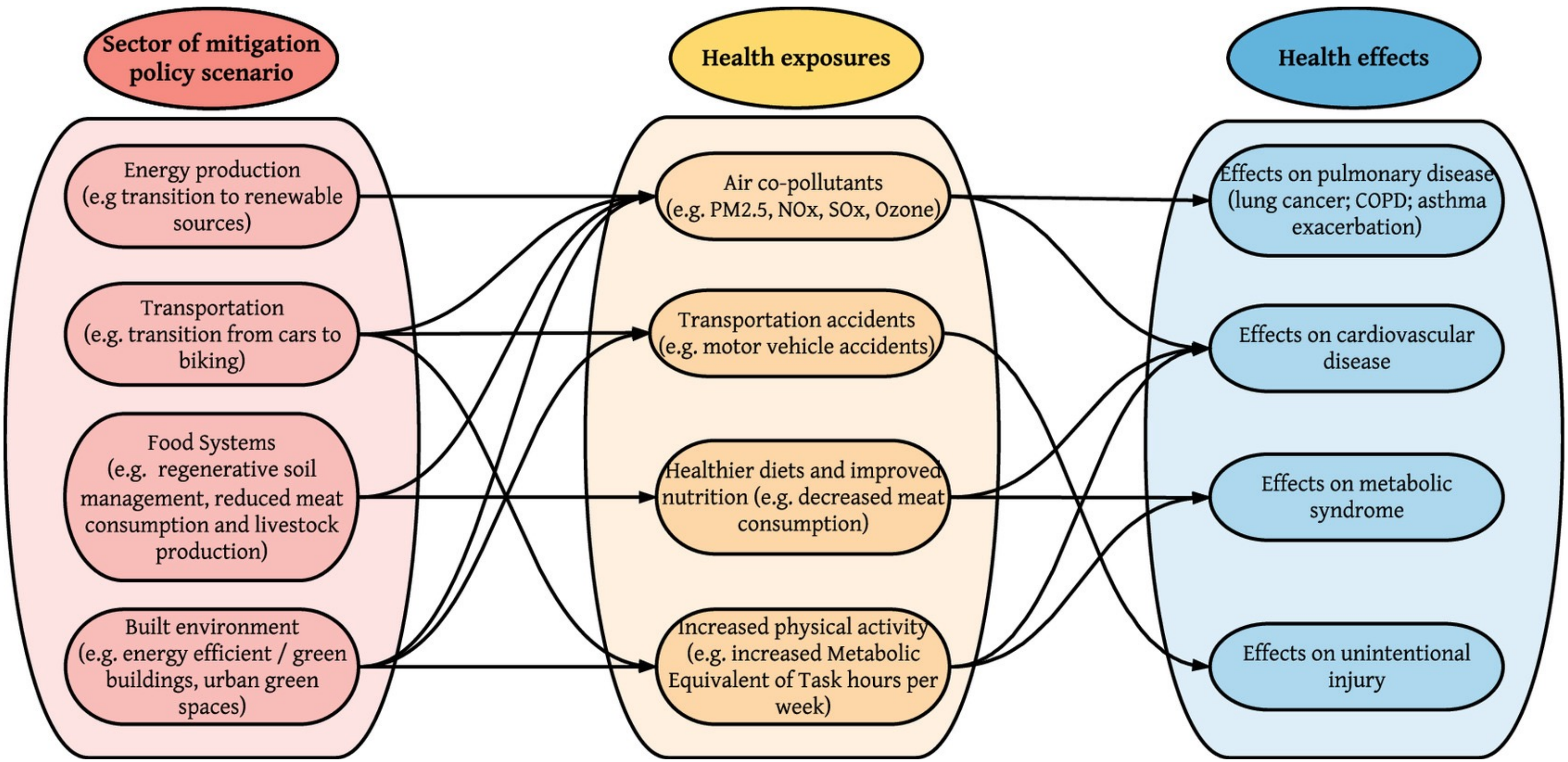
Meng YY, Yue D, Molitor J, Chen X, Su JG, Jerrett M. Reductions in NO2 and emergency room visits associated with California's goods movement policies: A quasi-experimental study. Environmental Research. 2022 Jun 3:113600. **https://pubmed.ncbi.nlm.nih.gov/35660569/**

Background:

- SoCal has the 2 busiest container ports in the US
- 2006: emissions reduction plan (ERPPGM) for goods movements
- While air quality improved, health impacts have not been directly studied

LA's BeST @USC

Idea of today's activity: Design a study to describe ***health co-benefits*** of climate change mitigation efforts



https://ehp.niehs.nih.gov/doi/10.1289/EHP6745

LA's BeST @USC

Plans for today's activity

Theme:
Research questions on *health co-benefits of* climate change mitigation efforts

LA's BeST @USC

Plans for today's activity

Theme:
Research questions on *health co-benefits of* climate change mitigation efforts

Group 1: Randomized controlled trial
Group 2: Cohort
Group 3: Cross-sectional
Group 4: Case-control
Group 5: Ecological