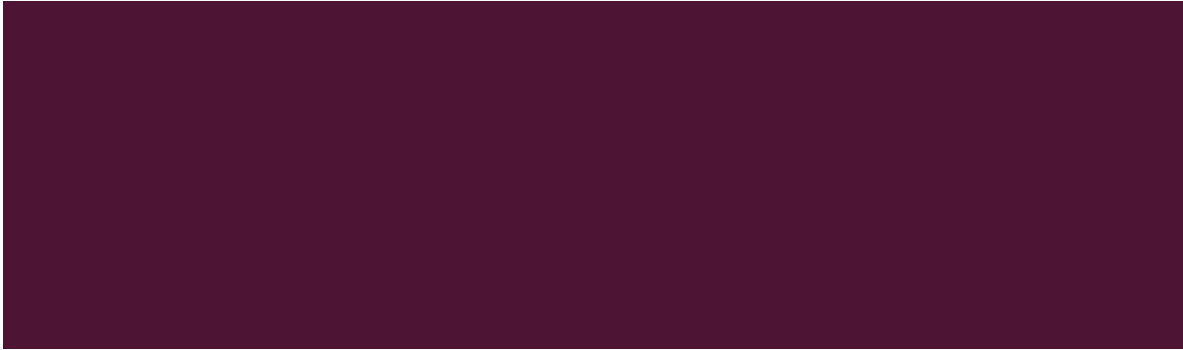


# PCOR RESEARCH DESIGN

DAY 2: THE SHARE PROJECT PCORI TRAINING



## START WITH RESEARCH INTERESTS

It is not uncommon for research interests to be directly or indirectly from things researchers have experienced in their lives, their loved ones lives, their neighborhoods and communities, and society as a whole.

Many people have a “What brought me here” story.

**What’s yours?**

## RESEARCH PROBLEM

**Turn that interest into a research problem.**

Try framing the interests, your “Why” into a solvable problem. Is it:

1. An unmet need?
2. Gap in services or care for people?
3. Lack of resources?
4. Lack of information or awareness about your health condition?
5. Need for policy change?

## RESEARCH QUESTIONS

### Turn that problem into a patient-centered research question.

Try turning that problem into an answerable question or set of questions that will focus research activities. Examples:

1. How severe is this problem among patients? Who is impacted the most?
2. What programs, services, resources, assets, policies are available to address this problem? What is lacking?
3. What risk factors contribute to the problem? What protects patients or community members from this problem?
4. What are the health consequences of this problem?

## FORMING RESEARCH QUESTIONS

Why are you here?

What interests do you have?

What problems do you see?

What are your research questions?

## IDENTIFY THE TYPES OF DATA YOU WILL NEED TO ANSWER RESEARCH QUESTIONS

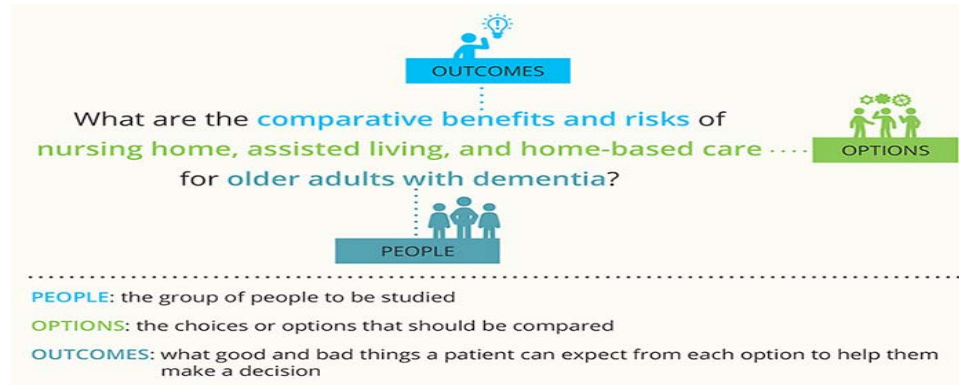
- Many different types of data can answer your research question
  - e.g., service provision, utilization, costs, outcomes, patient characteristics/demographics, patient knowledge/attitudes/beliefs/behaviors, community characteristics/demographics, community assets, community values and priorities, etc.
- It would be good to collect it all, but you need to consider the following when identifying what type of data you need:
  - Time frame
  - Available resources
  - Dedicated staff time and skills
  - Research interests
  - Research questions
  - Desired outcome/next step to solve identified problem

Now, take a moment to think through your team's research questions and what data you will need to answer them

- Use the research questions & data worksheet

Often, organizations skip this step and go straight to thinking about how to collect new data – often designing a survey. But they will later find out that they collected data they did not need to answer their research questions or inform the stated problem, or failed to collect the data they did need.

## EXAMPLE OF PCOR QUESTION



## IDENTIFY DATA SOURCES TO ANSWER YOUR RESEARCH QUESTIONS

- Start first with data that already exists
- Do a web search, look for federal/state/local data sources, ask your partners to see what data is out there and available to use

### Secondary Data

Data that has been collected for another purpose and can be made available to you for your data needs

### Primary Data

New data you collect from patients, community members, policymakers, etc.

## SECONDARY DATA - ADVANTAGES

Secondary data are a valuable tool for researchers.

- Secondary data is cheaper than collecting your own!
- Secondary data can help determine a baseline for your population and research interest
- Secondary data can help you compare your population of interest to other populations or geographies
- Using a secondary data source is quicker than primary data collection
- Secondary data may answer your research question!



Community



Health



Quality



Medicare



Hospital



Inpatient

## SECONDARY DATA - DISADVANTAGES

Secondary data sources may not be exactly what you need

- The data may be outdated
- When it's not publicly available, it may be expensive to gain access
- There are often restrictions to access
- Data quality varies and may not be obvious



Community



Health



Quality



Medicare



Hospital



Inpatient

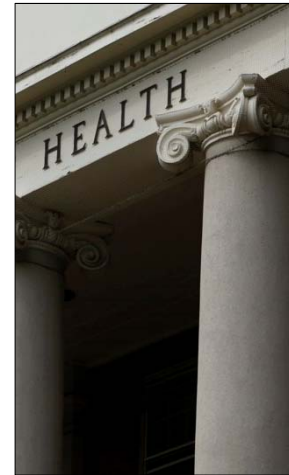
## CONSIDER THE SOURCE

Many organizations have already collected the information you need, or have compiled it and made it readily available.

Where you find your secondary data will influence the quality of the data and how applicable it is to your research.

Good sources of data include:

- Federal, state, and local government
- Schools
- Hospitals
- Advocacy groups and non-profit groups
- Colleges and universities



## EVALUATING DATA – HOW DO WE KNOW WHAT TO USE?

Evaluating the data you find is critical. This is done to combat **fake** or **misleading**, or **misrepresented** data that is readily available, or data that is **not useful for the needs** of your project.

The five criteria for evaluating data are:

1. Credibility
2. Specificity
3. Reliability
4. Generalizability
5. Timeliness

### USA CRIME STATISTICS ~ 2015



SOURCE: Crime Statistics Bureau - San Francisco

Donald J. Trump  
@realDonaldTrump

Follow

Fake crime statistics re-tweeted by Donald Trump, originally from a Neo-Nazi Twitter account. The "Crime Statistics Bureau – San Francisco" does not exist.

## CREDIBILITY

Credibility refers to the **source** of the data. Can you trust the research entity that produced the data or could it be distorted to fit their needs?

Whatever the source of the data you will want to investigate the source and examine the data for bias. Here are a few things to look for:

- **Money or “Kickbacks”**
  - Who paid for, sponsored, or funded the study? How much of a stake does the data source have in a specific finding?
- **Organizational Goals**
  - Research done by certain organizations (e.g. businesses, or political groups) may have **missions** that influence how they collect and interpret data.
- **Reputation-** What is the organization’s public image or reputation for their research?
  - Government and academic institutions are considered credible because research is conducted for the public benefit.

## CREDIBILITY EXAMPLE – ANTIVAX KICKBACK

- The belief that vaccines cause autism which has led many people to refuse to vaccinate their children – “Anti-vaxxers” was caused by a now unlicensed physician who we now know was not a credible researcher. – Andrew Wakefield



Source: The Nib. Full Cartoon found [here](#).

- Wakefield faked his data and created a fake condition: “Autistic enterocolitis”
- He and his family and friends planned to make money on new tests for “Autistic enterocolitis,” and a new MMR vaccine.
- Even though the research as been withdrawn and labeled fake, the Anti-Vaxxer movement is still going strong



## SPECIFICITY & GENERALIZABILITY

**I. Specificity** refers to data that is limited to a particular condition, or the impact on a specific population.

- Sometimes the data will align perfectly with your research, but sometimes it won't.

**II. Generalizability** refers to data on a specific population that is applicable to the general public other populations.

**III.** Specificity is similar to generalizability, but specificity refers to the information collected and generalizability refers to the population. Ask yourself:

- How close is the relationship between the data provided and the data you need?
- Does it make sense to apply information from this population to the patients I'm serving?
- Are there reasons why this information may not be generalizable?
- Do I have biases that may make this information seem useful but are possibly not true?

## SPECIFICITY & GENERALIZABILITY EXAMPLE



You are looking for information on the number of Vietnamese-American women who smoke for your study, but so far, you haven't found anything specific for this population.

- You find the number of Asian-Americans who smoke and the number of women in the United States who smoke, but nothing specific on Vietnamese-American women smokers.
- Next you find a study on information on the number of Chinese-American women who smoke.

Would this information be enough for you to work with? Is it specific to your population? Is it generalizable?

## SPECIFICITY & GENERALIZABILITY EXAMPLE



What you are looking for	What you've found	Concern
# of Vietnamese women who smoke	(1) # All Asian-Americans who smoke and % of all women in the United States who smoke	Not specific!
	(2) # Chinese-American women who smoke	Not generalizable!

- It's actually very common that the term "Asian American" will group people together, regardless of ethnicity & country of origin, despite very different backgrounds, socio-economic status and needs.
- With information (1) we we should not assume we know the ethnic breakdown of what is defined as "Asian American" or that it accurately represents Vietnamese Americans
- With information (2) identify personal biases, and reject false relationships that have been created with little cultural-awareness
- Instead, continue looking for available data, talk to community members to see what they think, or collect your own data

## RELIABILITY - CAN THE DATA BE TRUSTED TO BE CORRECT?

Reliability refers to the accuracy of the data.

- Has the research that produced the data been reproduced by other researchers?
  - Did other research studies get the same or similar results?
  - Research studies that result in similar data are said to "validate" or confirm the results of the other similar studies.
- How was the data collected?
  - Was it consistent with the mission/goals of the researchers?
  - Did the researchers adhere to ethical research methods?
- Is there some kind of bias in who will reply? For instance:
  - Did they conduct their survey in different languages if they need information about immigrants?

**You'll want to review the methods used to produce the data and find other studies that validate the data you want to use.**

## TIMELINESS – IS THE DATA *STILL* USEFUL?



Community Alliance for Research & Engagement (CARE)  
at the Yale School of Public Health

Timeliness refers to when the research was conducted and how rapidly the population or condition is changing.

Some data may remain accurate for a long time, while others are only good for a few years.

- For instance, the West Nile virus was a rapidly moving illness for which only the most current data is useful.
- Often there will be a lag time, especially with big studies such as the Census.
- Most comprehensive surveys will be a few years old by the time they are published.
- Even if it is out-of-date it may be the best we have right now [ex.BJS]
- Admit the limitations of this kind of data and try to supplement it with other closely related research.

## EVALUATING DATA

- Bring your own knowledge and expertise!
- Determining the quality and integrity of the data is an art and science.
- You can use these criteria as well to assess the quality of data that you collect.
- No data is perfect. Use your own judgment regarding the use of data you think is defensible.
- Understand the assumptions that went into the data collection.



<https://www.measureevaluation.org/our-work/data-demand-and-use>

## PRIMARY DATA: DETERMINE IF YOU NEED TO COLLECT NEW INFORMATION

- Look at your research question worksheet – which data types *cannot* be gathered from secondary data sources?
- **These are the primary data you may need to collect**
- Before selecting data collection method(s), be sure you need this data to answer your questions



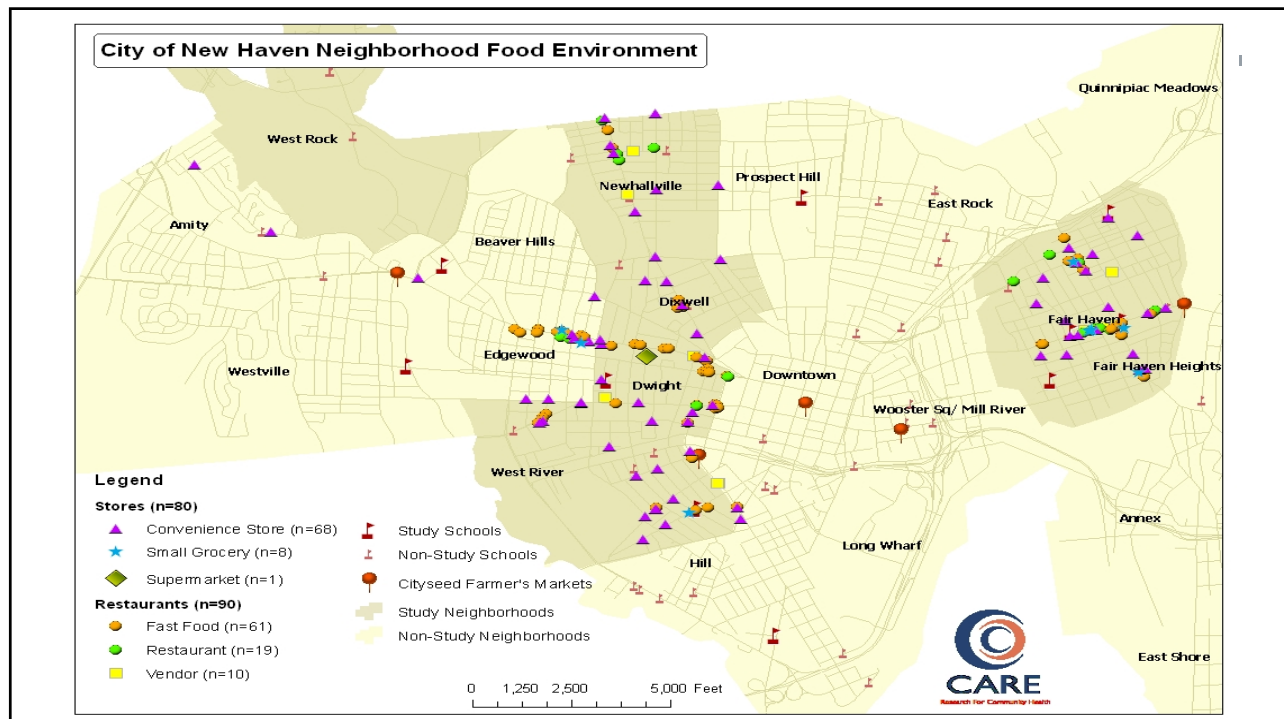
## SIX MOST COMMON DATA COLLECTION METHODS USED IN RESEARCH

- Asset Mapping
- Focus Groups
- Key Informant Interviews
- Surveys
- Community Forums
- Rapid Appraisal Techniques



## DATA COLLECTION METHODS: ASSET MAPPING

- An inventory of community health assets (resources, services, facilities, organizations, associations, programs, churches, schools, etc.)
- Usually represented by geographically mapped data
- Allows solutions to problem to identify gaps and build on existing assets
- Data can be used to develop, improve or advocate for additional resources or funding
- The process of asset mapping involves strong community involvement and promotes empowerment and ownership of a wide range of people
- May require a lot of time to survey community members or programs, to verify current addresses, and to keep it updated



## DATA COLLECTION METHODS: FOCUS GROUP DISCUSSIONS

- A series of discussions involving 8-12 people, selected to share their perceptions of a defined topic. Builds on group dynamic to stimulates participants to share experiences and opinions.
- Captures rich data in participants' own words
- Flexible to capture new ideas and issues
- Can be difficult to summarize and interpret results across groups
- Need to be creative when recruiting busy people
- Cannot be “representative” of the entire population
- Skilled moderator needed to keep respondents on track and make sure everyone gets to speak
- Bias may occur with unclear, leading, or loaded questions, or tone/body language of moderator



## DATA COLLECTION METHODS: KEY INFORMANT INTERVIEW

- A discussion conducted over the phone or in person during an interview
- Short answer or open-ended questions
- Allows respondents to share their opinions without the pressure of the group dynamic
- Allows interviewer to clarify questions and draw out thoughtful responses
- Important to select the “right” key informants so they represent diverse backgrounds and viewpoints
- Takes a while to administer; skilled interviewing needed to keep respondents on track
- Bias may occur with unclear, leading, or loaded questions, or tone/body language of interviewer

## DATA COLLECTION METHODS: SURVEY

- A questionnaire conducted over the phone, in person, online, or via mail
- Usually closed-ended questions (multiple choice, true/false or yes/no, brief numbered responses), but may include a few open-ended questions where participants can write or say what they choose
- Large amounts of data can be gathered from many respondents, and so can be used to “represent” the broader population
- Bias may occur with unclear, leading, or loaded questions, or tone/body language of surveyor

Neighborhood code \_\_\_\_\_ Participant ID \_\_\_\_\_ Net ID \_\_\_\_\_ Admin date \_\_\_\_\_

**Neighborhood**

**These first questions ask about your neighborhood. By your neighborhood, we mean the area AROUND your home that you could walk to in 10-15 minutes. We ask about your neighborhood to understand how it affects your health.**

1. How long have you lived in this neighborhood?  Less than 1 year  1 year  2 years  3 years  4 years  5 or more years

2. How long have you lived at this address?  Less than 1 year  1 year  2 years  3 years  4 years  5 or more years

3. Which of the following best describes the place in which you live?  I own it.  I rent it.  Live with family or friends  Other: \_\_\_\_\_

How much do you agree or disagree with each of the following statements?	Strongly disagree	Somewhat disagree	Somewhat agree	Strongly agree
4. Many stores, banks or ATMs, markets or places to go are within easy walking distance of my home.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. There are safe sidewalks and crosswalks on most of the streets in my neighborhood.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. There are places to bicycle in or near my neighborhood that are safe from traffic, such as on the street or on special lanes, separate paths or trails, shared use paths for cycles and pedestrians.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. Public transit can generally get me where I need to go.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. My neighborhood has several free or low cost recreation facilities such as parks, playgrounds, public swimming pools, etc.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. I see many people being physically active in my neighborhood doing things like walking, jogging, cycling, or playing sports and active games.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

## DATA COLLECTION METHODS: COMMUNITY FORUMS

- A series of public meetings focused on a defined topic
- Moderated to ensure that important topics are covered and time is used well
- Relatively inexpensive and easy to conduct
- Can also allow for educational content
- Allows for broad community and stakeholder participation in issue
- Allows for the gathering of many perspectives at once
- Participants may not be representative of the larger population, as those who attend may not reflect the entire community or target audiences





## DATA COLLECTION METHODS: RAPID APPRAISAL

- Observations and photography that can help to describe or visually depict community conditions. Observational tools or checklists and guiding questions for photos are used to create a research structure.
- Relatively easy and inexpensive
- Data can be gathered quickly
- Provide descriptions and visual imagery that give meaning to quantitative data
- Can be difficult to represent the entire community experience
- Can be difficult to interpret and summarize photographs and observation notes

## THINGS TO CONSIDER AS YOU SELECT YOUR METHODS

- What data collection method or methods would best collect the type of data you need, from the source you indicated
- Resources:
  - Time
  - Staffing
  - Skills
  - Funding
  - Computer technology
- Prior experience
- Potential benefits
- Potential challenges

Now, take a moment to think through the pros and cons of secondary and primary data, and the six data collection methods.

Decide which method or methods would be most appropriate for your PCOR project

- Use the research questions & data worksheet

## DATA INTO ACTION



**If your primary or secondary data makes a compelling statement, you can begin to formulate potential interventions and policies to address the needs.**

- Start with a brief report outlining the problem and what you found out about it.
- Include an asset map to demonstrate:
  - Resources already exist that address the issue
  - Identify spatial or geographic gaps in services or resources
- Disseminate the information
  - Make sure the brief and/or maps are user-friendly and in lay language
  - Organize a community forum to discuss findings and solutions



**FROM THE HEALTH D.A.T.A PROGRAM – DATA, ADVOCACY, AND TECHNICAL ASSISTANCE – FROM THE INTRODUCTION TO HEALTH DATA WORKSHOP**

Curriculum: <http://healthpolicy.ucla.edu/programs/health-data/trainings/Pages/health-data-workshop.aspx> Photo: <http://healthpolicy.ucla.edu/Pages>