

Health DATA Agenda

Session	Schedule
Session 1: Understanding Data <input type="checkbox"/> Commander DATA's Data Game	12:15 - 1:00
Session 2: Determining the Data You Need <input type="checkbox"/> Discussion <input type="checkbox"/> Debrief	1:00 - 1:30
Break -----	1:30 - 1:45
Session 3a: Finding Data <input type="checkbox"/> Exercise <input type="checkbox"/> Debrief	1:45 - 2:30
Session 4: Presenting Data <input type="checkbox"/> Discussion <input type="checkbox"/> Presentations <input type="checkbox"/> Debrief	2:30 - 3:30
Session 5: Ethics and Politics of Data <input type="checkbox"/> Group Discussion	3:30 - 3:45
Evaluation and Conclusion <input type="checkbox"/> Conclusion <input type="checkbox"/> Post Test <input type="checkbox"/> Evaluation	3:45 - 4:00

**Health DATA--Data Advocacy Technical Assistance
Introduction to Health Data Workshop**

**Curriculum and
Trainers Manual**

Developed by
The Health DATA Program

Of
The Center for Health Policy Research
University of California Los Angeles - School of Public Health

Funding for this project was made possible by The California Endowment

Introduction

Health DATA Workshop

Goals:

This session is designed to introduce participants to the mission and resources of the Center for Health Policy Research, including CHIS, and the Health DATA Program. In addition, the pre-test will give provide information about participants' prior knowledge of the topics covered in this manual.

Timeframe:

Activity	Equipment	Time	Scheduled Time
Introduction	<ul style="list-style-type: none"> ○ Sign in Sheet ○ Name Tags ○ “Commander DATA” ○ Health DATA Brochure/Flier ○ Training Manuals (one per person) ○ Photo Release Forms (one per person) ○ Training Agendas (one per person) 	10 min.	9:00-9:10
Pre-Test	<ul style="list-style-type: none"> ○ Pre-test Evaluation Form 	5 min.	9:10-9:15
Participant Introduction		15 min.	9:15-9:30
Total		30 min.	

Introduction:

Attending staff members introduce themselves. Staff members discuss the Center, the California Health Interview Survey (CHIS), and the Health DATA Program.

The **Center for Health Policy Research** applies the extensive expertise of faculty and researchers from UCLA's leading professional schools and academic departments to meet the needs of the nation, states, and local communities for health policy-related research and information. The Center conducts research on national, state, and local health policy issues, provides public service to policy makers and community leaders, and offers educational opportunities for graduate students and post-doctoral fellows.

The research, service, and education programs of the Center emphasize a community-and population-based perspective to improve health outcomes. This perspective focuses on: the

health status of population groups, systems of health care, and community-based interventions to prevent disease and promote health.

The **California Health Interview Survey (CHIS)** is the largest state health survey ever conducted in the United States, collecting information from 55,000 households drawn from every county in the state. Focusing on public health and access to health care, CHIS will provide:

- Statewide estimates for California's overall population, including important information not previously available on an array of racial/ethnic populations; and
- Local-level estimates for counties with populations of 40,000 or more. This will generate important information for local planning, and enable comparisons across counties.

Estimates for most counties will include breakouts by age group and, if the sample size is sufficient, by ethnic groups. The CHIS researchers have worked with officials in several counties to increase their sample sizes in order to improve the precision of the county-level estimates. CHIS will survey the state's population every two years.

Essentially, CHIS will provide everyone with easy to access data on the health status of Californians. Something you might all keep in mind as you prepare to use health data in your work.

The **Health DATA Program** is the largest public service program of the UCLA Center for Health Policy Research. The program was started because we wanted to help organizations gain influence and power in the policy-making arena. What do I mean by policy? One way to think about it is by looking at a board game. The rules on the inside of the box that dictate how many players, how many times you role the dice, etc. – that is policy. In the real world, these are the laws, regulations, and practices that citizens are expected to follow. Not everyone gets to decide what these rules are, but we want to help you become a more integral part of the policy making process so that you and the people you represent can affect how the rules are made and in turn, empower your constituents.

As a research center, we think that we can best help you by supplying you with the information or data that you need to further your mission. The kind of information that can help you to develop programs, target funding agencies and advocate for those you serve. For us, our expertise is in data, especially the kind that is scientifically measured, such as with a survey or a questionnaire. Our experience is that this kind of information can play a pivotal role in health policy, in particular because good quality research is considered objective and is often accepted by those who have the ability to affect health policy. Data can change people's minds about an issue:

- For example, a few years ago, a study was conducted to measure the cost effectiveness of prenatal care (\$1 spent on prenatal care yields \$3 of cost savings). The data were so convincing that now prenatal care is considered a very cost effective service that should be available to everyone. It changed the question from "Can we afford to?" to "Can we afford not to?"

- Data can also help to prove that a problem exists. The Latino Health Coalition conducted research into the health status of Latinas and issues their findings in a report. The research uncovered disparities in service to Latinas. Because of that report, political attention focused on the health issues of Latinas in California and now politicians see those disparities as a problem that needs to be addressed – especially given the fact that the number of politically active, or voting, Latinas has increased.

We want to increase your ability to use data and research to further your goals and work in much the same way.

Today we focus on five topics:

- Understanding data,
- Determining your data needs,
- Finding data,
- Presenting data, and
- The ethical and political aspects of data.

We are also going to try to have some fun along the way. But first...

Pre-Test:

Take five minutes to have participants complete the pre-test. Let them know that they will not be graded but that they will have an opportunity to re-take this test at the end to find out how much they gained from the workshop. Do not reveal the correct answers.

Participant Introduction:

Have participants introduce themselves. They should provide their name, the name of their organization, and maybe one of the following:

- An interesting statistic about themselves,
- Their reason for coming to the workshop
- A situation they have encountered where they were confused or misled by data.

Introduce Commander DATA (the Star Trek character), our mascot.

- Does anybody know who this is?
- Are there any Trekkies out there?
- If so, what is his character like?
- Do you have a favorite scene in which he appears?

Commander DATA is a character on the TV show and movies of Star Trek: Next Generation. He is an android, meaning he is a machine, who can process lots and lots of information. Yet because he is a machine, he cannot experience human emotion, and is constantly trying to understand what emotion is – acting lessons, stuff like that. My point is that data, like Commander Data, can describe human experience, but is not a replacement for it. Yet we can

use data as a tool in our work to improve the human experience – and that’s what has brought all of us here together today.

The Health DATA binders contain much of the information that we are sharing with you orally, and lots more that we don’t have time to talk about, but that you can refer back to when you need it. So don’t go crazy taking notes, we’ve got it all here for you. Okay, are we ready to start?

Critiquing Data

Goals:

This mock game show will provide participants with the opportunity to learn the vocabulary necessary to understand and evaluate the integrity of data. In addition, participants will be asked to critique data based on the credibility of the source, specificity of the data, applicability of the data to their own work, reliability, and timeliness. This will allow participants to identify gaps in data sources and determine the appropriate uses and limitations of health data when applying it to their policy and advocacy work.

Timeframe:

Activity	Equipment	Time	Scheduled Time
Commander DATA's Data Game	<ul style="list-style-type: none">○ Bellhop's bell○ Training Manual - Glossary○ Vocabulary board○ Flip chart or butcher paper○ "Thumbs Up"/"Thumbs Down" signs○ Prizes	45 min.	9:30-10:15
Total		45 min.	

Commander DATA's Data Game:

Divide the participants into 2 teams. Each participant should be given a thumbs up/thumbs down sign. The bellhop's bell will be placed on a central table where the host (the trainer) and Commander DATA, will be standing.

The host, reads the "answer" (a vocabulary word definition) and the contestants ring the bell for a chance to give the appropriate vocabulary word, in the form of a question. If the person who rings the bell gets it wrong, the opposing team is offered the chance to answer. If neither contestant can answer the question, Commander DATA will provide the correct answer.

After each question the host will provide an example of data. All participants are to decide on the integrity of the example given. If the data he/she presented is usable, they give it "thumbs up," if it is unusable, they give it "thumbs down." A brief discussion should follow each round so that the participants can understand some of the ideas they need to consider when analyzing data.

As the game is played, the facilitator should stress the Center's criteria for evaluating data (see debriefing notes). Ethical issues about data can also be raised in this activity.

For an example game scenario see Appendix A.

&
For vocabulary list see Appendix B.

Notes:

1. Credibility – Who produced the research/data?

- Who paid for, sponsored, or funded the study?
- How much of a stake do they have in a specific finding? What about their mission or constituents would motivate them to want the data to have a certain? Would this affect their interpretations of the data?
- What is the organization’s public image or reputation for their research?
 - Government studies and academic institutions tend to be considered most credible.
 - Research affiliated with or conducted by business groups, religious organizations or political organizations may have missions that affect the outcome.
- Every group will be called biased by their opponents. Research that is considered credible is not always conducted with an objective standard either.
- Sometimes you may want to use information of groups who share your values or mission because they are asking the questions that are probably similar to yours.

2. Specificity - What was collected?

- Refers to content of the study, does it really measure what you are after?
- Surveys are often seeking to answer specific questions or are looking for something in particular. How close is the relationship between what you need and the data they provide?
- What can you do if the information you have found isn’t exactly what you were looking for? For instance, you are looking for information on the number of Caucasian women who smoke. Statistics are available on the number of Caucasians who smoke and the number of women who smoke, but nothing specific on Caucasian women. Can you extrapolate from the data you find to get the answer to your question?

* Note: this is close to “generalizable”, but specificity refers to the information collected and generalizable refers to population

3. Applicability – What was the sample population?

- Refers to the population from which the data came.
- If the study was done in North Carolina, does it also apply to people in California?
- If it was a national study, can we use the same numbers for Orange County?
- If the study looked at Latinos, is there anything particular to say about Central Americans as a distinct group?
- How much can you “generalize” data to your constituents or service population? It is important to know the “who, what, why, when and where” of the data because it will determine how closely it matches your own need for data.
- Watch out for over sampling.

4. Reliability - How was the data collected?

- Reliability refers to how bias-free the study is constructed internally.

- Was the survey or research replicated? Did the researchers get the same or similar results?
- How was the data collected? Was it consistent with the mission/goals of the researchers? Did the researchers adhere to proper research protocols?
- Is there some kind of bias in who will reply? Did they do a telephone survey, calling households between the hours of 9 and 5, when they were looking for people with employer-based insurance? Did they conduct their survey in different languages if they need information about immigrants?
- Mail-in surveys traditionally have poor response rates, be aware of the validity of results with low response rates.
- Reliability can often be determined by the reputation of the group conducting the research.
- Were proper research methods used? (i.e. random sampling, focus group consistency, etc...)

5. Timeliness – When was it collected?

- 10 years ago? (i.e. 1990 Census) 1 year ago? 3 years ago, but there was a major change in program eligibility during that time? (i.e. welfare reform)
- Clearly, the more current the information the better – but keep in mind other factors when looking at data.
- Often there will be a lag time, especially with big studies. Most comprehensive surveys will be a few years old by the time it is published. Remember too that even if it is “old,” it may be the best we have right now because it has not been reevaluated recently. Admit the limitations of this kind of data us and try to supplement it with other closely related research.

Concluding Points

- The data you use to support your arguments will be based on these criteria as well. You are bringing your own credibility to the data, so it is important to set standards for yourself with regard to the data you plan to use.
- No data is perfect. Use your own judgment regarding the use of data you think is good enough and defensible, and you will have a rationale as to why it is. Knowing the assumptions behind the data you use will hold you to a good standard and build your credibility for being a source of information for policy makers.

These are criteria by which researchers judge their data and the studies of others. You can use these guidelines:

- To assess the quality of data that you collected,
- To assess the quality of other sources of data, and
- To critique the data of your opponents.

Appendix A

Examples for DATA's Vocabulary Game

Data (host): *Today's research topic is Access to Health Care. I'll give you a definition in the form of an "answer" and you give me the definition that corresponds to the "answer" in the form of a question. I will then give an example of the word and you, the audience, will tell me if the information is usable (thumbs up) or not usable (thumbs down) by holding up the appropriate side of the sign that is in front of you. By telling me whether or not the data is usable we will build a list of criteria for how to judge the quality of data. Okay, let's practice by going through an example first. The answer is.... "this tells how frequently an event is occurring and is usually expressed as 1 in 100,000."*

Contestant: "What is **rate**?"

Com. Data: *Right. The example is, "The reported incidence of AIDS in Los Angeles County for 1996-1998 was 22 cases per 100,000 people. This information was based on a ranked 3-year average of crude case rates using 1997 population. Is this information usable or not?"*

Audience: *usable (thumbs up)*

Com. Data: "We will then discuss why the data is usable or not usable. Does everyone understand?"

Com. Data: "Good. I think we've got it. Let's begin. We will be keeping score and the team with the most points will win a prize. So, let's begin. Can be referred to as the 'science of finding out'."

Contestant: "What is **methodology**?"

Com. Data: "That is correct. A point for Marie. (Turns to the audience). According to one study there are 1.85 million uninsured children in CA. They based this information on the March 1998 Current Population Survey. Is this information usable or trash?"

Audience: "Usable" (Raises their signs)

Com. Data: "Why?"

Audience: "Because it was done recently." "It uses information from the CPS."

Com. Data: "What if they used information from 1988?"

Audience: "It would be trash." "Too old."

Com. Data: "Would you still use the information today due to changes with welfare reform and the implementation of the Healthy Families Program?"

Audience: "Yes." Identify reasons. Has actually increased in the recent years as a result of these programs.

Com. Data: "Good. So timeliness, or when the data was collected, is one criteria for judging how 'good' the data are. What else?"

Audience: "This is the number that is still used today because it is the most recent that we have."

Com. Data: "And we can use the most recent numbers even if a little old because that is the best that we have. And, others are using the same number."

Com. Data: "Okay, next answer. The subset of people selected in some prescribed manner for study."

Contestant: (ding) "What is a **sample**?"

Com. Data: “Right.” A point for Jesse. (Turns to the audience). A tobacco company released a study of 1000 smokers’ families and found second hand smoke is not a health risk.”

Audience: (Raising their signs). “Trash”

Com. Data: “Why?”

Audience: “Because they only looked at 1000 families and that is too small a sample size.”

Com Data: “Okay, so sample size is important...we cannot make a conclusion based on 1000 families, or can we?”

Audience: “Not when all the other information says the opposite, and your study is the only one saying this, you need more evidence.”

Com. Data: “What else should we be asking?”

Audience: “How they got the sample and how representative it is.” “We need to know if there are biases in the way the sample was collected.”

Com Data: “What if we do not know that and they do not provide us with that information? What can we do?”

Audience: “Look at who the information is coming from, where the study was conducted.”

Com. Data: “Why does that matter?”

Audience: “Because I am less likely to trust how they got their sample b/c they have a clear stake in producing facts that say second-hand smoke is not a health risk so that people will keep smoking.”

Com. Data: “Good. So we can say that there information is not reliable b/c it may have sample errors and biases that sway the results. Most reliable data will tell you their limitations so you can decide if it is good data or not.”

Com. Data: “Okay, the next answer is, using data in a sample to make an enlightened guess of a population value.”

Contestant: (ding) “What is an *estimate*?”

Com. Data: “That’s right. A point for Jesse. (Turns to the audience...). DHS estimated that in 1998 about 4.0 million persons utilized Medi-Cal, based on their administrative data for people using Fee For Service (FFS), or enrolled in Health Care Plan’s HCP’s, or County Organized Health Systems COHS). They estimate that 4.96 people are eligible for Medi-Cal. Is this information you would use, or is it trash?”

Audience: (Raising signs). “Usable”

Com. Data: “What are you trying to measure, what question do you want an answer to if you are DHS?”

Audience: “How many people were on the Medi-Cal program that year.”

Com. Data: “Why does DHS want to know this?”

Audience: “To allocate money for the program next year. To compare with other years and look at trends. To look at how many people Medi-Cal insures and how many people are uninsured.”

Com. Data: “Based on those reasons are they trying to include as many people as possible? Would they include people that only used Medi-Cal for a single emergency use?”

Audience: “Possibly, to try and get more funding.” “To show that there are less people without insurance than there actually are.”

Com. Data: “Right. So the Center for Health Policy Research estimates about 3.29 million were covered by Medi-Cal using CPS data. Now, whose number would you be more apt to use?”

Audience: “Not sure” “You would have to look at the study and see how they got their number and how DHS got their number.”

Com. Data: “Good. Why is that important?”

Audience: “Because if DHS uses administrative data and counts every person who has used Medi-Cal even for a single incident it may over count those on Medi-Cal who have a regular source of insurance when they do not.”

Com. Data: “What do you want to know about the Center’s number?”

Audience: “Where the number came from, what data source they used, how they got the information?”

Com. Data: “What measures they are using is important. When you look at the data you want to know if they are measuring what they say they are measuring, this is called specificity. Now, which number would be more usable?”

Audience: “Well I would use the Center’s number because I consider those who use Medi-Cal once as still uninsured and with the work that I do we need a truer account of how many are uninsured.”

Com. Data: “Good. So that number and how it is gotten is more specific to your needs. They are measuring what they say they are measuring.”

Com. Data: “This type of research aims to collect facts and figures, information and data using methods such as the census, surveys, structured questionnaires, and interviews.”

Contestant: (ding) “What is **Quantitative** (Data)?”

Com. Data: “A community based organization working to increase Healthy Families and Medi-Cal enrollment has just released quantitative data from a survey they conducted in one zip code with 500 households. The data shows that the problem of uninsured children in this zip code is twice that than numbers for Los Angeles County as a whole. Is this information that you would use or not”

Audience: (Raising signs). “Usable” “Not sure”

Com. Data: “Why would you say that you would not use the number?”

Audience: “Because it was done with only 500 households and the organization has a vested interest to show there are more uninsured in the county.”

Com. Data: “Okay, so you may not see them as a credible source of information. Whereas you may see the information done by the county DHS is more credible? Why?”

Audience: “Because it is done with a larger sample size and by the county for the entire county. And I think that is more credible data.”

Com. Data: “Great. But is poverty and uninsurance distributed equally in the population here in Los Angeles County?”

Audience: “NO”

Com. Data: “Somebody said they would use this information, tell me why.”

Audience: “Because that may very well be the case in that neighborhood and the organization that sees the people knows it is so, that is why they found that information.”

Com. Data: “Good. So by saying that this data is credible because the group producing it has worked in this neighborhood for the last ten years and has experienced this reality would work. What else could you do to add credibility?”

Audience: “Make sure that people know this information is only speaking for this neighborhood and not the whole population.”

Com. Data: “Good. We may not be able to say this goes for all children in the county, but for this neighborhood, definitely. With that in mind you could also show other data from the Census or state Department of Finance with statistics on the socioeconomic conditions of this neighborhood that show it is different from the county as a whole. And information on how poverty/working class people have higher rates of being uninsured in the population.”

Com Data: “Next answer. These data are named not numbered and extensively used to get a more in-depth understanding of a situation. Usually done through focus groups, case studies, and observations.”

Contestant: (ding) “What is **qualitative** (data)?”

Com. Data: “Correct. A point for Marie. (Turns to the audience...). The UCLA Center for Health Policy did focus groups with low-income women and found that lack of information due to language and culture were barriers to quality care. Is this information that is usable or trash?”

Audience: (Raising signs). “Usable” “I would use it because I hear the same thing every day from people I see every day.” “Well, I would want to back that up with some numbers.” “How many women were in the focus groups?”

Com. Data: “So, what does the information tell us right off?”

Audience: “That low income women are not accessing care because they may not speak the same language as the doctor.”

Com. Data: “Does this hold true for all low-income women?”

Audience: “No, because not all low-income women may see a doctor that speaks a different language.” “Not all low-income women do not speak English.”

Com. Data: “For those women that are most likely not to speak English, such as Latinas and Asian Americans can I say that language is a barrier to accessing care.”

Audience: “You may not be able to say for all but you may be able to generalize for most, at least that is what I see.”

Com. Data: “Good. The biggest question to ask with qualitative data is how much can I say from this information to generalize to the whole population. Generalizability is a big factor with qualitative research. But if the target population of the study matches the demographics of your population then you should use this information and remember that you are not making assumptions for the whole of a population but for most. This helps to paint a picture of what is going on and why low income women may not be accessing health care. This may also get others motivated to start collecting or demanding this information in a research study. Focus groups often allow us to describe what the problem is for further study.” Want to incorporate Delight’s comment about taking the bridge over-should we have a visual here?

Com. Data: “Okay. The answer is, A particular incident or fact of an interesting nature.”

Contestant: “What is **anecdotal** (information)?”

Com. Data: “Right. A point for Jesse. (Turns to the audience). Before the state assembly a mother testified in the interest of a bill how the lack of insurance for her son with chronic asthma was taking a financial and personal toll on the family. The family cannot afford to pay for the inhaler and medication that the son needs, he is in and out of the emergency room, and has not been diagnosed with a learning disability. Is this information usable or trash?”

Audience: (Raising their signs). “Can be usable”

Com. Data: “Why is that?”

Audience: “Because people like to hear these personal stories. It does a lot to add sympathy and put a face with the numbers.”

Com. Data: “Good. But what does this one mother’s experience say for the rest of the population?”

Audience: “Well, it may not be asthma but many people without health insurance that I have seen have similar stories about how not having medical coverage has negatively affected their family, themselves, their kids.”

Com. Data: “Okay, so we may be able to use this information to highlight people’s experiences. But can this one incident speak for all people without health insurance?”

Audience: “Probably not. But we are trying to make a point and draw on people’s sympathy and emotion for support so this makes sense.” “And it is based on a truth and situations that many people face who do not have health insurance, especially among low-income populations asthma is a real problem.”

Com. Data: “Good. So we know this is not generalizable but may enhance our message, and they are powerful messages that motivate people to act. What else can you do with this type of information to add credibility?”

Audience: “You can show how this experience is not isolated or show that it is common.”

Com. Data: “Right. And we can do that by adding some quantitative information to show the extent of this problem in the community. This will lend credibility and generalizability to the issue. Well, the star has the most words on her side so she wins a lovely t-shirt.” Show the shirt and thank them for participating.

Appendix B

Vocabulary List

- Estimate (n)
- Qualitative
- Quantitative
- Population
- Anecdotal
- Methodology

In Binder inserts:

- Case study
- Focus groups
- Population
- Over sampling
- Survey
- Questionnaire
- Random
- Variable
- Cross-section
- Longitudinal
- Sample
- Cluster sampling
- Response rate
- Reliability
- Credibility
- Rates
- Range
- Median
- Mean vs. Average
- Correlation
- Causation
- Indicator
- Statistical significance
- Percentage
- Proportion
- Specificity
- Generalizable
- Timeliness

Determining the Data You Need

Goals:

During this popular-education-style activity participants will “map-out” their organization and its constituents. Through exploration and discussion, participants will identify the “web” of people, organizations and governmental entities that affect the political climate and, in turn, the activities of their organization. This exercise will help participants learn to identify relevant and appropriate data and information sources that are necessary for them to actively participate in policy changes that affect their organization.

Timeframe:

Activity	Equipment	Time	Scheduled Time
Assignment		10 min.	10:15 – 10:25
Exercise	<ul style="list-style-type: none"> ○ Charts: <li style="padding-left: 20px;">What do you need the data to do? <li style="padding-left: 20px;">What message will the data deliver? <li style="padding-left: 20px;">What kind of data does your audience need? ○ Butcher paper ○ Markers ○ Color legend 	20 min.	10:25 – 10:45
BREAK	-----	10 min.	10:45 – 10:55
Group Exercise		25 min.	10:55 – 11:20
Debrief		40 min.	11:20 – 12:00
Total		105 min.	

This exercise will utilize the three charts noted in the timeframe above. The first, “What do you need the data to do?” will help participants clarify the kind of data necessary to meet their organizational GOALS. The second, “What kind of data does your audience need?” is intended to help participants craft a convincing argument based on their target AUDIENCE. Finally, the third, “What message will the data deliver?” will instigate participants to consider the VALUES that must be addressed in order to motivate their intended audience – persuasion/motive.

Assignment

Before beginning a search for data, it is important for you and the other staff at your organization to outline and define the questions you need the data to answer. This information will help you

identify appropriate data and save you a lot of time. As a research center we often get requests for data, and many times the people who call us have only a vague idea of the type of data they need.

Briefly go over the three charts, “What do you need the data to do,” “What message will the data deliver,” and “What kind of data does your audience need,” to give the participants a general idea of the topics that will be covered in this session. Do not spend too much time, as these concepts will be covered in greater detail in the group exercise.

Exercise

If there are a manageable number of participants, this exercise can be done individually or in groups. If the participants are from different organizations it might be useful to let them do the activity individually. If participants are from the same organization, groups may provide them with an opportunity to discuss organizational needs and plans that relate to the exercise.

Pass out a sizable piece of butcher paper to each participant or group. Each person or group should draw themselves in the context of their work. They could include:

- People from their service population
- People who they wish they could recruit into their service area
- Other community based organizations that provide services similar to theirs
- Community based organizations that work in conjunction with them
- Local government agencies with whom they interact
- Barriers to service – linguistic, cultural, ethnic, etc...
- Political figures who have an impact on their work
- Funding Agencies

After about ten minutes, have a few participants share aspects of their drawings.

- What did they include?
- Why did they choose to include these items?
- Were there things they thought they should include, but did not? If so, why?

Tape a couple large sheets of butcher paper to the wall. Select one of the pictures and position it on the butcher paper so that there is room around the picture to draw around it. Utilize this drawing to ask the participants about their work, their constituents, their funding, and the problem in ways that will slush out some of the other aspects of their organization that they may have failed to include, such as:

- Who else might agree with you? Who else has a stake in this scenario?
- Who are your opponents?
- What are the other pressures on the local government agency? Who or what influences the local government agency?
- What other steps could your organization take to influence this policy making body?

- If your community (i.e. Los Angeles) does not provide enough services and your organization is overextended in terms of your service population, what is the impact on neighboring communities (i.e. Long Beach, Riverside and San Bernardino)?
- What is the public’s opinion on this issue? How were they influenced to believe this?
- Who will be opposed to your plans and why?
- What do we have to do to convince the decision makers to spend the money on the problem we have laid out here? Who else might care about this?
- Are there coalitions established that could help with this issue? Who else might be brought in on this?
- What can we do to learn more about these organizations? What influences them? What other pressures do they face?

The following discussion will supply participants with a framework from which to begin the process of determining the data they need.

Group Exercise

Now that you have “mapped out” the players who will have an impact on your organization’s activities, it is time to determine the kind of data you need in order to address the various needs of these people and organizations. What arguments would you make, and what information would you need to support your argument. Be prepared to:

- State the problem accurately and dramatically;
- Show how your solution would alleviate the problem;
- Illustrate the negative consequences of not doing it;
- Refute arguments the opposition will use against you.

As you think about what information you need to support your mission or goals, there are three things to keep in mind:

- What message will the data deliver?
- What do you need the data to do?
- What does your audience need?

What message will the data deliver?

- cost
- quality
- access
- equity
- rights

What is the message that the data will deliver? What will motivate the decision makers into action? Different things will convince different people. For some, the statistics alone will motivate them; for others a more personal story of overcoming hardships will be the key. In either case, the statistics and the stories are most effective when they appeal to someone’s values. Remember what I said about policy being like the rules of a board game? People who write

those rules are putting into place their values, their beliefs about what would make the game fair. Their values determine “how much money you start with in the game” and “how many turns you get to take when you roll double sixes.” Your job has been to get them to use their power to put your values into those policies that set the rules of the game. If you can supply data that not only accurately describes what people experience, but also data that motivates them into action by appealing to their values or belief system, then you have some very powerful tool with which to achieve your goals. Data can be more than just information; it can be a convincing way of conveying your values and vision of society.

Your data should also appeal to people’s logic by convincing them of your argument. You may often need data to do the following:

What do you need the data to do?

- Define the problem
- Show that your solution alleviates or solves the problem
- Show the negative consequences of not using your solution
- Measure the outcomes

At any given time, you may need the data to do all these things, some of these things, or just one. Often, though measuring the outcomes is a form of evaluation and a way to hold someone accountable for their actions or inactions or to prove that your solution alleviates or solves a particular problem.

What data does your audience need?

- elected officials, juries, media, general public – overall perspective (the forest)
- committee staffs, judges, special interest groups with legislative analysts – more specific information, other impacts (individual trees)
- agencies, submitted testimonials for the court, academics – details, things not on the surface (acorns and roots)

In our experience, the data you need not only depends on what you are trying to do, but also whom you have to convince. Some of this depends on the amount of time you have to convince them of your position or the depth of the information that you decide is appropriate. In general, the level of complexity you decide upon when presenting data is dependent upon the type of people you are trying to convince.

- The Forest—the Big Picture – Politicians, the public and the media are audiences who tend to need a piece of information that tends to be very descriptive and easy to understand; is often an overall perspective – from a big picture point of view. For example, in California, a child is more likely to die from a gunshot wound than in a automobile accident (Californians for Responsible Gun Laws).
- Individual Trees—Some Detail -- committee staffs, judges, and special interest groups with legislative analysts tend to need more detail than the big picture. These individuals want to know what kind of trees are in the forest or how many trees per square acre. They may want to know if they clear this part of the forest, what does that do to the ecosystem, etc. This

information will have more layers to it; often the audience understands the general ideas, but does not understand the details. For example, homicide is the number one cause of death for people between the ages of 15 to 19 in California, men account for 78% of these fatalities (OSHPD).

- Acorns/Roots—Specific Details – Government agencies, court officials, and academic institutions often need data to be more academically focused or statistically driven data to understand and critique. This type of data may require a high degree of accuracy because funding or planning decisions will be made based on those numbers, so it needs to be reliable because it will be scrutinized carefully. Like roots, they want to know details that you would not see clearly on the surface. For example, at a mean medical cost of about \$17,000 per injury, the 134,445 gunshot injuries in the United States in 1994 produced \$2.3 billion in lifetime medical costs. That figure is in 1994 dollars and uses a real discount rate of 3%. (JAMA 08/04/1999).

Have the groups spend about 15 minutes thinking about these concepts in terms of their drawings and the needs/goals of their organization. They should jot down thoughts and questions they have, these will be addressed when the group discusses this activity in the debriefing.

Debrief

Each groups should present their arguments and their data needs. Test their arguments with the rest of the participants and see if people think those arguments seem appropriate. Test it with different audiences and see if the values message and/or the data needed would change. If time, allows, you might suggest some places participants could go to find the data they need. Talk about the different data needs of the audience.

Debrief Questions

- With each group, go through their arguments first. So what were the arguments you came up with? Did anybody else have a similar argument? Did anyone use something different?
- What was your thinking behind your argument? Why did you choose to focus on that?
- Would that argument work with a different audience, such as...? How would it need to change?
- What data are you going to need to convince them?
- Where will you get this data? How would the data change with a different audience?
- What data is your opposition using? Where are they getting it? What's wrong with it?

Finding Data

Goals:

This interactive exercise will teach participants to think critically about the process of finding data. As students collect, summarize, and interpret information relevant to the public health problem they are given, they will also encounter some of the real dilemmas that confront those seeking data, such as time constraints, differing data, difficulties with funding and other barriers to finding the data they need. Through this exercise students will learn to narrow their data search and efficiently determine the most appropriate sources for their given data need.

Timeframe:

Activity	Equipment	Time	Scheduled Time
Exercise	<ul style="list-style-type: none"> ○ 10 Data Stations or “Data Providers” ○ Container for “Data Needs Scenarios” ○ Training Manual Localizing Data Worksheets ○ Chart, “How to localize data.” Organizational List 	25 min.	12:50-1:15
Debrief		45 min.	1:15-2:00
Break	-----	10 min.	2:00–2:10
Total		40 min.	

Possible titles for Data Stations/Data Providers:

- Research Centers: The Center
- Advocacy group: Latino Coalition for a Healthier CA, Families USA, Electronic Policy Network, League of Conservation Voters, etc...
- Governmental Agencies: Department of Health and Human Services, the Centers for Disease Control, The Environmental Protection Agency; the Department of Labor or smaller units from any of these agencies.
- State, County, or City Health Departments and Local Governmental Agencies
- Educational Institutions: Universities, Schools/ School Districts
- Community Groups: Economic Development Committees, Workforce Development Agencies, Chambers of Commerce
- Surveys: U. S. Census, CHIS

Exercise:

Have the participants split into groups of two or three. Pass the container of “Data Need Scenarios” around the room. Have each group take one.

The participants will be given 25 minutes to try and find the information and data necessary to answer the question they chose, unless of course they find information that tells them otherwise.

The possible scenarios participants may face:

- There will be some questions that are rather easy and you will only have to go to one location to find the answer.
- There will be some questions that will require participants to piece together information from a variety of locations. They will have several solutions but none that “perfectly” answer their question. This scenario can be used as an opportunity to discuss how to localize data, for example.
- There will be some questions that will require them to develop “proxies” for what they need.
- There will be some that address the frustrating issues of lack of time and resources. For example, one information source can provide participants with “set-backs” such as:
 - “will give you that information in 2 months;”
 - “This data is only available in publication, not electronically. Please send \$19.95 to us and expect 3-4 weeks for delivery;”
 - “your funding has been cut in half, lose five minutes of time;”
 - “I don’t have that but check with the CDC. Someone there should have that information;”
 - “Boss just came in to let you know that they changed their minds. Don’t need that statistic anymore;” or
 - “Server is down. Wait all afternoon.”

Debrief:

Begin by asking how many groups found exactly what they were looking for and ask how this process went (was it easy or hard; what were some of the difficulties they faced?). Their frustrations will probably focus on resources, access, and quality of data – depending on what was provided at the 10 data stations.

When trying to find data there are three themes that arise and frustrate the CBO in their quest for the “answer”. One is that **resources** are limited meaning that the CBO may lack the internal organization in order to fully utilize data; often the organization does not have the time, knowledge base, and/or people power to access the data. The other is **access**; the limited quantity or non-existence of certain data can be very problematic. In addition, getting much of the data can also be very difficult. The last theme related to finding data is the **quality** of the data. It can be difficult to determine the reliability of information, especially if time constraints are involved in the process. Also important to note is the issue of piecing together data from a variety of sources.

Resources:

- Had to go to more than one location, ran out of time, etc.
- “Bookmark” favorite web sites that are useful for you. Our binder lists the web addresses of some sites that may provide you with some help getting started.
- Plan to spend time looking for data when you do a campaign. Be realistic about how much time this is going to take you. Getting the data you need can take weeks, or months. You should plan accordingly.
- Had to go to more than one location to find the information and ran out of time. This is a frequent problem in scenarios such as this; other demands often blur one’s ability to critique statistical information.

Access:

- The data is not accessible for the organization because administrative data is not in format you need for policy goals.
- Are you not getting the answer because no one is asking the question? How can your organization and others advocate for change? Who are natural allies to your advocacy cause? Where would you receive resistance?
- Found “raw” data but don’t know how to convert that data into something that is comprehensible. Look primarily for the organizations that publish user-friendly data. Otherwise you may be wasting your time.
- Does anyone have access to this information? Is there a broader answer or reference to be made from the question? For example, if you have a question that calls for a specific piece of information to be gotten about a county and you cannot find that information but can find the information for the state, what can you do in this situation? (Localizing Data) Can you have that information ran just for your county? Who would be able to do that for you? How similar is your county to the demographics of the state and can you document that, or state versus country, etc.?
- What are some other possible factors that would assist in helping you to make your point even if you do not have access to the “correct answer”? Discuss ways in which a proxy could be used to find the data needed to answer our data inquiry.

Quality:

- Did anybody check the quality of the data they found? How can you be sure that the information you pulled together is a “good” answer to your question?
- Is it okay to piece together information from various sources? What does that say about the quality of our data? What are some situations where it would not be advised to piece together information from various data sources? How can you be sure that the information you pulled together is answers your question?
- Are there any markers or “rules” that we could use to measure the reliability of this approach? Where has this been done effectively? It is okay to make inferences as long as you are using credible sources.

Methods for “Localizing” Data

Then ask about what happened with the groups that didn’t find what they were looking for. What did they find instead? Refer to chart, “How to localize data.”

How to “Localize” Data

Method #1 – Make Your Own Estimate

Method #2 – Paint a Picture

Method #3 – Ask a Researcher

Method #1 – Make Your Own Estimate

How-to examples are in the binder, but basic steps include:

1. go to the lowest level of the most reliable data you have (state, maybe)
2. compare that level’s demographics to the demographics of yours, notice what is obviously different
3. are the differences likely to be related to the thing you are looking for? (you may need to make some reasonable assumptions about those differences)
4. may be able to use the information you have to calculate your own estimate
 - You can do this, but remember to protect your credibility when you put out a number (are you sure you want to do this?)
 - What will the estimate be used for? Is it appropriate to estimate in certain instances?
 - Sometimes no number is better than a bad one; sometimes a fuzzy one is better than none. You must decide.
 - Be prepared to defend the information you use and methods.

Method #2 – Paint a Picture

- What other information can you get that would come close to describing exactly what you want? Is that information easier to get?
- Is what you want hard to measure? What other things can substitute for it or approximate it?
- Not having the exact number is not the end of the world. Sometimes you can paint a better picture without it.
- Sometimes these substitutes are called “social indicators,” “proxies,” or “surrogate markers.”

Method #3 – Ask a Researcher

1. If you liked their study, sometimes they can give you an estimate based on their data.
2. If their study was close but not quite, sometimes they have data that did not get published and they may have what exactly what you need.
3. Sometimes they have it but are restricted from providing it because of a confidentiality agreement, which is a contract or ethical standard that is designed to protect the identity of the people who participated in the study. If you ask for something very local, it is sometimes possible to identify the actual person that data came from and that would jeopardize their privacy.

- Asking a researcher can begin to establish a longer-term relationship with him or her. It may pay off in the long run.
- Be persistent. They are busy and may require constant, polite prodding.
- Expect that it will take time, many researchers have moved on to their next discovery by the time data from their last experiment becomes available to the public.
- When you do get data this way, pay attention to any caveats the researcher places on the data. These are the limitations of the data you are receiving, the assumptions behind them, etc. You will need to know this to understand the quality of your data.

Presenting Data

Goals:

After this role-playing exercise, participants will gain hands on experience utilizing the knowledge they gained in a previous session on *Determining the Data You Need*. This exercise will teach participants to effectively present accurate demographic, statistical, programmatic and scientific information to professional and lay audiences. Specifically they will present data, as well as an accompanying values argument, to a mock audience comprised of their fellow trainees.

Timeframe:

Activity	Equipment	Time	Scheduled Time
Assignment		10 min.	2:10–2:20
Group Exercise	<ul style="list-style-type: none">○ poster board (one for each group)○ markers○ copies of the four scenario pages (A,B,C,D) following this session.○ large face timer/clock (optional)○ Training Manual – Chart Types	20 min.	2:20–2:40
Presentations		40 min.	2:40–3:20
Debrief		20 min.	3:20-3:40
Total		90 min.	

Assignment:

1. Divide the class into Four Groups (approximately 10 minutes per presentation)
2. Distribute one Scenario Handout to each group.
3. Read the instructions to the participants, remembering to stress
 - groups will have to present both a values argument and the data
 - groups should consider the audience they are addressing
 - everyone should be ready to share thoughts and comments with their peers about their presentation
4. Time permitting, facilitate a group discussion about the experience. (What did everyone learn?)

Group Exercise:

For our final exercise of the day, we are going to be role-playing. Everyone will be divided into groups and each person will play a member of a community-based organization, which provides public health advocacy to their community. Those who are not presenting at the time will serve as the person or persons to whom the presentation is directed. Each group will be given a hypothetical scenario and data for a public health problem confronting your service population. You will have ten minutes in your presentation, remembering who is in your and that all presentations must include a values judgment. After your presentation, the rest of the class will provide you with feedback. You will get some markers and poster board to make charts or whatever visuals you need. The idea here is to prepare an argument for an influential individual or group with a values-based bottom line that your data can support.

Presentations:

Notes to trainers for presentation:

- Raise questions about the quality of the data presented.
- Play the devil's advocate, argue about their values assumptions (i.e. if they make arguments about rectifying inequities, make scarcity and deserving/undeserving arguments – “data shows that immigrant families are much less likely to have employer based insurance, so this government program should be expanded to include them. It will save us money in the long run and it's the right thing to do.” “Yes, but taxpayers aren't willing to pick up their tab right now, and besides we should focus on taking care of our own.”)
- Do occasional annoying things – fall asleep, get interrupted all the time, be belligerent, etc.
- Fight data with data; get into dueling data situations.

Notes to trainers for specific scenarios:

Scenario A:

Scenario B:

Scenario C:

Scenario D:

Notes to trainers for feedback session:

- Make sure comments are not counter-productive. Keep them positive.
- Don't forget their presentation technique. Was their visual aide a benefit? Did they speak clearly and concisely? Were the arguments persuasive? Did they “fend off” your mock attacks?

Recommendations for Participant Feedback

- What were the most convincing things you saw or heard?
- What moved you? What did they seem to have trouble communicating?
- Focus some on how they presented the data. Was everything accurate? Were the visual aids appropriate given the audience? Did the presentation benefit from the visual aide(s).

- Where were the pitfalls?

Debrief

- Point out other aspects of their audience that they might not have considered.
- Were there other values arguments that might have been more effective?
- Did their data come from the best sources?
- Criticize constructively. Stress that their efforts were successful and draw on the positive outcomes.

Ethics and Politics of Data

Goals:

Guided by a short lecture on the topic, participants engage in a discussion of the ethical and political aspects of data. In doing so, participants will learn to apply ethical principles to the collection, maintenance, use, and dissemination of data and information about their own policy and advocacy work.

Timeframe:

Activity	Equipment	Time	Scheduled Time
Discussion	<ul style="list-style-type: none">○ Charts outlining the points of discussion: Define your code of ethics. Hold yourself and others to high standards. Hold researchers and policy makers accountable for their data.	10 min.	3:40 – 3:50
Total		10 min.	

Discussion:

Data can be viewed, as objective truth proved by science and/or as a rhetorical tool in the political process. For this reason, we want to spend a little time discussing the ethics and politics of data.

Sometimes we come across situations where we see data being used in a misleading way, for instance:

- people sometimes use good, quality data in ways that were not intended by the researchers; they may present only the parts of the data that serve their purposes
- people may also use “bad” data because it supports something they are trying to do
- people use anecdotal information, a story they know about, and generalize it, as if that were the experience of everybody in that same situation
- people use “science” to justify what they believe, because science is “not biased”
- people ignore data that runs counter to what they believe, without realizing that valuable information can be gained from opposing data

We want you to use data in your work, but encourage you to think about your own code of ethics.

We also want you to realize the many people use data in a misleading way. You should be prepared to tell people when you believe their data is not based on sound decision making and bad research. It is important for all of us to set a high standard for others and ourselves when data and research can play such a deciding role in policy formulation and, in turn, people’s lives.

Discussion points:

- Is there a “hidden agenda” inherent in the research? What is the agenda of those conducting the research? What about those who use the research?
- In policy, presenting data with too restrictions or caveats can diminish the effect of data on policy implementation.
- In policy process, you are trying to be the best source of information possible – to build trust among policy makers so that you will have influence. A lot of that trust depends upon your credibility as a source of information. High credibility may include that you be accurate, accessible, consistent, and reliable in your data.
- Do not ignore data that contradicts what you are trying to accomplish. Pay close attention to your opponents data too. Often times, if you understand what they have “up their sleeve” it will help you create counter arguments and prepare for their attacks on your own data. Use the criteria we developed to determine how “good” the opposing data is.
- As a research center, we strive in every way possible to come as close to the truth as our methods allow. We also try to make our methods “transparent” so that you can clearly see the sets of assumptions that we are using. This is how we built our own credibility. Be suspicious of anyone who is not willing to reveal his or her methods to you.

Conclusion

Goals:

The post-test will provide you with feedback regarding the participants increased knowledge about the use of health research data in their work. This will provide students with some time to reflect and think about ways they can utilize their understanding of data in their own work.

Timeframe:

Activity	Equipment	Time	Scheduled Time
Conclusion		20 min.	3:50 – 4:10
Post-Test & Evaluation	<ul style="list-style-type: none">○ Post-Test Evaluation Form○ Evaluation Form	20 min.	4:10 – 4:30
Total		40 min.	

Conclusion:

- Thank everyone for their time and effort
- Explain the Technical Assistance Program to Participants
 - Refer them to the informational sheet on TA
- Provide contact information to the Health DATA Program and the Center.

Often time, data can be used in a misleading way:

- sometimes people use good, quality data – but only part of it
- sometimes people use not very credible data because it supports the very good thing they are trying to do
- sometimes people use anecdotal information, a story they know about, and generalize it out as if that were the experience of everybody in that same situation
- sometimes people use “science” to justify what they believe, because science is “not biased”
- sometimes people ignore data that runs counter to what they believe, without looking carefully for valuable lessons that might be in it

We want you to use data in your work, but encourage you to think critically when determining the data you need, when looking for data and when presenting it. We also want you to let others know when you believe they are using data in a misleading way. It is important for all of us to set a high standard for ourselves and others when this kind of information can play such a deciding role in policy and people’s lives.

Lastly, we want you to hold researchers accountable. The politics and ethics of data are complicated and you will have to decide for yourself the effect these aspects of data will have on

your own work. We only hope that we made you feel more at ease about understanding, using and applying research data.

Post-Test & Evaluation:

Take five minutes to have participants complete the post-test.

Have participants complete the evaluation form before they leave or have them send it back to you if they need to leave.