Region IV Public Health Training Center
A MEMBER OF THE PUBLIC HEALTH TRAINING CENTER NETWORK

Leading Public Health: Data-Driven Leadership

Episode 6: Quantitative Methods & Analysis

Liz Kidwell (LK): Welcome to Leading Public Health, a podcast from the Region IV Public Health Training Center at Emory University. Through this podcast, we explore current leadership challenges, strategies, and ideas to help build the capacity of the current and future public health workforce. In this series, Data-Driven Leadership, we explore the essential role leaders play in incorporating fairness into program evaluations. You will gain valuable insights from experts who share practical tools and strategies for measuring and enhancing program impact. This series features pre-recorded sessions from a project ECHO Initiative offered by the Region IV Public Health Training Center, the Injury Prevention Research Center at Emory, and Safe States Alliance. In this episode, we hear from Dr. Jen Gathings, Senior Research Associate at ETR Services, as she shares her expertise in program evaluation, applied research, health impact assessments, and capacity building. With over 12 years of experience, she provides best practices for quantitative data collection, the advantages of mixed methods approach, and strategies for ensuring data collection and analysis reflect community needs. All resources shared in the episode will be linked in the show notes. So be sure to check that out after you listen. Let us get started.

Jen Gathings (JG): A little bit about me. I grew up in the Carolinas. I went to North Carolina State University. I am a sociologist by training. I studied race, class, and gender inequalities as my specialization. Since graduating, I have, well, even while I was working on my doctorate, I worked with a wonderful organization that I am still with. And we do a lot of evaluation, applied research, training, and technical assistance for various kinds of partners, non-profit partners, folks working in the education sector, and some private organizations as well. So, I have been doing this for more than 12 years at this point, and most importantly, Lover of Cats, Performing Arts, Science Museums, and Traveling.

The organization that I mentioned is called ETR Services. We are based in Durham, North Carolina, and we are a mission-driven organization, and our mission is to work with partners and programs that are supporting underserved communities. So the direction for today won't talk a bit about quantitative data, talk a bit about mixed methods, and try to infuse those discussions with some ideas about how to center equity in data collection and analysis. And I'll bring in some examples from various projects to just help make some of these cases. Then also, we want to end with some tips and strategies for successful and useful analyses, because we know that the most useful analyses are ones that actually they get used. So we want to talk a little bit about how to prepare for success.

So just to start from a common framework, when I say quantitative data, I'm talking about data that are information that are captured in numerical form. Typically, we can categorize quantitative data in two broad areas, categorical data, which will have a limited number of possible values. This could be something like a Likert scale, where somebody is sharing their agreement or disagreement with a statement using a scale from one to five or one to seven. Then we also have continuous data, and these kind of data have many possible values. So here, think about weight or height or some other kind of measurement, where there is a whole lot of variation or range.

There are multiple places where we typically pull quantitative data from. First and foremost, if you're doing evaluation, then you might be able to pull some of that information from your program documentation. And that can include looking at your administrative data and your fidelity trackers, participation trackers. It could also include your evaluation data. So if you've administered surveys to people in your priority population, if you have observation data from your intervention, if you've collected qualitative data as evaluation data through your project, sometimes you can quantify that data as well.

There are also official data that we commonly use, and I've included just a few examples here, but there are so many that are out there. WISCRS is one, that's the Web-Based Injury Statistics, Query and Reporting System, the National Maternal and Infant Health Survey, and M-I-H-S, or Census data also would be an example of official data. And then there is also a place that I learned about in grad school that honestly, I don't use very much, but it could be a wonderful source of data, and that is, or those are data repositories. So thinking about things like ICPSR or that Inter-University Consortium for Political and Social Science, the Harvard Dataverse or Statistica, something like that, Statistica being offered by Johns Hopkins. So lots of places out there where we can pull quantitative data.

So I want to pause and just ask a couple of questions for you. The first one, I want you to think about what are some of the ways that you previously are currently use quantitative data and health equity evaluations or even other kinds of evaluations. And as you're thinking about that, I want you to think about some of the challenges that you've encountered with using quantitative data.

So if you work with quantitative data, you have probably realized there are some advantages and there are disadvantages to using this kind of data. Oftentimes using quantitative data can be quick and a more cost efficient approach for evaluations or other kinds of research. So sometimes using quantitative data is great. It reduces some of that cost in comparison to collecting qualitative data, but it's not free, right? Sometimes, you know, there is a burden for collection with going out and doing surveys, especially if you want a sufficient sample size and you want to make sure you're hitting all of the demographics that you are interested in.

We also know quantitative data can be good for answering certain kinds of questions like what or how many, but they are not as great answering other kinds of questions like the ones related to why or how. Typically, we look to qualitative data to fill in those blanks and help us better understand how something is happening or why it might be happening. And then also, many people consider quantitative data to be more concrete with minimal possibility for review or bias. So they see it as being more legitimate. The findings you read from it are more legitimate because there is not subjective interpretation like what we would see with qualitative research or evaluation. But again, like quantitative data, the data collection can also be time consuming, but again, not as time consuming as qualitative approaches.

One of the things we talked about prior to this session was where do you go for quantitative data for health equity? And I think that is a question that a lot of people are asking right now and some people are out there trying to answer. And I will say when that question came up during our pre-session, the first thing that came to my mind was, I don't know that there is a single place to go for really great health equity indicators. But I do share some resources here and in the resource guide for some places where you might start to look and see folks who are doing some of this work to compile these data sources for health equity have uncovered anything that aligns with the area in which you work.

Here is a little snippet from some work from Baylett Health who is putting together a state health equity measure set with funding from the Robert Wood Johnson Foundation. What you see over on the right-hand side includes health status measures so related to behavioral health conditions, birthing people, infant, perinatal, and family care. And then the last one is chronic pain conditions. They also include a set of healthcare measures that measure healthcare access and affordability, behavioral health conditions, birthing people, infant, neonatal, and family care, chronic physical conditions, oral health, and prevention. So if you're working in any of those areas, check the resource sheet. There will be a link for you to check out the measures that they have identified as potentially helpful.

A lot of the quantitative data that I've ended up working with has come from, you know, a secondary data source. So it's secondary quantitative data. And I think it's important to recognize that there are many challenges with using secondary quantitative data. Many sources of official data in particular, so data that are collected by government or other offices, a lot of times these data are collected without explicit concern with or for equity. So there may be data collection bias in that they did not oversample specific populations that really should have been oversampled to make the data useful for looking at health equities or inequities.

We oftentimes are working with limited information about the variable details, the context. There might be variables of interest that are missing from the dataset that we would want to have access to or information on categorization issues also come up pretty frequently, particularly with response categories and how something like race or ethnicity might be captured in their dataset. It might not align with how we're thinking about race or ethnicity or gender or sex in our own work. And then oftentimes these datasets are outdated as well because we know collecting large swaths of data, it takes time, it takes effort. And oftentimes the data that are available are at least several years old or more.

There are ways we can think about mitigating some of those limitations of secondary quantitative data. We can think critically about the data collection methods, the variable definitions and any potential sources of bias. Can also consult with community members and experts and other colleagues in the interpretation of the data. We can name data limitations and we can also think about triangulating and our research or our evaluations. And there's multiple ways we can think about triangulating and research or evaluation. We can think about data triangulation where we're using multiple sources. We can think about methodological triangulation when we're using multiple methods. We can think about investigator triangulation when we have multiple researchers or evaluators that are working on analysis. And we can also think about theory triangulation where we're working for multiple theories.

So thinking about some of the benefits and some of the limitations of quantitative data. You know, mixed methods really fills in some of those gaps for those limitations. And mixed methods are typically strategically integrated or combining rigorous quantitative and qualitative research or evaluation methods to draw on the strengths of each. And that's coming from NIH. So utilizing mixed methods can yield multiple or different perspectives for us. It can also help us validate what we're seeing in our data. It builds comprehensive understanding. So not only can we speak to how many or to what magnitude something is occurring, but we can also begin to help provide some explanations for why or how something is occurring.

We can begin to speak to process. We similarly, you know, it helps us explain those statistical results in greater depth and it can provide more contextualized measures for us as well. So getting started with mixed methods. There are several questions you might want to ask yourself. What is it that you want to know? What will be the detailed quantitative, qualitative and mixed methods research questions or evaluation questions you hope to address? What kinds of data will you collect? What will be quantitative? What will be qualitative? Are there opportunities to collect quantitative and qualitative data that speaks to the same indicators that you can use for triangulation? Which rigorous methods you're going to use to collect those data and engage with your stakeholders or partners? And how you engage or integrate the data in a way that allows you to address your first question.

So once you've done some of that initial thinking, you can begin to think about the kind of method that you would want to use for your mixed method design. And there are several that are pretty common. The convergent parallel design is a design where you're collecting your quantitative and your qualitative data at the same time. And you're using those data for comparison purposes or relating them to each other to arrive at an interpretation. With an explanatory sequential design, you start with the quantitative data collection first and then you follow up with the qualitative data collection. And again, that's for the purposes of interpretation. And we'll flip it for the exploratory sequential design where you might start with your qualitative data first and you're gonna build up to collecting quantitative data through a survey or some other approach and you'll analyze those data together, again, thinking about interpretation. There's also something called an embedded design where you might be conducting a large study, perhaps it's a large quantitative study over several years, but at some point during your larger multi-year study, you decide you're going to collect some qualitative data to help you with part of that study. Maybe it's interpretation or refining measures or something like that. We would call that an embedded design.

So one example that I can share from my own work comes from a health impact assessment that we did of paid leave, where we were looking at paid leave policies in the state, particularly paid leave provided by employers. And we really wanted to better understand what was the impact of having access or not having access to paid leave provided by an employer. And we also wanted to begin to have some discussions about safety leave for victims and survivors of intimate partner violence as well. And initially we had planned for this to be a qualitative HIA. We wanted to really focus on collecting stories. We wanted to hear from working parents with small children. And we really wanted to sort of humanize paid leave.

Unfortunately, this happened during the pandemic where, you know, things shut down. And for the university where IRB was approved, they suspended qualitative data collections, no in-person data collections. And just a lot of what was going on in the world made it incredibly hard for us to carry out the project in the way that we had initially intended. So we stepped back and said, okay, well, let's do something that's more mixed methods. And we reduced the number of interviews that we conducted with working parents to a more manageable number. And we did a parallel study with many of the official sources of data and secondary quantitative data sources that were available.

So that meant we looked at the general social survey to get ideas about general attitudes toward paid leave. We looked at the intimate partner violence survey to get a better understanding of IPV as it affects communities and be able to lift some of those statistics up as well. We looked at pregnancy risk assessment monitoring system data or a PRAMS data, survey of income and program participation data, which is SIP data, the US Bureau of Labor Statistics, we included data from their data sources as well. And of course the US Census data too, so that we could begin to look at national trends, we could look at state trends, and we could begin to do some estimation for the number of parents in our state that would benefit from having access to paid leave through their employer.

So I wanna talk just a little bit about analyzing quantitative and mixed methods data for health equity. One of the, I think, simplest strategies to start with is disaggregating data based on key variables. And how you think about doing that disaggregation is really important. So sometimes we might just think about disaggregating by race. For example, the figure over on the left comes from some work that we did with an organization that was situated in the black belt. So they're in the southern part of the US and they had done this incredible work around health equity and COVID, where part of what they were doing was engaging deeply with the community.

During that deep engagement, they were collecting surveys from people about their vaccination status. They were providing information to people about COVID vaccinations and asking about the impact that information had on their thoughts about

COVID. And they also asked the series of questions about local health hospitals and health clinics, the quality of those services that were available to them, how far they had to drive, things like that. There are over 34,000 cases in this data set that spanned, I think it was 12 states. Measured on a scale from one to five, where one means very poor, five means very good. We wanted to begin to disaggregate this data in ways to help them tell their story of what they were seeing, what was going on in these very different communities. And they're also interested in race. So we wanted to make sure we were looking at racial disparities and health equity.

We did a series of analyses and I thought it would be interesting to just show some of the strategies we use because you begin to see, as you're disaggregating, you begin to see different stories emerging. One thing to keep in mind, this does typically require a large sample size. And to do this well, you typically need a sufficient subsample sizes as well. For example, we were looking at that health clinic and hospital quality rating by race. We disaggregated for the whole region. And because we wanted to center black communities, we looked at black versus non-black. And we had just stopped here. We would walk away from our findings saying, "Oh, well, black respondents typically provided higher ratings for hospitals and health clinics in the area."

But as we began digging, we found that that wasn't always the case. And we know context matters. Individuals are in relationship with others, within a community, and that exists within a society. So we thought first we would begin to disaggregate by counties in some of these states. And you'll see here that we still disaggregated by race, but we also disaggregated by county and began to do some comparisons. And we found that in, I think it was seven out of the 12 counties that actually non-black residents were providing higher ratings than black residents for the local hospitals and health clinics. And in six of those instances, the findings were statistically significant. So that's one example where you wanna continue to ask those questions and not just stop when you are disaggregating by one variable.

We also know intersections matter. We are individuals that are embedded within numerous systems, context, and historical moments. We all carry a multitude of intersectional identities. So as we're thinking about intersectionality and these intersecting identities in our work, we should also begin to think about ways we can begin to analyze those intersections. And there are a couple of strategies. The first one that came to mind for me was thinking about dummy variables and interaction terms and survey data. And again, here you do need a pretty large data set. And this is a more sophisticated approach, but it generally involves modeling interaction effects to explore the differential impact of intersecting identities.

And this is also, I'll say, an improbable example. We'll look at depression scores as our dependent variable, medication use being Prozac as an independent variable and also receiving therapy as an independent variable. And what we see is that for individuals that took Prozac alone, that their depression score went down. We also see that for individuals that took Prozac and received therapy, their score went up. So what this is communicating, and again, this is not very likely to see in the real world, but what it's communicating is that there's something uniquely different or important that's happening when we combine medication use with therapy. So it's a cumulative effect that's very different than what we would expect from seeing from one of these variables alone, or their impact on the dependent variable alone.

So this is capturing that idea of an interaction effect. So when you have these two variables that are appearing together, there's something special that happens. And there is a special effect that we can measure if we wanted to do some modeling. Typically you do this by adding dummy variables to a statistical model like OLS. And then you create a product term that captures that unique synergy for your interaction variables. An easier approach is disaggregating at intersections, and that is perfectly valid as well. So we're still looking again at the hospital and health clinic quality data, this time by gender. And you'll see, which is pretty typical, that females provide slightly higher ratings than males. And in this particular data set, the partner also captured non-binary as measure of gender. So we see non-binary individuals had a slightly lower rating than both males and females.

So if we begin to disaggregate by race, again, we're looking at some of those intersections. We see that a different story emerges again. We see that non-black, non-binary respondents report lower quality ratings than black non-binary respondents. And it's a pretty significant difference there as well. In fact, it's statistically significant as well. So just wanted to share this and highlight how you're thinking about disaggregating your data is really important. And as you're thinking about these different kinds of intersections, there's opportunities to dig deeper and go further to really think about how individuals with intersecting identities are experiencing an intervention or healthcare or local hospitals. It's really important to bring those stories out too.

So moving into our final section, just thinking about how to make our data analyses useful. We know that evaluations get used when they are relevant, when they're accessible, and when they support data-driven decision-making. So just a few tips. Don't think about analyzing your data. Don't wait to think about analyzing your data until you're analyzing your data. How you want to analyze your data can have implications for how and from whom those data are collected. So I'm sure this is something that we all realize on some level. Equity is not only an outcome, but it's a process. And doing health equity work not only includes the kind of work, the sub-stume kind of work we're doing and the kinds of analyses we're doing, or the kinds of changes we want to see in our communities, but it's also about how we do that work too.

And it's really important to work with partners to identify analyses that need to be conducted that will, and I'm gonna use some jargon, but the analyses that need to be conducted that will help move the needle in your community. There are a couple of ways that I like to do this and they start very early in the process. One is using an evaluation matrix, which I know many people do. But the purpose of this not only is to help us with planning and clarifying our thoughts and thinking through what it is we wanna ask and those indicators and where we're gonna find our data and how we're gonna analyze those data. But we also share this back with our partners early on in the process and use it for facilitating conversation about the project and their data needs.

So we bring in that expertise and try to provide a framework initially, but have found it to be so valuable to walk through this and facilitate discussion around the evaluation matrix when our partners, because our partners are going to be the experts in their program. And oftentimes the experts with serving their particular priority populations as well. So if we're not doing some of that initial planning with them, we're really missing out on some of their expertise too. So this is one great way to make sure you're planning for useful analyses.

Another way is stakeholder mapping. Oftentimes we do this early on in a project as well. I am using the term stakeholder mapping because that's how CDC has created this particular resource that I reference. We've used this in the past with partners when we were working with several different counties who are also working with a network of partners within their counties. And we would walk them through an activity asking them to think about who their primary stakeholders or partners are, as well as their secondary partners and tertiary partners. After they think about who they are, we ask them to think about how their partners might be affected by the evaluation results and how they might contribute to the evaluation. And we use those questions as ways to begin to really dig deeper to get a sense for what kind of data would be useful, not just for the partners we're working with directly, but for their community to help push things forward. And then we use that to help co-create key questions that we should address in the evaluation.

So again, this is something that's done pretty early on, probably like far sooner than you're thinking about doing your data analysis, but can be a really great source of information for thinking about what data collect and also thinking about how you're gonna analyze those data once they come in. And if you can't tell another strategy I like is just talking. So could all fashion conversation, like when we are working with a partner, we are constantly asking, what do you wanna learn from this evaluation? How will the data and the findings generated through this evaluation be used? Who do you think your partners are that are going to read this? Who do we want to read this? And we also ask what information is gonna be necessary for continuous program improvement as well.

So just a few key takeaways. Quantitative data can help us answer great questions about magnitude of health inequities, but it may not help us understand how or why those inequities emerge or persist. So this is a great opportunity for us to think about using mixed methods where we're combining quantitative and qualitative data. And doing so can help us overcome the limitations of each of those approaches singularly. And through thorough planning, open communication with partners and key stakeholders or partners, and some intentionality, we can really begin to do some planning for successful, help equity evaluations that not only get done, but they get used.

EK: We hope you enjoyed this episode of "Leading Public Health." A podcast from the Region IV Public Health Training Center at Emory University. We value your feedback, so please take a minute to complete the evaluation located in the show notes. Thank you for joining us.

This project is supported by the Health Resources and Services Administration (HRSA) of the U.S. Department of Health and Human Services (HHS) under grant number UB6HP31680 "Regional Public Health Training Center Program." The contents are those of the authors and do not necessarily represent the official views of, nor an endorsement, by HRSA, HHS or the U.S. Government.