Technical Assistance Webinar
Identifying Appropriate Data Sources for Community-Level Evaluations
Tier 1B Grantees and Evaluators

Moderator: Amy Farb
April 5, 2016
12:00 pm CT

Coordinator: Welcome and thank you for standing by. At this time all participants are in a listen-only mode. During the question-and-answer session, please press star and 1 on your phone. Today's conference is being recorded. If you have any objections, you may disconnect at this time. I will turn the meeting over to Ms. Amy Farb. You may go ahead.

Amy Farb: Hi, everybody, it's Amy from Office of Adolescent Health. Welcome to today's Tier 1B Webinar. This Webinar builds directly off of our last Webinar. This time our help from Abt is going to offer us some considerations for selecting data sources and helping you understand some key features of relevant data sources.

So today we are lucky enough to have Kim Francis, the Project Director, we have Randall Juras and today we're joined by somebody new to you all who is Sara Donahue and she's a data expert so I will defer to her all of your questions and anything we can't answer or that, you know, is so specific really

to just your project, we can try to handle later offline so let me go ahead and turn it over to (Kim) and we'll go ahead and get started.

Kim Francis: Great, so I'm just going to go over what we're planning to cover today so what we're going to start with Dr. Randall Juras. He's going to provide a short overview kind of slight review from the last review of how to select outcomes and co-variates in the context of community-level evaluation designs.

And then we'll turn it over to Dr. Sara Donahue who's going to guide us through some of the key considerations and issues for identifying administrative data sources and we'll finish by discussing some specific datasets that could be used for different types of outcomes and co-variates in a little bit more detail.

And then there will be time for questions at the end and we'll also just like last time there's a question, a Q&A box on your screen where you can type-in questions to us during the Webinar and we're going to plan to stop once about roughly halfway through to see if we can take a couple of questions that have been submitted through that box.

And then depending on how much time we have left at the end we'll take a few more from there as well as open-up the phone to see if anyone wants to ask a question directly.

So we hope that by the end of today's session you'll understand the key factors in selecting data sources to measure community-level outcomes and co-variates and that you'll know more of the content, the timeliness and the accessibility of some particular datasets that could be used and be better able to assess the appropriateness of different potential data sources for your own evaluations.

So each of you is likely to have a unique situation when it comes to identifying the right data source and sources so today we're aiming to give you a high-level introduction to some of the key issues but we might not be addressing your specific situation but hopefully we'll provide some information you can take with you and use to make some decisions about your own data and situations and with that I will turn it over to (Randall).

Randall Juras: Thank you, (Kim), hi, everyone. I'm going to take a few minutes here to walk you through the kinds of data that we expect most of you will use in your evaluations. Most of what I have to say is a review of material from either the grantee orientation or the first Webinar but I want to go through just sort of to ground everyone in where we're at.

Then my colleague Sara Donahue will take a deeper dive into potential specific data sources so as you know, the purpose of an impact evaluation is more or less to assess whether an intervention affected some outcome compared to what the outcome would have been without the intervention.

Now you can approximate that counter-factual what the outcome would have been without the intervention by looking at what happens in similar communities that did not implement the intervention.

For that reason the typical QED requires two kinds of data: first, you need data on the outcome of interest to determine whether or not it changes; second, you need data that will allow you to determine whether your treatment and comparison communities are similar.

We're going to call this latter kind of data co-variates in this Webinar mostly because they usually appear on the right-hand side of a regression model.

Most data sources do not have information on all of the outcomes and/or all of the co-variates that you would ideally want to use in your study.

It is possible and in fact it's quite common to get data on outcomes and co-variates from different data sources. You can combine data from more than one source as long as you can identify the same kinds of communities in each data source so that you can link them together. (Sara)'s going to talk a bit more about that later in the Webinar.

So there are lots of considerations that go into choosing the most appropriate outcomes to study in your evaluation. For example, you can only assess the program's impact on outcomes that you can get from some data source.

But before getting too far into the data, I do want you to take a step back and remember that the outcomes you assess should be things that your program is intended to affect, in other words, things that appear as short or long-term outcomes at the right-hand side of your program's logic model.

It's worth putting some effort into finding data on outcomes that are actually meaningful for your program rather than just the outcomes that you happen to be able to find in datasets so here are some examples of the kinds of outcomes that we've seen in your logic models.

These include obviously teen pregnancies and births. Also they include health outcomes such as the prevalence of STIs, sexual risk behaviors such as contraceptive use as even academic outcomes which might be particularly relevant for something like youth development program.

For reasons that (Sara) will describe in detail, it will be very difficult for most of you to use pregnancy as an outcome in these evaluations, having to do mostly with the timing of data availability.

So we encourage you to look at shorter-term outcomes and in particular shorter-term outcomes that might be antecedents so if pregnancy so that you can assert that if you affected those outcomes through your intervention that possibly you'll have a longer-term effect on pregnancy and these are things like academic or sexual risk behavior.

Before moving on, I do want to emphasize that data on most of these outcomes will be measured at the individual level. For example you mentioned whether or not an individual woman is pregnant, not whether a community is pregnant.

But for evaluation purposes your data will need to be aggregated to the community level because your programs are intended to for example chance community norms and therefore affect community-wide outcomes such as the teen pregnancy rate, the STI rates or the high school dropout rate measured community-wide.

If you use different data sources for different outcomes and/or co-variates you need to make sure that the data can be aggregated to the same community level. Now having said a couple of words about outcomes, I want to talk a little bit about co-variates. Returning to the research design, QED compares the outcomes in the treatment group with outcomes in a similar comparison group with emphasis on the word similar.

It's basically by observing by comparing observable community-level attributes, i.e., co-variates that you would determine whether your treatment and comparison groups are indeed similar.

So in addition to assessing whether the groups are similar, you can use co-variates also if they're not exactly similar to statistically adjust for any differences across the two groups, for example by using propensity score matching or Mahalanobis Matching which we talked about in the last Webinar.

So how do you know which co-variates to use? Well, a big goal is looking at these observable characteristics is to convince skeptical readers that the two groups in other words the treatment group and the comparison group really were similar before the program implementation started.

Therefore, any differences between the two groups at follow-up must be due to the intervention and not to other factors so the best co-variates to use are factors that you would expect to be correlated with your outcome of interest.

Doug Kirby at the National Campaign to Prevent Teen Pregnancy I believe published a list of study findings on the antecedents of adolescent sexual behavior in a booklet called Emerging Answers and these include things like education, this community-wide education rate, the unemployment rate and the crime rate along with a bunch of individual-level characteristics that could potentially be aggregated into the community level.

So if you take a look at that, you'll find very good things to use as co-variates in order to convince your readers that you're controlling for the right things. Now finally many of the administrative data sources that (Sara)'s going to talk

about in just a second here are only available for certain time periods or with some delay.

Quite often administrative data sources on vital statistics are only available on a calendar year basis and so for that reason I do want you to have the timeline in your head as we move forward. As you know, these are five-year grants that you're working on. They were awarded in July of 2015. They'll end in roughly September 2020.

You're currently in the planning year, Year 1 working on your evaluation plans which will have to be submitted and approved this summer. You'll be implementing starting this summer 2016 and ending in the summer of 2020.

You will not be able to assess the impact of your program through the summer of 2020 most obviously because you'll need time to write the final report but also because data have some lag and so we expect that the most likely period during which you'll be able to actually assess the impact of the program is for three years between about January 2017 and December 2019.

Any data that you're collecting on a calendar-year basis would only go through December of 2019. Some of it may be available on March 2020 which is probably the latest point at which you can successfully incorporate data into your evaluations.

So with that in mind, now I'm going to turn the presentation over to (Sara) who will talk you through potential data sources and other considerations.

Sara Donahue: Hi, thanks (Randall) so my name is Sara Donahue and I'll be going through some more details on different types of administrative data sources that you

might be aware of in thinking about some of the considerations for their use in your evaluations.

So to reiterate, because these interventions are at the community level, the outcomes and the co-variates are also at the community level and so we're going to talk now about some of the initially about the considerations for using administrative data sources to measure these community-level outcomes and co-variates.

Just as a reminder for many of you work with administrative data on a daily or regular basis and these data are very useful sources for community-level data. Rather than collecting information from each individual in the community through a survey or other type of data collection, you can use existing information that's been gathered by community programs and other initiatives to assess the outcomes and co-variates of interest for your evaluation.

Administrative data are defined as data that are collected as part of routine operations for our programs and here you'll see some of the data collection mechanisms that are used for administrative data. They include ongoing surveillance so for example to count live births in an area over time, to study and to measure and monitor program or service utilization.

So for example educational records of students in a school or school district and then another category of administrative data are data that are captured as part of needs assessment or other types of research activities and that that include surveys about knowledge or behaviors.

So some key features of administrative data for this evaluation are that the data already exist and as you know the FOI states that we cannot be surveying youth so using existing data is important and the data are able to be

aggregated at the community level that you will need for the program evaluations.

As (Randall) mentioned and you all know the ideal outcomes for these evaluations are looking at pregnancy as well as births but data on some of the outcomes might not be available to you in the timeframe of interest for conducting your evaluations.

Some other data on relevant outcomes such as sexual risk behaviors or academic outcomes might be difficult to access in your community so a key challenge for the work that you'll be doing in the evaluation is to determine how you can balance the relevance of the outcome whether it's meaningful outcome that the program expects to impact versus the timeliness and availability of data to measure that outcome or those outcomes.

And unfortunately there aren't any simple solutions but through the rest of this Webinar we will be reviewing some specific administrative data sources and talking about the key considerations so that you can make some informed decisions about which outcomes you will measure in your evaluations and obtaining the data that you'll need to do that work.

Three of the key considerations to be aware of when selecting and utilizing administrative or secondary data source to measure the outcomes and co-variates for this evaluation are those level in units of the data, the timeliness and the quality of the data and together these will help you determine the feasibility of using various data sources for evaluations.

So to start with the concept of level and units, the evaluations need to measure outcomes and co-variates at the community level by which we mean the level to which the program is scaled for your work so when selecting administrative

data to use in the evaluations, the data needs to be available at that geographic level that's used to define the community.

And across the different grantees you have used several different geographic units to define the program communities so programs are using one or more counties, one or more ZIP Codes, that's the most often but then some projects are defining the communities as a city or more than one city, school districts or one or more neighborhoods or boroughs and there are even some other definitions that are used by some projects.

A key point here is that the data sources for the treatment and the comparison communities must have the same geographic units so if it's a set of counties and it needs to be a set of counties, if it's a city, it needs to be a city.

So what you'll need to do is to clarify and understand the populations that are represented by the data sources that you identified to make sure that that data source will be able to measure the outcomes and co-variates of interest in the population of interest so that's both geographic and around the characteristics of the population.

So for example the age range of the teens' birth data needs to match the age range of the teen birth outcome that you've identified in your logic model. Another factor to consider is timing. Timing is always everything and (Randall) just went through the timeline a bit with you.

When and noted that when we're working with administrative data sources for the evaluation, you want to be sure that you use data sources that cover the period of project implementation but administrative data are typically not produced for use in real time so there is a lag between when the data are generated and when they'd be available for the evaluation.

You do need to be able to access the data in time to use it in the evaluation analytic and reporting activities which will likely be in that first half of 2020 with the final evaluation report due in September of 2020 if not before.

So one consideration might be that because vital statistics data are a likely source of information for you for the teen birth outcomes, that can be a starting point for thinking about when data might be collected for the measures of interest.

And so you can see that in the center here the evaluation data collection of January, the calendar year of 2017, 2018 and 2019 which does not overlap identically with the implementation years but does cover the time period of project implementation.

And so knowing the timing of the data for key outcome measures can serve as a reference for considering the timing and availability of other data sources on other outcomes and co-variates and you need to clearly define and explain the time period for evaluation and how that relates to the time period of the intervention in the evaluation approach.

One key point on this timing concept is that the age group for the outcome measure will likely differ from the age group in the intervention because of cohort characteristics and the timing of data availability so to explain that a little bit, the effective programs implemented in 2017 with 14 to 16-year-olds can be described with outcomes data in 2018 on age 15 to 17-year-olds.

Those 15 to 17-year-olds in 2018 are the ones who are 14 to 16 in 2017 so you just need to consider that cohort aging as you plan your evaluations. Another factor to consider is evaluating the quality of data sources. This is important

because the data are going to be informing decisions that are made and they need to be credible and useful to your stakeholders.

Most of the administrative data that will be used in this evaluation were not gathered for the purpose of the evaluation so and the reason that they were gathered may or may not be in line with the goals and the needs of this evaluation.

So understanding data quality and how the data are produced and we'll go into that in a little bit more will help you determine whether the data source is a reasonable one for using in the evaluations or any considerations that you might need to make during the evaluation's design and implementation.

So thinking about the fitness of youth of the administrative data, you can assess data looking at dimensions such as relevance which we talked about a little bit, accessibility and interpretability or whether the data are clear.

Data coherence and accuracy refers to seeing whether the data are consistent over time and comparable across geographic areas which is a key consideration for you in looking at the treatment and comparison communities with the same measures and then assessing credibility will involve looking at whether the data source produces high-quality and reliable data.

There are some existing quality assessment tools that can be used to review administrative data and we have an example for you in our resources that will be at the end of this presentation and available on the Website.

And then you can also discuss the data characteristics including limitations directly with the agency or organization that created or owns the data and

work with them especially if they've had experience of having somebody use the data for other evaluation activities.

We talked about the aspect of linking datasets because of the data for the evaluation will come from different data sources so for example health outcome data might come from the health department or clinics.

Service utilization data might come from a clinic, some education data might be coming from schools or schools districts and then co-variate data would come from an existing source like the American Community Survey.

So in this example you see illustrated for our visual friends that the different data sources need to be the same in measuring the same factors in the treatment and the comparison community. Here is the example of showing a school district in two different treatments and the comparison.

And that means that the rates are reported for the same age groups and at the same geographic level and then the data sources all get linked together using that geographic identifier so here are the school districts and then the data can be put into a regression model for evaluation purposes.

Understanding how administrative data are collected and reported and when they are available can help you determine which data might be appropriate or accessible for your evaluations so now we'll turn to discussing some specific datasets for some of the key outcomes and co-variates in a bit more detail. (Randall) mentioned a few of the outcome measures.

This table is illustrating some of the examples that were noted in the grant applications. These outcomes would of course be calculated at your defined community level so for example a county or set of counties and on the right-

hand side there's an administrative data source which lists some potential data sources for measures.

So vital statistics data from health departments are a source for birth data. The youth risk behavior surveillance system is the source of information on youth behaviors and knowledge and then healthcare utilization data can be obtained from a range of different sources and we'll talk a lot more about some of these data sources in a few moments.

There are other outcomes that were either less frequently noted in the applications or are more difficult to measure and they can be studied as part of your evaluation activities if the data are available in both the treatment and the comparison communities.

As we've talked about teen pregnancy rates are clearly an outcome of interest for these interventions but the data might not be easily accessible in the timeframe that's needed for the evaluations.

STIs can be studied. It's important to note that if more teens are accessing healthcare as a result of the program, these rates might be expected to go up and then the latter category here includes skills, self-efficacy, academic outcomes, other behaviors and community involvement or community service.

And those are all potentially relevant outcome measures for the work that you're doing and as (Randall) noted looking at the logic model to see which outcomes your program is expected to impact.

In addition to the baseline value of the selected outcome, a number of co-variates can be used for the first activity of matching a comparison community

to the intervention community and then these co-variates would also be used in analysis.

So here's some examples of community and family factors that were described in the grant applications either you described them in the community description or in the section on evaluation design or both and as with the outcome measures, these co-variates are to be calculated at the defined community level.

They include characteristics such as income characteristics, unemployment or poverty, demographics, for example racial and ethnic composition of the community. There are other administrative data available on factors including crime and education and family characteristics can also be measured at the community level using administrative data.

And the two data sources that we have noted here are the American Community Survey which we'll go into in a bit more detail and then the behavioral risk factor surveillance system is another source of information about for adult risk behaviors.

So to go into some more detail on some of these data sources, the first will be on the vital statistics data which is a source of information on teen birth and pregnancies.

So each state and territory has a registrar of vital statistics and they're responsible for collecting the vital records data about natality and mortality and so physicians and other hospital personnel have to report live births, fetal deaths and induced termination of pregnancy to the vital statistic registrars at the Department of Health.

And this reporting process follows procedures and definitions that have been established by the National Center for Health Statistics and then there's some state-specific or location-specific modifications for certain definitions and reporting timeframes and then the state registrars compile and publish the data.

The birth data are typically available between nine months to a year after the close of the calendar year but in some states there's even a longer delay. We do know that the birth data that are reported tend to be very complete and they're submitted relatively timely in a relatively timely way.

Most states have electronic or Web-based birth registrations and this leads to very accurate provisional statistics so while the provisional statistics are not as complete as the final year-end statistics, they do offer the ability for us to monitor health and assess the impact of interventions on a much timelier basis.

And it's becoming more and more common that using this provisional data is an option for researches who need more access to more timely data. The quality of the data is very good so each state knows its own data batch so it's very important to connect with your state and health department to understand the quality of provisional data and whether provisional data are available.

So we recommend connecting with them about this but on this slide we're showing that the lag could be about three months so the for example provisional data on the births that occur in the second half of 2016 might be available as soon as March of 2017 and then so forth looking at the next calendar years.

Some states do provide provisional aggregate data to the public and if they're not routinely request released by the state, then researchers can request the provisional data.

(Unintelligible) are defined by the National Center on Health Statistics as the sum of the number of live births, reported induced terminations of pregnancies and report of fetal deaths of all gestations and so calculating pregnancy rates requires compiling multiple datasets and for this reason the pregnancy data are traditionally not available from state vital statistic departments or state departments of health for several years.

You can explore options for pregnancy data with your local partners but we expect that this outcome is unlikely to be able to be measured for this evaluation. It's very important to remember that even if you're able to access pregnancy data for the treatment community, to do the evaluation you will also need to have data on the comparison community so connecting with data providers in both communities will be important.

Many of you access vital statistics data when preparing your applications so you're likely familiar with the resources that publish these data. There are some national Websites, the National Center for Health Statistics publishes county-level birth data.

The Annie E. Casey Kids Count is a phenomenal Website. It has data from a lot of different administrative data sources and including birth data and it's a good source of information for both teen birth data at the county level and for some states at the city level and they do have some teen pregnancy data but the timeliness really depends on the state so it's state by state.

The state vital statistics departments are really the best source for having the most timely data. The Website for state vital statistics usually has links to the reports that they have produced and some states have created some tools for creating custom tables or maps for data in counties of interest.

This here is an example of a report that was generated that I just generated to show some data for counties that turn teen birth rate data by county and again connecting with the state vital statistics department is the way to go for identifying access to data certainly at the lower level for example lower geographic level for example a ZIP Code.

Moving on to another data source, the youth risk behavior surveillance systems or YRBSS is a source of information on teen sexual risk behavior and other teen behaviors. This survey monitors six types of health-related behaviors and the surveys are conducted in selected public, middle and high schools every two years either by the CDC or by local and state government agencies.

There are some large urban school districts that participate in the survey as well. This survey is designed using a specific sampling methodology to allow CDC to generate national and state estimates and it's really important to know that the National YRBS data are not the aggregate of the state YRBS datasets.

And similarly the state's data and local data are not subsets of the national dataset so this is a survey designed with a sampling frame to create these population-based estimates. Some communities do use the YRBS methodology to conduct their own survey on these same topics.

So with the YRBS survey that's conducted by the CDC and states are available the summer, the data are available the summer after the survey is

fielded so for example the spring 2015 survey data will be available this June and the data are checked as they're compiled to assess data quality and the response rate does vary by location.

When the response rate is less than a standard of 60% the data are not weighted and that means that they're not included in the population-based estimates that are calculated for the state or the nation.

The unweighted data represent only students who completed the survey but with weighted data it's possible to say for example X percent of students in this state received HIV education.

There are also some reliability checks that have been implemented in the survey to assess data quality so overall it's a rather high-quality survey and data source and it does address relevant outcomes.

It is conducted in some of the locations where the projects are occurring and the data can be requested from individual schools and you might be able to access some information on data and measures before the intervention starts.

However, the state data that are routinely collected and generated are not the appropriate geographic level for this evaluation and because of the sampling design of this survey, individual schools might not be selected to participate in the YRBS survey and the upcoming years will be 2017 and 2019 and so individual schools may or may not be selected for one or both of those years.

And similarly the data might not be available even if it's available in the intervention school, the treatment school might not be available in the comparison community.

So it's just important to know that is the data were used from the YRBS for an individual school or multiple schools that because the survey is designed to be a representative sample of the United States and individual states, only the unweighted individual school-level data would be appropriate for use in this evaluation.

And those data are only going to represent the students who participated in the survey so just a lot of considerations about this data source but definitely worth exploring because it does cover relevant topics.

There's some examples of the output that can come on for the survey. The CDC Website currently has data from the 2013 survey and the 2015 survey should be available later this spring and so they publish the data at the state level and for certain cities and counties.

The state and lower geographic-level data are owned and controlled by the Jurasdictions that conducted those surveys and so sometimes giving CDC permission to distribute the data if requested but others just manage the distribution of the data themselves and so if a local so talking with your local provider would be the way to learn about the availability of the data.

If a local community has implemented its own survey using the YRBS methodology the data would be available from that organization. Health service utilization data is another category of outcomes that might be of interest for your program. One example here is TITLE X program.

It's the only federal program dedicated to family planning and preventive healthcare as many of you know and it's implemented through grants to public health departments as well as other agencies and community health organizations.

And all grantees are required to report data on their program users and service providers as well as utilization of health services and their report of using something called the FPAR the family planning annual report and apologies to FPAR and I don't know the acronym well enough.

So just looking at the timing of those data they are submitted by February for the previous year and the data are reviewed a number of times for quality and then provided in summary form and individual regional offices receive grantee specific summary reports so this graphic is showing some of the potential availability of data for the time period relevant to the project implementation.

So we're happy to talk with you all further about health services utilization data that might be available in your communities including this Title X data but updated from other service providers and discussing that availability is something worth exploring if that is an outcome of interest.

So just to summarize some of these considerations that we've discussed thus far about selecting administrative data on outcome measures, just a reminder that the outcome measures need to relevant, they need to be linked to that logic model that you have developed.

The outcomes information needs to be available for both the treatment and the comparison communities. It needs to be aggregated to that community level at which you are scaling. Data need to be timely enough that you can capture it in the amount of time needed to produce the evaluation report and the data need to be reliable and of high quality.

So after you've selected your outcome measures, reviewing the administrative data sources to determine feasibility is the next step. So we'll shift now to looking at some sources of information on administrative data for co-variates and as a reminder from the last Webinar when you talked about the comparison group, this is a local and focal comparison group.

And comparability of the communities is important for both the outcome measurement as well as the measures of baseline trends that are used to match the communities so when looking at baseline equivalence of the comparison group, you should be focusing on the pre-intervention data points that will help establish baseline trends used for the matching purposes.

So the match community or communities will be similar to the treatment community in the years prior to the start of program implementation this year, this summer so we've talked about some of the sources of information about outcome measures but looking at demographic data to look at community comparability one of the resources is the American Community Survey.

And this is an ongoing survey that samples about 3-1/2 million addresses each year or 1 in 38 U.S. households and the survey topics cover many measures that are of interest and relevant for this evaluation including demographic characteristics as well as education, employment and income characteristics.

The American Community Survey or ACS data are compiled in a number of ways including one-year estimates, three-year estimates and five-year estimates. These are all released annually. They're available about nine months after the close of the calendar year.

And as you might expect the five-year estimates combines 60 months, five years of data for the area and those are best used when you're analyzing very

small populations or looking at Census tracts or even ZIP Codes. The one-year estimates have data from the previous year for areas with populations over 65,000.

So they're less reliable given their timing but they are best used when you need data in a more timely way and when you have a large population, they're an appropriate source of data.

You can access the ACS data a number of different ways. The American FactFinder Website is the Census Bureau's Website and it provides access to the data at several geographic levels including the county and the city.

For those who are looking at communities to find school districts, the American Community Survey data at the school district level have been compiled by the U.S. Department of Education and they're available through something called the school district demographic systems data dissemination program and will have some information on how to access that at the end of this presentation.

And then some of these other Websites that have compiled administrative data from a number of different sources will also include ACS data at the county level and in some cases at lower geographic levels and some of the organizations that you work with already or some state partners might also compile this community-level data.

Here are just a few examples of visualizations of the ACS data. On the left this is by the school district showing family poverty level. This one was generated using that school district's demographic system data dissemination program and then on the right is an example of an output table obtained from the American FactFinder Website.

And here was an individual city looking at educational attainment characteristics in that location so here we're showing an example of how the demographic data can be reviewed to match the implementation community with the comparison community.

So you see that there are historic data compiled here on poverty in the community, the racial ethnic composition of the community and the percent proportion of single-parent households in the community.

So using this information you can identify whether the two different communities have similar trends and you can see here that these two match-up quite well. They're two different counties, the implementation county and a candidate for the comparison county so everything lines-up pretty well except for that percent Hispanic is a bit different.

So these data could be used with outcome data on outcomes of interest from the similar baseline and pre-intervention time period to confirm whether the baseline equivalence between the two communities is adequate is both local and focal.

There are some other administrative data sources that can be used for co-variates. These include community-level needs assessments and data from community partners and those were used quite frequently in your grant applications and if those data are collected on an ongoing basis, the relevant measures can certainly be used for evaluation.

But we really want you to be very cautious in considering this because you need to remember that the data must be available for both the implementation community which you may have studied in depth and accessed in depth but

then also for the comparison community so the same data need to be available in both sites.

The National Survey of Family Growth is a source that many of you are familiar with because it includes measures that are relevant to this topic and the evaluation including contraceptive use and risk behaviors.

This survey is very high-quality, low amounts of missing data and really done very well and has been going on for quite a long time. There are some problems with this data source in that it's not very timely and the geographic level may not be appropriate for your work.

Right now the most recent data that are available are from surveys conducted during 2011 through 2013 and the data are reported back to the public at the national level. It is possible to request county-level data but it does require a special application to the National Center on Health Statistics and the Census research data centers.

So as mentioned a couple of these sources that are general data resources that you can explore when looking at administrative data so the Annie E. Casey Kids Count Website, county health rankings is another Website with aggregated information from some administrative datasets as well as the health indicators Website warehouse Website.

And many individual locations that have at the state or the city level or county level have developed data bases that include information relevant to your evaluations and could be sources of information on the outcomes and/or the co-variates of interest so a few summary points before we look into some questions that you all have had.

To review that acceptable data sources are those where the outcomes and the co-variates are measured in a reasonable and credible way and where the data are available on your measures for both the treatment and the comparison community.

And the data need to be timely enough so that you can access data about the time period of the intervention within enough time to conduct your analyses to prepare your evaluation reports.

You'll be taking the datasets that are selected for use in this evaluation and linking them together using the community geographic identifier so those that identifier will mix together the data sources from different the data from different administrative data sources and then using that you can prepare regression analyses.

We've described some of the administrative data sources that are readily available and the key advantage to these administrative data sources is that they often provide historic information as well as the outcomes that you'll be studying over the next couple of years over the course of the intervention.

And just to come back to this concept of when you're selecting administrative data, you'll need to strike that balance between relevant data for relevant outcome measures and the timeliness and accessibility of data so you can review the options for accessing data while reviewing the potential outcome measures and the potential co-variates in the development of our evaluation plan.

So we have listed these a few resources for you all. We've covered so much information today and these resources will be available on the (unintelligible)

Website and they include resources on selecting outcomes and co-variates that (Randall) discussed.

And then some more direct links for identifying and accessing administrative data including these repositories as well as some source-specific links to vital statistics Websites, the YRBS, the American Community Survey and then here as I mentioned there are some tools available and some references to review when evaluating data quality so it's just some examples here.

I think what we'll do now is turn it back to (Kim) with (Amy) who has been monitoring the questions.

Kim Francis: Yes, and questions are coming in so I'll read a few that we've gotten through the Q&A box and then we'll see if there are any folks waiting to talk over the phone so the first question is a question for OAH. Would OAH allow no cost extensions so that final evaluation reports could be submitted after birth or pregnancy data are made available?

Amy Farb: Yes, so this is Amy Farb from OAH so at the end of the grants which, you know, end the last day I guess of June, you have 90 days to turn-in some final reporting but if you thought that that 90 days would not be enough for you to finalize your report, you could put in a request for a no-cost extension and we would certainly be very supportive of that but we don't just as a caution we don't have the approval of that.

That's strictly through the Office of Grants Management and it's up to them but we would absolutely support that if you didn't think you could finish it in the 90 days following the end of the grant.

Kim Francis: Okay, next question is about the YRBS. It's a little bit of a design question as well so for the YRBS can we utilize school-level data for the purposes of this grant and I think here (Randall) might want to weigh-in but it depends on how you're defining your community scope.

Randall Juras: Yes, I think that's right. If you're defining your community as a much broader area than a school so for example if you're defining your community as a city and you've decided that for the purposes of evaluation cities are the appropriate unit to be looking at, then using data on one individual school within that city is probably not appropriate.

But if you are really scaling-up your program or programs, just to the school level then I don't see any reason from a design standpoint that you couldn't use the YRBS to assess outcomes for that school.

On the other hand YRBS data aren't necessarily available for the same school at different points in time and may not necessarily be available for your school during the follow-up period as I understand the YRBS so that may or may not work.

Kim Francis: Uh huh, and this is a similar question that gets at this question of a level of saturation of a community and the data availability so the question is can you provide some strategies to deal with the situation where the community data available is for a larger scope than the treatment scope, that is, not all school districts received the intervention.

Randall Juras: Yes, so there are a number of well, there are a number of ways to think about this question. How you choose a community for the purposes of evaluation depends on a lot of things and often represents a trade-off.

You want to try to select a community that really represents the community conceptually that you are trying to scale-up to program to. In an ideal world you would scale-up the program to affect kids in the communities, you've defined it and you would then get data that was aggregated to the level only of that community.

And of course it would be fine if you're not actually affecting everyone in the community. You could only be directly affecting half of the kids in the community. You can still use data at that community level to assess the impact of the program on the community.

Of course it won't always be possible to find data which is aggregated to the same level that you conceptualized as the community you're scaling-up to. In that case you have to, well, you have to think about whether or not it would the impact estimates you would get from using that, you know, would be meaningful.

So for example if you were doing a program in three out of six school districts in a community that you had data for and you were affecting nearly all of the kids in those school districts, then one thing that you could do is to use data on your community realizing that not all of the kids will be affected, that you'll really only intends to affect about a third of the kids in the community as you're defining it for the evaluation because of the data requirements.

And interpret your impact findings appropriately so if you found that the program caused a decrease in teen pregnancy by X percent, you might then interpret your results to say well that was an X percent decrease on average across all of the kids in the community, 2/3 of whom we don't anticipate will be affected.

And so really, you know, we think that there is all this, you know, there is all the impact is probably three times as large as that. There are a number of ways that you can go about doing that from a technical standpoint but the basic idea so that's one option is to use the data that you're talking about for technology more broadly defined community and scale your impact estimates.

Another thing you know, probably preferable would be to try to find some data source which is measured or aggregated at the level of the community as he would really like to define it which may involve more searching or more legwork or maybe impossible but it may be worth trying to do that.

Kim Francis: Okay, great. This question has to do with baseline trends and comparison communities so I don't know if we'll be able to answer this here but we'll give it a shot so for baseline data where we're just looking at this year or do we need to look at the prior five-year trend.

For example is a community comparable if one community had a teen birth rate of 45 and another had one with 45 but one of those had a five-year trend of increase in birth rates and the other had a trend of declining birth rates. How would we define comparable in this situation.

Randall Juras: I think that the answer to that basically gets back to this idea of doing an evaluation and trying to convince a skeptical reader that your treatment group and comparison group looks similar.

And I think in the case that's described in the question, if you know that the treatment group is trending in a much different direction that the comparison group, even though they have similar levels of the outcomes one year prior to the intervention, I think as a skeptical reader in that case would say that well

those communities are not comparable and they would probably not believe your results.

And if you only use one year of data, then a skeptical reader might say well, we don't know the way that the communities are trending and so I don't believe that these communities are necessarily comparable.

So I think the answer to the question is that a research design where you're using where you're comparing communities based on long trends - five or more years - pre-baseline or stronger studies than studies where you're only conditioning on one year or two years of pre-baseline data.

And that will always be the case because if you only have one year of baseline data skeptical readers will come back to you and argue that perhaps the communities weren't really comparable.

Kim Francis: Okay, I'm going to read one more from the Q&A box and then we'll go to the phones. Would it be acceptable to use one level of aggregation for one outcome measure and then a different level of aggregation for a second or a third outcome measure?

Randall Juras: So in other words defining communities differently for different outcome measures, presumably I'm going to assume that whoever asked the question was asking that because different outcomes are available at different levels of aggregation and that I think the answer to that is probably complex but the short answer is yes, you could possibly do that.

You would have to be very careful in presenting and interpreting your results knowing that some of the levels of aggregation are more appropriate than other levels of aggregation so if you have some primary analysis and some

secondary rather exploratory analyses, you probably want your primary analysis to be the one where you're measuring the outcome at the level of aggregation that you really think is the most appropriate.

Kim Francis: Okay, can we see if there are folks waiting to ask a question over the phone?

Coordinator: If you'd like to ask a question, please press star and 1 on your phone. Be sure to unmute your phone and record your name clearly. Again star and 1 with any questions from the phone. One moment while we wait for the first one. One moment, we do have a question. We have a question from (Daniel). Go ahead.

(Daniel): Hi, I was just wondering how you all feel about if I'm going to use a lot of in-school settings as my implementation sites, the idea of using schools not receiving the program but within the same school district as the comparison sites.

Randall Juras: I think that if I'll take a first stab at that answer and let Amy and (Kim) and (Sara) chime-in as needed. I think that it's probably I would be a little bit skeptical if that's a sign.

If you're scaling-up the program to a fairly large degree within the schools that you're implementing in, so if you really feel like you're saturating several schools in a school district, then you would hope I think that kids in the other schools in that school districts might be affected too.

On the other hand it will probably vary on a case-by-case basis. If you're in a school district where the schools are you know it to be the case that the schools are spread-out and/or there's just very little contact between kids at

one school and kids at another school then you might be able to consider a design like that.

But it is a little bit dangerous especially in the context of a grant program where the intent really is to try to change community norms. If you're successful changing community norms then you would really expect that that, you know, probably the other schools in the school district are affected and therefore wouldn't make a very good comparison group.

(Daniel): Thank you.

Coordinator: Our next question comes from (Angela). Go ahead.

(Angela): Hello. I know that Emerging Answers was cited as a place to take a look at the risk in protective factors and in Chapter 3 Doug Kirby on Page 52 indicated that there were more than 500 factors known to increase or decrease the chances of risky sexual behavior and he actually refers to some of the things as antecedents and some of the factors that actually are sort of in existence and that are very difficult to change.

It seems as if some of the more available sources that were referenced today are those things that Dr. Kirby indicated are difficult to change and the ones that these programs are more intended to affect those related to sexual beliefs, values and attitudes are most strongly related to sexual behaviors.

Those the only place we're finding those to be available are in the YRBS and the YRBS is not available at the level that we want it to be so wondering some of those factors that we really don't see a strong correlation that were mentioned today. What advice do you have we do with that?

Kim Francis: This is (Kim), I'll take a first shot at that so just to clarify I think so the variables that were referred to from the Emerging Answers document are not expected to be things that you're changing.

Those are for ideas for you to think about what kinds of co-variates you want to use to establish like baseline equivalence of your treatment and comparison communities so I just wanted to clarify that.

And then the dilemma that you raise I think is the key issue that these evaluations are facing and so on the one hand you're being encouraged to look at these more proximal outcomes but then you're not finding the appropriate data sources for those and you know, I think this is something that we're sort of continuing to discuss with OAH and what the best courses of action is.

And I think that on the one hand you'll want to look at the teen birth rates and then also look at to the extent possible some of these antecedent behaviors as available. I don't know if Amy if you have any insight into this at this point.

Amy Farb: I can jump in. This is why you also have, you know, in the Tier 1Bs we have three components to the evaluations, its performance measures and the implementation study and then the impact study so we would expect that in the key informant interviews and focus groups and all the other things you'll do for your implementation study you might get at some of the outcomes that you are asking about but at the end of the day the Tier 1B strategy is much larger than just these (eddies) that you're implementing.

It's much larger than what might be going on in a single school or a group of schools. There's so many components to it including the community mobilization and, you know, involving the youth, making health referrals, seeing if you can get clinics to more use-friendly, right?

There's so many more components going on in there that we really want to see at a larger level if there's been this cultural shift in health and so we are a teen pregnancy prevention program so we're never going to get out of looking at pregnancies and birth but we do encourage you to look at other outcomes in addition like academic outcomes, other health outcomes, those kinds of things at the rigorous level so hopefully that helps answer some of your question.

(Angela): Thank you.

Coordinator: Our next question comes from (Lauren). Go ahead.

(Lauren): Hi, I was wondering about picking comparison counties. We have a grant in Tennessee that probably covers around 2/3 of this state and in terms of counties and then there's also two other Tier 1B grantees that cover three counties total between them.

When picking a comparison county could you because we're constrained by the number of counties in Tennessee that are currently receiving services under this grant, would we be able to use a county outside of the state that we're implementing in if it is a good fit?

Randall Juras: My one word answer to that is yes. Yes, if there's no, I mean, you want to pick a comparison group that's as local and as focal as you can make it but there is a trade-off between those two things often and the situation that you're describing where there just aren't a lot of places you can go within a state to find a comparison group, you'll have to trade-off a little bit of local for more focal.

If you think this is a very appropriate comparison county in Kentucky or in Alabama or in Georgia or someplace, then yes, I think that, you know, I'd try to stay within the same sort of cultural and geographical regions so you want to stay within the Southeast for example but neighboring states are a great place to draw comparison counties from.

(Lauren): Great, I also have a second question if possible.

Sara Donahue: I might just jump in really quickly, this is (Sara). I would just encourage you to make sure because of the data sources will be then from other states, just really having good conversations about the comparability of the administrative data for the same outcome measure that's captured by two different systems so the state vital statistic department in Tennessee versus that in another state.

Are going to have some nuances around, you know, timing of data availability or processes, quality, those types of things to just getting an understanding of that would be a consideration to make given that you'll need to have the same data for both the treatment and the comparison communities.

Randall Juras: Yes, and if there's any question about whether the data are really truly comparable it would be a good idea to do something like a, you know, comparative interrupted time series or some other design so that you can really control for any baseline differences and subtle baseline differences in the data sources that you're using between the two places.

(Lauren): Can I ask another question?

Randall Juras: I think you could go ahead and ask a second question.

(Lauren): Okay, wonderful. My second question is how about communities that we are looking at for instance in our service area we have 26 counties that received OAH grants last for the first cohort. How would you look at that around ones who have had services over the past five years?

Randall Juras: Sorry, I'm not sure that I caught quite all of that. Would you mind repeating it?

(Lauren): My question is what about some of the counties that we have in our grant that have been receiving services over the past five years, how would you want to account for counties that have been receiving services under OAH for the past five years?

Randall Juras: I think the only way to really convincingly do that is to then find a comparison group that also includes counties that have been receiving OAH services over the last five years and are currently not receiving services.

If no such group exists then it might be the case that those are just very difficult counties to incorporate in the evaluation and if you have a much wider pool of counties in your treatment group, if you have more treatment communities, some of which were not receiving OAH services before then you may want to consider just not including those that have been receiving OAH services in the evaluation.

In other words narrow your evaluation sample to just the counties that are newly receiving services.

(Lauren): Great, thank you.

Coordinator: Our next question comes from (Anne). Go ahead.

(Anne): You indicated that comparison community is one that does not receive OAH services. What about other services that might be funded by state programs or local programs? Is it just the OAH services and specifically Tier 1B bringing to scale that would be the defining intervention or would we also consider other incidences?

Randall Juras: Remember that the counter-factual is the definition of counter-factual is what would have happened in the treatment communities if they had not received your intervention and it's possible that what would have happened in your counties if they hadn't received the intervention if that they would have had some underlying teen pregnancy prevention programming that was for example funded by other sources.

And so your basically there are different research questions that you can answer, right? You can answer the research question about what, you know, what was the impact of the program relative to the business's usual counterfactual which is usually the research question you answer.

Another research question you could think about answering is what's the impact of the intervention compared to what would have happened if there had been no intervention, if there had been no teen pregnancy prevention programming in this case.

I suspect that it will be difficult to find, I mean, I guess it depends on exactly where you are but it will probably be difficult to find communities that have no teen pregnancy prevention programming at all so you do probably want to look for communities that don't have a really robust teen pregnancy prevention curriculum or programming because probably in your communities

there wouldn't have been such a thing and that's why they received OAH funding.

But it will probably be impossible to avoid so I think areas that have some programming so you just need to describe that as, you know, business as usual counterfactual and take it into account when interpreting your results to your impact estimates would be estimates of how effective your intervention was compared to what other services might be available in the community not relative to no programming.

(Anne): Great, that's very helpful, thank you.

Coordinator: No further questions holding.

Kim Francis: All right. Just checking if there are any other questions so some folks were asking if these slides will be made available or be e-mailed out and the answer is yes, they should be available I believe at least within a week both on max.gov and on the OAH Website.

And with that if there's no further questions, if you have specific questions that come up about your design, about, you know, the data sources and comparison groups, please continue to direct those to your project officers and they'll connect you with either someone else at OAH who could answer your question or with one of us at (apt) associates and we'll work on getting your response. All right, well thank you all very much for your time and participation today.

Coordinator: Thank you. That concludes today's conference. You may disconnect at this time. Thank you.

END