

Welcome to the 3rd webinar in UIHI's Evaluation Series.

My name is....

I am a clinical psychologist trained in a Scientist-Practitioner model and I support local health programs in evaluating program outcomes and determining opportunities for growth.

I am excited to be supporting this webinar as it meets my own personal goals for building evaluation capacity for our tribal and urban Indian communities.

However, the topic of analyzing data is very complex and requires a lot of additional training and support.

Therefore, this webinar will provide you with an overview data analysis with a focus on providing you with additional learning resources to support more depth on the topics presented in this webinar.

A resource guide has been developed and will be disseminated to those of you that registered after the call. Please use this guide to continue your learning on this topic of analyzing evaluation data.

Learning Objectives

1. Define data analysis
2. Describe the general process for analyzing qualitative data
3. Describe the general process for analyzing quantitative data
4. Define reliability and validity and describe how it relates to data analysis

Urban Indian Health Institute, a division of the Seattle Indian Health Board



Review objectives

There will be quiz questions presented throughout this webinar. Please take the time to respond to this quiz as it will support your learning through this course.

What does it mean to analyze data?

- Data analysis is the process of condensing raw data into something that is usable such as:
 - Reports
 - Graphs
 - Charts
 - Word Clouds

Urban Indian Health Institute, a division of the Seattle Indian Health Board

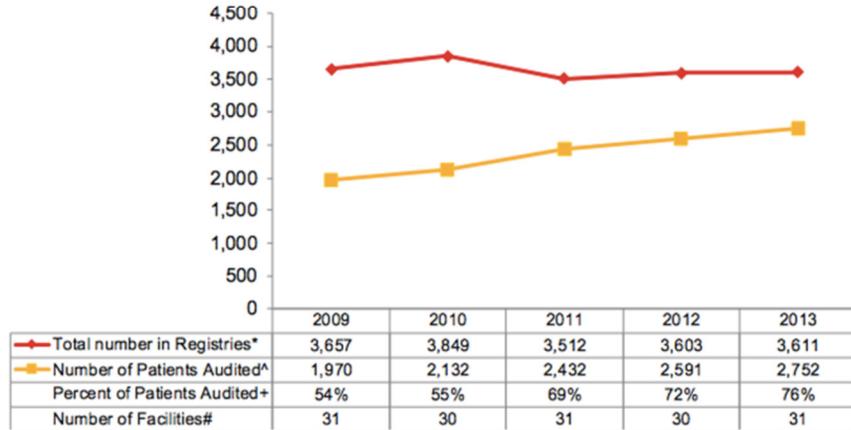


Review slide

Many ways data can be presented.

Graphs and Charts

Patients in Diabetes Registries, Urban Indian Health Organizations, 2009 - 2013

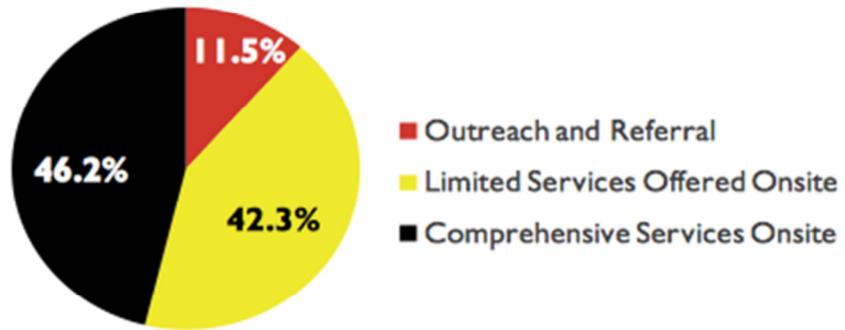


Source: Urban Indian Health Institute

Urban Indian Health Institute, a division of the Seattle Indian Health Board



Pie Chart



Source: Urban Indian Health Institute

Urban Indian Health Institute, a division of the Seattle Indian Health Board



Preparing to Analyze your data

- Describe the problem
- Create research question or hypothesis
- Develop method and tools to collect data
- Collect data using selected and/or developed tools
 - Surveys, Assessments, Structured Interviews/Observations
- Develop a database file to aggregate data
 - SPSS or other statistical program
 - EXCEL
 - Access
 - ATLAS

Urban Indian Health Institute, a division of the Seattle Indian Health Board



The first 2 UIHI webinars in this series supported you in understanding how to develop a logic model that supports you in:

- 1) Describing the problem,
- 2) Create a research question, and
- 3) Develop the methods and tools to collect data.

Once you have collected that data using the methods selected it is time to move that data from individual form to an aggregated form.

There are a number of software tools you could use for aggregating data and I have provided a few examples.

SPSS/EXCEL for aggregating and analyzing quantitative data. There are others like SAS that is a common software used by researchers.

Excel is basic but can get the job done for a local evaluation and is usually already on the computers as most programs use Microsoft.

Access and ATLAS for aggregating qualitative data. Although ACCESS is Microsoft it is not usually included in the standard Microsoft Office package and may have to be purchased.

While you can use some quantitative software tool to organize qualitative data (like Excel), it is not very functional for analyzing that data. While I have used SPSS for aggregating qualitative data, the analysis really must be done by hand as there are limitations to analyzing words and phrases in a quantitative software package.

Aggregating Data

- Combining individual data **by variable** so that it can be analyzed and presented as a whole group (**summary statistics based on those observations**)

Urban Indian Health Institute, a division of the Seattle Indian Health Board



Want to do a term check just to make sure everyone understands what aggregating data means. This is the first step toward analyzing that data is to combine all the individual data by each variable so that it can be analyzed across the whole group of participants.

Once you aggregate the data in one place and space (like an Excel database) then you can run statistics across all those data points (note that data points could also be observations and other qualitative information)

Variables

- Variable: Characteristic, number **or** quantity that can be measured or counted.
 - Categorical variables: two or more categories
 - Gender; location of residence
 - Ordinal variables: in an sequential order
 - Grade in school; satisfaction rating
 - Interval variable: like ordinal but equal distance between data points
 - Temperature; Herth Hope Index
 - Ratio variable: like interval but true zero.
 - Income; Blood Sugar

Urban Indian Health Institute, a division of the Seattle Indian Health Board



Now you know what aggregating data is, I want to make sure everyone understands what it means to aggregate data by variables.

What is a variable?

Review Slide **Variable**.

There are several types of variables and it is important to know what type of variable you have. This will help you determine what type of statistics you can run on that data for the purpose of analysis. First I will review the types of Variables and later I will help you understand what types of statistics you could run on each type.

Review Types **Categorical**

Independent vs. Dependent Variable

- **Independent:** variable that you or nature manipulates
 - Ex: Participation in a new treatment program
- **Dependent:** the effects of the independent variable
 - Ex: Participation in new treatment effects a persons outlook on life (item from the Herth Hope Index)

Urban Indian Health Institute, a division of the Seattle Indian Health Board



In addition to understanding the type of variable, you also need to understand if that variable is an independent or dependent variable.

Review Independent and Dependent variable Slide

So, an independent variable is one that you or nature controls.

You can control such things as whether or not someone receives a new treatment.

An example of an independent variable that is controlled by nature could be something like depression. You can't control if someone is depressed or not, but nature can. When nature is controlling a variable researchers refer to this as a quasi-experimental study because there are some things that you as the researcher can't control. Therefore it is not considered true experimental unless you are manipulating some variable yourself.

A dependent variable then, is the effect (and/or outcomes) of that independent variable, like the data that you collect from the participants that are in the program.

For example, let's assume we want to see if a new treatment helps people with depression that adds Yoga and Mindfulness to the standard treatment for depression. We can randomly select and assign individuals with depression to participate in one of two groups. One group used the additional pieces (Yoga and Mindfulness in addition to CBT). The other group only provides CBT but does not use Yoga and Mindfulness. Let's say that we give all participants, regardless of the intervention group they are in the Herth Hope Index before and after the intervention. When we give the Herth Hope Index before the intervention, this is called collecting baseline data. In other words, where do people start treatment in terms of their level of hope. When we give the Herth Hope Index after the intervention (called a post-test or assessment) we can examine the changes from baseline to after treatment on levels of hope using statistics. We will talk more about these statistics in the slides to come but I wanted to give you a real example of a potential project that highlights what is an independent vs. a dependent variable.

It is also important to note that, in this example provided, we are implementing a true experimental study because we use

- 1) Random selection of participants and assignment into two groups
- 2) We are manipulating an independent variable, which is the assignment of participants into two different treatment programs.

So while nature is controlling who has depression to be a part of the study (quasi-experimental) when you as the investigator create some variable within that to be manipulated (like which treatment people receive), then it becomes a more true experimental design.

(like whether or not a person has depression)

How we analyze data

- Two types of data analysis:
 - Qualitative
 - A way of capturing and recording attitudes and feelings
 - “This webinar is great!” –Webinar participant
 - Quantitative
 - Data that is expressed in numbers
 - 9 out of 10 people agree this webinar is great

Urban Indian Health Institute, a division of the Seattle Indian Health Board



As previously described there are two types of data that we can analyze: qual and quant.
Review Slides

We will go into much more detail in the next few slides about these two types of analysis

Qualitative Data



Urban Indian Health Institute, a division of the Seattle Indian Health Board



We will start with qualitative analysis.

But before we do I want to bring some acknowledgment to one of the community projects I have had the real honor to work on. The photo you see came from the Fresno American Indian Health Project, one of the urban Indian health programs. What you see in the photo is a real session of coding qualitative data and the participants you see in the photo are local community members. They are parents, uncles, aunts, grandparents and none of them had formalized training in evaluation before we taught them how to code qualitative data.

I want to bring this to your attention to acknowledge the critical role that you communities can play in supporting local evaluation. Not only can they bring a human resource but they can also bring improved reliability and validity of the data as there are challenges in coding qualitative data if you are not familiar with the local community. In this particular community we trained community members to code the data and they spend literally days working on as local volunteers.

As a researcher from the outside I recognize that I have my own lens to see the world. I could miss some critical piece of information or misrepresent something because I don't know the local language and context. Community members can support coding and interpreting qualitative data in way that is in line with local language, norms, and believes.

You can learn more about this by reviewing some of the archived presentations that I have supported on Community-Based Participatory Research through the SAMHSA (SPELL OUT) Tribal Training and TA center, Evaluation Learning Collaborative.

Qualitative Data

- Data expressed in descriptive terms (words, phrases, photos, etc.).
- Examples of sources for Qualitative data:
 - Surveys (open-ended questions/fill-in responses)
 - Interviews
 - Focus Groups
 - Photo Voice
 - Project Activity Documents

Urban Indian Health Institute, a division of the Seattle Indian Health Board



<http://archive.learnhigher.ac.uk/analysethis/main/qualitative.html>

Qualitative Analysis

- Look for recurring patterns or themes
- Code of categories and sub-categories
- Compare codes, looking for consistencies, differences, and patterns
- Codes can be based on:
 - Themes, topics
 - Ideas, Concepts
 - Terms, Phrases
 - Keywords

Urban Indian Health Institute, a division of the Seattle Indian Health Board



Analyzing qualitative data requires an examination of themes or patterns in the data. There are a number of theoretical approaches to analyzing qualitative data but all methods include coding of data into various categories and sub-categories. Codes can be based on....REVIEW

Qualitative Analysis Example

- Process Evaluation: Assesses the program process, implementation, and/or service delivery
 - Effectiveness?
 - Efficiency?
 - Continued Quality Improvement?

Urban Indian Health Institute, a division of the Seattle Indian Health Board



Let's walk through an example of a typical project that would use qualitative data.

REVIEW SLIDE

Any Project Report Form

Date/Time: 2-2-15	Event Name: Coding/Evaluation Subcommittee Team Meeting
Who Attended? Jami Evelyn Seith Sam	How Many Attended? 4
What Happened?	This was the final meeting of the evaluation subcommittee to code the data from the last completed focus group. The team spent approximately 4 hours coding transcripts
Barriers to Overcome:	The time commitment has been challenging with conflicting schedules
Next Steps:	Converting data coding table into Word or Excel and manually copying and pasting transcripts into themes and subthemes. Send off to Jami as soon as possible
Notes/Observations:	My observation is the group really owns this data. They feel proud that they have completed all the coding in a little over 2 months. The coding team recognizes their hard work in organizing the thoughts of their peers to support the local needs assessment
Submitted By:	Sam

Urban Indian Health Institute, a division of the Seattle Indian Health Board



My typical approach to process evaluations is to have program staff documenting the work using simple project activity report forms.

This is an example of one of these forms for a project in which I was supporting a process evaluation.

As you can see, the report form documents a number of data points, though mostly qualitative. For a project, staff may generate hundreds of report forms in one year that must be aggregated and analyzed to determine the effectiveness and efficiency of the program, in addition to using data for quality improvement purposes.

Each of these activity reports would be coded for themes and subthemes and would be organized in some database so that all the data can be examined together.

Meeting Date	Event Name	Summary	Theme	Sub-Theme
10/14/14	Community Council	Focus group updates, survey updates	Community updates	Data quality and evaluation capacity
11/19/14	Community Council	Updates on youth and adult surveys, update on recruiting efforts for focus groups	Community updates	Community and areas understanding
11/30/14	Youth Council	Recruitment night for youth council	Development of program	Community participation
12/17/14	Community Council	Survey completed, Focus group update, youth survey update	Community updates	Community participation
1/28/15	Eval Subcommittee	Training for Coding Team; IRB; coding process/documentation	Workforce development	Community participation
2/2/15	Eval Subcommittee	Data and evaluation /Coding training	Workforce development	Community participation
2/5/15	TA Site Visit	Family-Driven Training	Workforce development	
2/8/15	National Webinar	LGBT training	Workforce development	
3/22/15	Youth Council	advertise youth survey	Community updates	
3/26/15	Youth Council	Program development for youth focus group	Group & Community Und	Data quality and evaluation capacity
4/8/15	County Training	Patient-Centered Health Home Initiative webinar	Workforce Development	
4/8/15	County Training	Suicide prevention training	Workforce Development	
4/17/15	Parent Training	Nurturing Parent Program Class	Program	Cultural Inclusion
4/29/15	Partnership Building	Meeting with Veterans Affairs	Collaboration	Information sharing

Urban Indian Health Institute, a division of the Seattle Indian Health Board



Now this slide shows you how we aggregated data using excel. Again, excel has limitations in analyzing qualitative data but can serve as a tool for organizing that data. As long as our themes and subthemes are consistent we can actually count the number of coded themes by type.

In this example, you see that the report form submitted by Sam highlighted in red has been coded into the theme Workforce Development and a sub-theme of Community Participation. This helps us understand what the impact of the activity.

Qualitative Data Codes in Aggregate

Staffing	10
Workforce Development	69
Development of Program	52
Outreach	19
Community Updates	44
Barriers	7
Community Participation	38
Information Sharing	19
Collaboration	38
Information Sharing	33
Cohesion	8
Data Quality and Evaluation	32
Community Inclusion	21

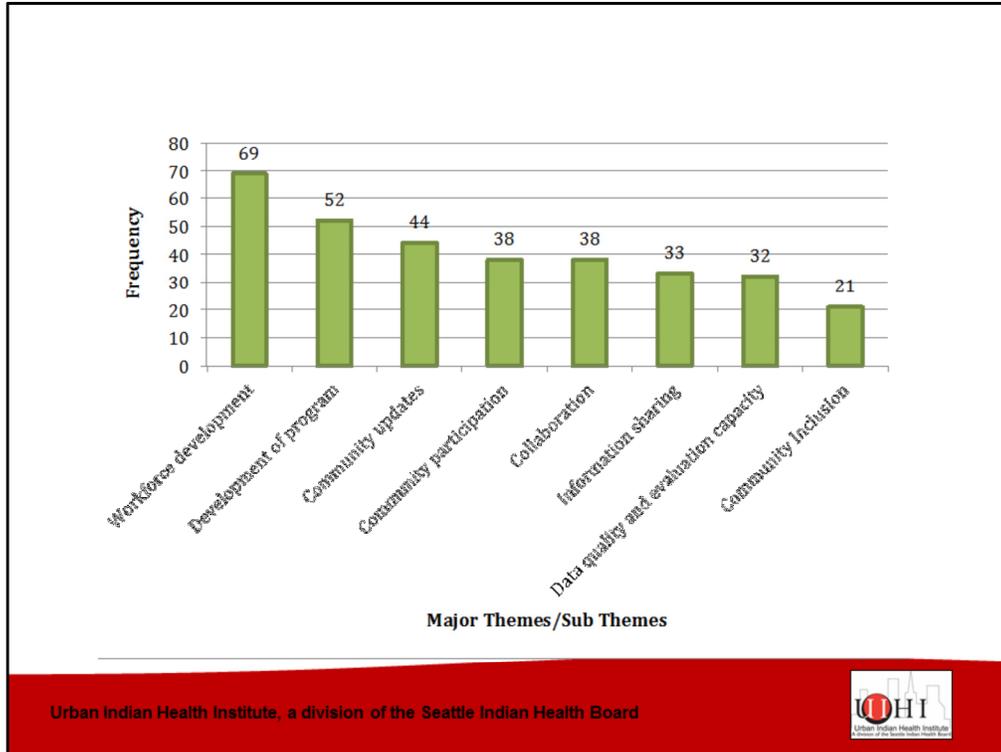
Urban Indian Health Institute, a division of the Seattle Indian Health Board



This slide then shows how to take this data one step further and determine the frequency of the various themes and subthemes.

It is important to note that some researchers chose not to turn qualitative data into anything quantitative as you see in this slide, and there is disagreement about this method.

However, I typically get some frequency count of the various themes to get some idea about what themes are rising to the surface most frequently in the program. Then I will prepare a written narrative that describes in more detail the range of activities that occurred around that theme, making sure to highlight any of the program successes, challenges, barriers, and/or areas needed for growth and quality improvement.



This slide shows how that data can be turned into a graph form.

Limitations of Qualitative Analysis

- Usually fewer people studied
- Less easy to generalize
- Difficult to make systematic comparisons
- Dependent on the skills of the researcher and coders

Urban Indian Health Institute, a division of the Seattle Indian Health Board



Review limitations.

Quantitative Data



Urban Indian Health Institute, a division of the Seattle Indian Health Board



Now we will move on to quantitative data but I want to direct you to a photo taken from one of the programs I work with that found creative and youth friendly ways of capturing quantitative data using glass beads instead of a likert scale.

Some of the staff from the Friendship House of American Indians in SF used skittles in a similar fashion for a youth conference.

I really encourage programs to engage youth in developing any assessment methods and protocols to ensure these are youth-friendly.

Quantitative data

- Data expressed in numbers
- Sources of Quantitative data include:
 - Questionnaires/Surveys
 - Number of participants/individuals with experiences
 - Standardized assessment tools
 - Social media data
 - Dollars/Cents

Urban Indian Health Institute, a division of the Seattle Indian Health Board



Review slide

Variable Type By Stats

Stats	Nominal	Ordinal	Interval	Ratio
Frequencies	YES	YES	YES	YES
Percentiles	NO	YES	YES	YES
Mean/Standard Deviation; SEM	NO	NO	YES	YES
Add/Subtract	NO	NO	YES	YES
Division/Multiplication	NO	NO	NO	YES
	Categorical	Categorical OR Continuous	Categorical OR Continuous	Continuous

GraphPad Software, Inc. 2014 (modified)

Urban Indian Health Institute, a division of the Seattle Indian Health Board



Remember back to the slide on the different types of variables?

This slide helps you see what types of statistics can be used on the different variables.

Categorical variables (also called nominal): two or more categories

Example of GENDER

Review Chart

Ordinal variables: in an sequential order

EXAMPLE: Grade in school

Review Chart

Interval variable: like ordinal but equal distance between data points

EXAMPLE Herth Hope Index

Review Chart

Ratio variable: like interval but true zero.

EXAMPLE: Blood Sugar

Review Chart

Descriptive Statistics

- Mean: the sum of the values divided by the total numbers in the set (**average**)
 - Ex: 1, 2, 3, 4, 5 = $15 / 5 = 3$
- Median: the middle value that separates the higher half and lower half
 - Ex: 1, 2, **3**, 4, 5
- Mode: the most frequent value in a data set
 - Ex: 1, **2, 2, 2**, 3, 4, 5

Urban Indian Health Institute, a division of the Seattle Indian Health Board



There are two major type of statistics that you could use. The first is descriptive statistics.

Descriptive statistics are used to describe the basic features of the data in a study. They provide simple summaries about the sample and the measures. Together with simple graphics analysis, they form the basis of virtually every quantitative analysis of data.

Descriptive Statistics include
REVIEW SLIDE

Descriptive Statistics Continued

- Variance: the average of the squared difference from the mean (square root of variance=Std Dev)
- Std Deviation: the measure of how spread out the numbers are (the amount of variation in scores)
- Range: the difference between the largest and smallest values
 - Ex: 1, 2, 3, 4, 5 Range= 4

Urban Indian Health Institute, a division of the Seattle Indian Health Board



Review Slide

Inferential Statistics

		Dependent Variable	
		Categorical	Continuous
Independent Variable	Categorical	Chi Square	T-Test/Higher Order (3 groups +)
	Continuous	Higher Order	Higher Order

Example 1:

Independent Variable= Intervention
 Dependent Variable= Gender
 Statistic=Chi Square

Example 2:

Independent Variable= Intervention
 Dependent Variable= Herth Hope Index
 Statistic= T-Test

Urban Indian Health Institute, a division of the Seattle Indian Health Board



Inferential Statistics are the other type of statistic that you would use in evaluation. Inferential is in its name. It means statistics that help you draw inferences or conclusions about the data, such as treatment A is better than Treatment B.

There is one other way to identify variables.

We have talked about the different types of variables including (nominal, ordinal, interval and ratio).

We have also reviewed the difference between independent and dependent variables.

But variables are also identified as categorical or continuous. This is an important distinction for determining what type of inferential statistics you could use.

Categorical means something that can clearly be put into a category, like whether or not someone is depressed or not.

Continuous means something that is not in a category, but rather, on a continuum, like the Beck Depression Inventory or the Herth Hope Index. These measures tell you the range of where someone falls on each of these constructs. In other words, continuous doesn't tell you if they meet criteria for depression but does tell you how severe the depressive symptoms may be.

As you can see from this slide, the designation of a variable as categorical or continuous will help you determine what types of inferential statistics you can use.

In Example One let's say you have an intervention that serves as the independent variable and you want to know if males and females experience that intervention differently. In this case, you have two categorical variables and the most appropriate statistic would be a Chi Square.

In Example Two, let's say you have a new intervention (such as adding Yoga and Mindfulness to CBT) as your independent variable and you want to know if people that receive this new intervention do better in increasing their hope for the future than do those that receive treatment as usual, which is CBT alone. If you look at the slide you will see that having one categorical variable (which is the new intervention as compared to the treatment as usual) and one continuous variable (in this case the measure of Hope using the Herth Hope Index), then you would use a t-test.

REVIEW EXAMPLE

Chi Square

- Used to determine if two categorical variables differ from one another
 - Example: Do males/females differ in their choice of treatment (home-based vs. clinic-based)?

	Treatment Type		ROW TOTAL
	Home Based	Clinic Based	
Female	250	150	400
Male	100	300	400
COLUMN TOTAL	350	450	800

Urban Indian Health Institute, a division of the Seattle Indian Health Board



So now that we have reviewed the types of inferential statistics you will likely use in program evaluation, I will provide a little bit more information about what these statistics actually do.

A Chi Square
REVIEW SLIDE

By eyeballing this it looks like there might be differences between males and females on their treatment preferences but we will run a Chi Square to determine if these groups are indeed statistically significant.

I will review significance and related terms later in this presentations so hold tight.

T-Test

- Can be determined if TWO sets of data are significantly different from each other
 - Independent t-test (different groups)
 - Ex: wait-list control and treatment group
 - Dependent t-test (same group)
 - pre/post or matched pairs

Urban Indian Health Institute, a division of the Seattle Indian Health Board



The t-test then is REVIEW SLIDE

There are many different types of T-test you could run but they will generally fall into two categories for local evaluation projects.

REVIEW INDEP and DEP T-test.

There are resources in the resource guide that will help you learn more about calculating t-tests.

Higher Order Statistics

- ANOVA
 - Compares categorical independent variable with 3 or more groups on continuous dependent variable
 - Example: Compare 3 types of treatment by hope
- Regression
 - Analyzes the relationships between a continuous independent variable and an continuous dependent variable
 - Example: Relation between age and hope

Urban Indian Health Institute, a division of the Seattle Indian Health Board



There are higher order statistics that researchers use but these require very advanced training in research design and statistics. I include a brief review of a couple of common statistics just so you have a greater understanding of the depth of this topic.

Review slide.

Quantitative Analysis Example

- Pre/Post assessment for each youth
- Culture Camp
– Herth Hope Index



Urban Indian Health Institute, a division of the Seattle Indian Health Board



Now let's go through an example of a typical quantitative analysis project that you might have for your local evaluation.

Let's assume that you have a pre-post assessment that that you are giving to youth before and after a culture camp and you want to see if the culture camp improves the hope for the future.

Study No. 101

HERTH HOPE INDEX

Listed below are a number of statements. Read each statement and place an (X) in the box that describes how much you agree with that statement right now.

	1 Strongly Disagree	2 Disagree	3 Agree	4 Strongly Agree	Score
1. I have a positive outlook toward life.				X	4
2. I have short and/or long range goals.			X		3
3. I feel all alone.			X		2
4. I can see possibilities during difficult times.		X			2
5. I have a spiritual belief that gives me comfort.			X		3
6. I feel scared about my future.			X		2
7. I can remember happy/joyful times.			X		3
8. I have deep inner strength.			X		3
9. I am able to give and receive caring/love.			X		3
10. I have a sense of direction.			X		3
11. I believe that each day has potential.			X		3
12. I feel my life has value and worth.			X		3

© 1989 Kaye Herth
1999 items 2 & 4 reworded

Urban Indian Health Institute, a division of the Seattle Indian Health Board



Here is an example of a completed Herth Hope Index for an individual.

This data must be scored and entered into a database so that the statistics can be run.

Note that it is important to check and double check scoring. Many tools use reverse items. The items circled in red are the reverse items for the Herth Hope Index and can see that we scored them backwards (in other words, a youth that agreed on item 3 (I feel all alone) would be scored as 2 rather than three as this item is reversed and a larger number rating is actually worse on this item, rather than better as with the more positive oriented questions.

Developing a database

1. Describe your gender: Male _____ Female _____
Other: please explain
2. How old are you?

Herth Hope Index	Strongly Disagree	Disagree	Agree	Strongly Agree
1. I have a positive outlook toward life				
2. I have short and/or long range goals				
3. I feel all alone				

Kaye Herth 1989, http://www.npcrc.org/files/news/herth_hope_index.pdf

Urban Indian Health Institute, a division of the Seattle Indian Health Board



Now lets say you have 10 youth that completed the Herth Hope Index. You will also need to make sure you have the important demographics you need, like age, gender, or other important descriptive variables. This will help you learn more about the data and determine if any of these other variables make a difference.

For example, you might find that the culture camp does well for younger youth but not for older youth or visa versa.

You have to make sure you ask these questions and can match the demographic data up with the outcome data, in this case, it is the Herth Hope Index as the dependent variable.

Developing a database

Participant ID	Gender	Age	Outlook	Goals	Alone
101	1	16	4	3	2
102	2	19	3	4	1
103	2	14	4	3	2
104	1	12	2	3	3
105	2	15	3	4	2

Urban Indian Health Institute, a division of the Seattle Indian Health Board



Here is a snapshot of how we would take the data and enter it into Excel or some other database.

As you can see, each participant is given a unique ID and all of their data is entered across that row.

All the categorical variables must be coded into numbers. In this sample database, 1 is used for female and 2 is used for male.

For participant 101, it is a female, age 16 and all of her data is entered across the row for the Herth Hope Index items.

Analyzing the data

- T-Test for Dependent Samples
 - Same youth take Herth Hope Index pre/post camp
- Is it statistically significant?
 - P-Value: Probability of making Type I errors (i.e., say there is something significant when there is not)
 - Science Norm p is less than or equal to .05 (95% chance you are correct)
 - Practice Norm- depends

Urban Indian Health Institute, a division of the Seattle Indian Health Board



So now how would we analyze this data once we get it all entered into the database?

In this example....
Review Slide.

Magnitude of Effect

- Clinical Significance vs. Statistical Significance
 - Effect Size: strength of the effect
 - Large number= great effect
 - 0.80
 - Moderate number=moderate effect
 - .050
 - Small number = limited effect
 - 0.20

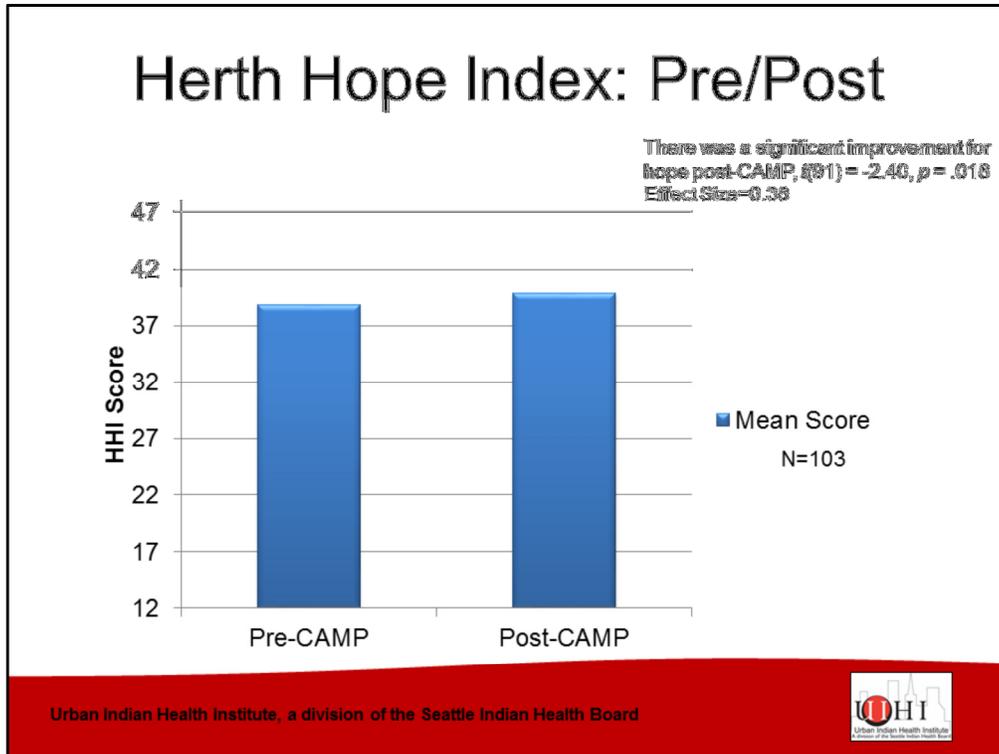
Urban Indian Health Institute, a division of the Seattle Indian Health Board



Now, just because something is statistically significant doesn't necessary mean it is clinical significance.

REVIEW SLIDE.

There are online resources to support you in calculating effect sizes and some stat packages, like SPSS, have mechanisms for this too.



There was a significant improvement on youth's overall feeling of hope post GONA but we have to ask ourselves, is this difference from pre to post clinically meaningful. If youth are only improving by a few data points, it might not be.

Limitations of Quantitative

- Numerical descriptions rather than detailed narrative
- Preset answers may not reflect how people really feel
- Structural bias

Urban Indian Health Institute, a division of the Seattle Indian Health Board



Review slides

“the development of standard questions by researchers can lead to 'structural' bias and false representation, where the data actually reflects the view of them instead of the participating subject.”

Reliability and Validity

- **VALIDITY**

- The degree to which results can be extrapolated outside of the study or research
 - Evaluation Examples: For the question in the Herth Hope Index "I have a positive outlook on life," you would expect individuals, regardless of where they are at who have a positive outlook on life to respond with a strongly agree or agree.

- **RELIABILITY**

- Repeating a piece of research in order to establish the reliability of its findings (ex: a ruler is a reliable measure of length)
 - Evaluation Example: For qualitative research projects, the researcher uses a standard procedure and script to interview participants

Reliability and Validity



Reliable
Not Valid



Valid
Not Reliable



Neither Reliable
Nor Valid



Both Reliable
And Valid

Urban Indian Health Institute, a division of the Seattle Indian Health Board



1st Target: This is reliable and not valid. The participant hits the same spot on the target each time but does not hit the bulls eye

2nd Target: Valid but not reliable. If you average all the shots, it would average to the bull's eye so it would be a valid measure.

3rd Target: Neither valid nor reliable. The shots are random and are not at the target. Neither reliable to valid.

4th Target: Both reliable and valid. It is valid because it hits the middle of the target and reliable because it consistently hits the near the same spot every time.

Threats to Reliability/Validity in Data Analysis

- Unit of Analysis not clearly defined or delineated (Reliability)
- Method of analyzing data doesn't match the data (Validity)
- Violating the assumptions of the statistical test used (Validity)
- Disagreement among coders about what data means (Reliability)
- Statistical Regression to the Mean (Validity)
- **Data Entry Errors** (Validity)

Urban Indian Health Institute, a division of the Seattle Indian Health Board



Unit of Analysis not clearly defined

Can you clearly tell the difference between variables or are there areas of overlap within the variables being collected?

Review Slide

Statistical regression to the mean:

Regression toward the mean involves outcomes that are at least partly due to chance. Over time, people have a tendency to regress to the mean or average on any variable and therefore can create a threat in data validity. In other words, It may look like nothing really happened in the intervention because the participants regressed to the mean. Usually creating a non-significant finding when their actually might be something significant in the data.

Data Cleaning

- Process of detecting or correcting inaccurate data
- Look for min and max values in responses
 - (if responses are 1 through 4, 44 shouldn't be a response)
- Look for any missing values and code as something different (ex 999)
- Look for spelling errors
- 20% random check for errors (consider rechecking all if high error rate exists)

Urban Indian Health Institute, a division of the Seattle Indian Health Board



Review slides

A word about Standard Deviation

- Standard Deviation
 - Large number= less reliable
 - Mean=38, Standard Deviation=15
 - Small Number= more reliable
 - Mean=38, Standard Deviation=5

Urban Indian Health Institute, a division of the Seattle Indian Health Board



Review slide

Questions and Contact information

- Please ask any questions or if you have questions later please contact:
- Contact Information:
 - Jami Bartgis, CEO
 - email: drbartgis@onfireassociates.org
 - Alex Denning, Consultant
 - email: alex.d.denning@gmail.com
 - Website: www.onfireassociates.org



Urban Indian Health Institute
Seattle Indian Health Board
P.O. Box 3364
Seattle, WA 98114

Phone: (206) 812-3030
Fax: (206) 812-3044

Email: info@uihi.org

Website: www.uihi.org